



# Red Hat Enterprise Linux (RHEL) Driver Release Notes

RHEL 8.4

# Table of Contents

- 1 **Overview**..... 4
  - Supported HCAs Firmware Versions..... 5
  - SR-IOV Support..... 5
  - RoCE Support ..... 6
  - VXLAN Support ..... 6
  - DPDK Support ..... 6
  - Open vSwitch Hardware Offloads Support..... 6
- 2 **Changes and New Features** ..... 7
- 3 **Certifications** ..... 8
  - RHEL NIC Qualification..... 8
- 4 **Known Inbox-Related Issues**..... 9

## List of Tables

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

# 1 Overview

These are the release notes of Red Hat Enterprise Linux (RHEL) Inbox Driver. This document provides instructions on drivers for NVIDIA® Mellanox® ConnectX® adapter cards used in a RHEL Inbox Driver environment.

This version supports the following uplinks to servers.

Table 1: Supported Uplinks to Servers

HCA	Uplink Speed	Supported Driver
BlueField®2	<ul style="list-style-type: none"> <li>Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE, 100GbE and 200GbE</li> </ul>	mlx5_core (includes the ETH functionality as well), mlx5_ib
ConnectX-6 Lx	<ul style="list-style-type: none"> <li>Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE, 100GbE and 200GbE</li> </ul>	mlx5_core (includes the ETH functionality as well), mlx5_ib
ConnectX-6 Dx	<ul style="list-style-type: none"> <li>Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE, 100GbE and 200GbE</li> </ul>	mlx5_core (includes the ETH functionality as well), mlx5_ib
ConnectX-6	<ul style="list-style-type: none"> <li>InfiniBand: SDR, EDR, HDR</li> <li>Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE and 100GbE</li> </ul>	mlx5_core (includes the ETH functionality as well), mlx5_ib
BlueField <sup>a</sup>	<ul style="list-style-type: none"> <li>Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE, and 100GbE</li> </ul>	mlx5_core (includes the ETH functionality as well)
Innova™ IPsec EN	<ul style="list-style-type: none"> <li>Ethernet: 10GbE, 40GbE</li> </ul>	mlx5_core (includes the ETH functionality as well)
ConnectX-5	<ul style="list-style-type: none"> <li>InfiniBand: SDR, QDR, FDR, FDR10, EDR</li> <li>Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE, 56GbE<sup>b</sup>, and 100GbE</li> </ul>	mlx5_core (includes the ETH functionality as well), mlx5_ib
ConnectX-4	<ul style="list-style-type: none"> <li>InfiniBand: SDR, QDR, FDR, FDR10, EDR</li> <li>Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE, 56GbE<sup>b</sup>, and 100GbE</li> </ul>	mlx5_core (includes the ETH functionality as well), mlx5_ib
ConnectX-4 Lx	<ul style="list-style-type: none"> <li>Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, and 50GbE</li> </ul>	mlx5_core (includes the ETH functionality as well)
ConnectX-3/ ConnectX-3 Pro	<ul style="list-style-type: none"> <li>InfiniBand: SDR, QDR, FDR10, FDR</li> <li>Ethernet: 10GbE, 40GbE and 56GbE<sup>b</sup></li> </ul>	mlx4_core, mlx4_en, mlx4_ib
Connect-IB®	<ul style="list-style-type: none"> <li>InfiniBand: SDR, QDR, FDR10, FDR</li> </ul>	mlx5_core, mlx5_ib

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

b. 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.

# Supported HCAs Firmware Versions

Red Hat Enterprise Linux (RHEL) 8.3 driver supports the following Mellanox network adapter cards firmware versions:

Table 2: Supported HCAs Firmware Versions

HCA	Recommended Firmware Version
BlueField-2	24.30.1004
ConnectX-6 Lx	26.30.1004
ConnectX-6 Dx	22.30.1004
ConnectX-6	20.30.1004
BlueField (Technical Preview)	18.30.1004
ConnectX-5	16.30.1004
ConnectX-4 Lx	14.30.1004
ConnectX-4	12.28.2006
ConnectX-3/ConnectX-3 Pro	2.42.5000
Innova IPsec EN	14.22.1002
Connect-IB	10.16.1002

# 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 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 (# ib sm virt enable). <b>Note:</b> This capability is not tested over Inbox environment and considered Tech Preview.

## RoCE Support

Table 4: RoCE Support

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

## VXLAN Support

Table 5: VXLAN Support

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

## DPDK Support

Table 6: DPDK Support

Driver	Support
mlx4	Mellanox PMD is enabled by default.
mlx5	Mellanox PMD is enabled by default.

