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.
Table of Contents

Document Revision History ........................................................................................................... 5

1  Firmware Burning ...................................................................................................................... 6

2  Port Type Management .............................................................................................................. 7
   2.1  