## 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

Component	Feature/Change	Description
Driver mlx5	Hairpin Support in Switch Mode	<b>[ConnectX-5 and above]</b> Added support for receiving tunneled traffic from the uplink port, after which it can be decapsulated and sent back to the uplink port and possibly encapsulated with new tunnel information.
	RoCEv2 Flow Label and UDP Source Port Definition	<b>[ConnectX-4 and above]</b> This feature provides flow label and UDP source port definitions in RoCEv2. These fields are used to create entropy for network routes (ECMP), load balancers, and 802.3ad link aggregation switching that are not aware of RoCE headers.
	MPLS-over-UDP Hardware Offload Support	<b>[ConnectX-5 and above]</b> Added support for encapsulation/decapsulation hardware offload of IPv4 traffic over MPLS-over-UDP. This can be used in networks with MPLS routers to achieve more efficient routing.
	IPsec Crypto Offload	<b>[ConnectX-6 Lx and above]</b> Added support for IPsec crypto offload. IPsec offload is a set of features aimed at reducing the CPU overhead of using IPsec. It targets the heavy crypto (encryption/decryption) operation, offloading them to the hardware. The NIC handles all encryption, decryption, and authentication, leaving the rest (replay, encapsulation, and decapsulation) to the software IPsec implementation. This approach has the benefits of increased flexibility in software, at the cost of some performance overhead.
	TLS RX Hardware Offload	<b>[ConnectX-6 Dx and above]</b> Added support for hardware offload decryption of TLS traffic over crypto-enabled ConnectX-6 Dx NICs and above.
	Add messages when VF-LAG fails to start	Added messages to provide human-readable feedback on configuration issues causing VF LAG to fail.
	TC dump and OVS revalidation optimization	Optimized TC dump and OVS revalidation performance by introducing a new mode of TC filter dump operation named "terse dump" mode. In this mode, only parameters necessary to identify the filter (handle, action cookie, and so forth) and data that can change during the filter lifecycle (filter flags, action stats, and so forth) are preserved in the dump output while everything else is omitted.
mlx5	General driver update	Driver base Upstream Kernel v5.9
mlx4	General driver update	Driver base Upstream Kernel v5.9

libpcap	Support sniffing offloaded (RDMA) traffic with tcpdump	Added support in libpcap to sniff user-space traffic (RoCE and raw Ethernet) using tcpdump.
rdma-core		Updated to version rdma-core-32.0-4.el8
mstflint		Updated to version mstflint-4.15.0-1.el8
VMA		Updated to version libvma-9.2.2-2.el8
UCX		Updated to version ucx- 1.9.0-1.el8

---

## 3 Certifications

### RHEL NIC Qualification

The following RHEL and NIC combinations successfully passed RHEL NIC qualification covering OVS functional, OVS non-offload, OVS-offload, and OVS-DPDK:

Adapters	RHEL Versions
ConnectX-4 Lx, ConnectX-5	RHEL 8.0 - 8.x
ConnectX-5 Ex	RHEL 8.3 - 8.x
ConnectX-6, ConnectX-6 Dx	RHEL 8.3 - 8.x



**Note:** More details can be found at the Red Hat page: “Network Adapter Fast Datapath Feature Support Matrix” at [access.redhat.com/articles/3538141](https://access.redhat.com/articles/3538141)



## 4 Known Inbox-Related Issues

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

Internal Ref.	Bugzilla Ref.	Description
2482177	-	<b>Description:</b> RDMA device name for VFs may change after resetting all VFs at once.
		<b>Workaround:</b> Either reset interfaces one by one with a delay in between, or use a network interface naming scheme with predictable interface names, such as NAME_PCI or NAME_GUID: copy /lib/udev/rules.d/60-rdma-persistent-naming.rules to /etc/udev/rules.d/ and edit the last line accordingly. Note that this will change interface names.
-	-	<b>Description:</b> RPM package <b>kernel-modules-extra</b> is required for supporting various OVS Hardware Offloads.  To use OVS Hardware Offloads, make sure to install the <b>kernel-modules-extra</b> RPM package which provides various kernel modules that are required for supporting this functionality.
2345747	1890261	<b>Description:</b> RHEL installer fails to start when InfiniBand network interfaces are configured using installer boot options.
		<b>Workaround:</b> Create a new installation media including the updated Anaconda and NetworkManager packages, using the Lorax tool. For more information on how to do so, please see <a href="#">here</a> .
		<b>Keywords:</b> PXE, IPoIB, InfiniBand
	1816660	<b>Description:</b> When the NUM_OF_VFS parameter configured in the Firmware (using the mstconfig tool) is higher than 64, VF LAG mode will not be supported while deploying OVS offload.
		<b>Workaround:</b> N/A
		<b>Keywords:</b> ConnectX-5, VF LAG, ASAP <sup>2</sup> , SwitchDev
	1816660	<b>Description:</b> An internal firmware error occurs either when attempting to disable single-root input/output virtualization or when unbinding PF using a function (such as ifdown and ip link) under the following condition:  VF LAG mode in an OVS offload deployment, where at least one VF of any PF is still bound on the host or attached to a VM.
		<b>Workaround:</b> Unbind or detach VFs before you perform these actions as follows.  1. Shutdown and detach any VMs.  2. Remove VF LAG bond interface from OVS.

Internal Ref.	Bugzilla Ref.	Description
		<p>3. Unbind VFs, perform for each configured VF:  <pre># echo &lt;VF PCIe BDF&gt; &gt; /sys/bus/pci/drivers/mlx5_core/unbind</pre></p> <p>4. Disable SR-IOV, perform for each PF:  <pre># echo 0 &gt; /sys/class/net/&lt;PF&gt;/device/sriov_numvfs</pre></p> <p><b>Keywords:</b> ConnectX-5, VF LAG, ASAP2, SwitchDev</p>
1284047	-	<p><b>Description:</b> Bandwidth degradations due to PTI (Page Table Isolation) in Intel's CPU security fix.</p> <p><b>Workaround:</b> 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 grub.conf.</p> <p><b>Keywords:</b> Performance</p>
1610281	-	<p><b>Description:</b> Setting speed to 56GbE on ConnectX-4 causes firmware syndrome (0x1a303e).</p> <p><b>Workaround:</b> N/A</p> <p><b>Keywords:</b> ConnectX-4, syndrome</p>
1609804	-	<p><b>Description:</b> Kernel panic during MTU change under stress traffic.</p> <p><b>Workaround:</b> N/A</p> <p><b>Keywords:</b> Panic, MTU</p>
1578022	-	<p><b>Description:</b> OVS offload: fragmented traffic is not offloaded.  When sending traffic with packets bigger than MTU, traffic runs but is not offloaded.</p> <p><b>Workaround:</b> N/A</p> <p><b>Keywords:</b> OVS offload, fragmentation</p>

## Notice

This document is provided for information purposes only and shall not be regarded as a warranty of a certain functionality, condition, or quality of a product. NVIDIA Corporation ("NVIDIA") makes no representations or warranties, expressed or implied, as to the accuracy or completeness of the information contained in this document and assumes no responsibility for any errors contained herein. NVIDIA shall have no liability for the consequences or use of such information or for any infringement of patents or other rights of third parties that may result from its use. This document is not a commitment to develop, release, or deliver any Material (defined below), code, or functionality.

NVIDIA reserves the right to make corrections, modifications, enhancements, improvements, and any other changes to this document, at any time without notice.

Customer should obtain the latest relevant information before placing orders and should verify that such information is current and complete.

NVIDIA products are sold subject to the NVIDIA standard terms and conditions of sale supplied at the time of order acknowledgement, unless otherwise agreed in an individual sales agreement signed by authorized representatives of NVIDIA and customer ("Terms of Sale"). NVIDIA hereby expressly objects to applying any customer general terms and conditions with regards to the purchase of the NVIDIA product referenced in this document. No contractual obligations are formed either directly or indirectly by this document.

NVIDIA products are not designed, authorized, or warranted to be suitable for use in medical, military, aircraft, space, or life support equipment, nor in applications where failure or malfunction of the NVIDIA product can reasonably be expected to result in personal injury, death, or property or environmental damage. NVIDIA accepts no liability for inclusion and/or use of NVIDIA products in such equipment or applications and therefore such inclusion and/or use is at customer's own risk.

NVIDIA makes no representation or warranty that products based on this document will be suitable for any specified use. Testing of all parameters of each product is not necessarily performed by NVIDIA. It is customer's sole responsibility to evaluate and determine the applicability of any information contained in this document, ensure the product is suitable and fit for the application planned by customer, and perform the necessary testing for the application in order to avoid a default of the application or the product. Weaknesses in customer's product designs may affect the quality and reliability of the NVIDIA product and may result in additional or different conditions and/or requirements beyond those contained in this document. NVIDIA accepts no liability related to any default, damage, costs, or problem which may be based on or attributable to: (i) the use of the NVIDIA product in any manner that is contrary to this document or (ii) customer product designs.

No license, either expressed or implied, is granted under any NVIDIA patent right, copyright, or other NVIDIA intellectual property right under this document. Information published by NVIDIA regarding third-party products or services does not constitute a license from NVIDIA to use such products or services or a warranty or endorsement thereof. Use of such information may require a license from a third party under the patents or other intellectual property rights of the third party, or a license from NVIDIA under the patents or other intellectual property rights of NVIDIA.

Reproduction of information in this document is permissible only if approved in advance by NVIDIA in writing, reproduced without alteration and in full compliance with all applicable export laws and regulations, and accompanied by all associated conditions, limitations, and notices.

## Trademarks

NVIDIA, the NVIDIA logo, and Mellanox are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. Other company and product names may be trademarks of the respective companies with which they are associated.

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

## Copyright

© 2020 NVIDIA Corporation. All rights reserved.

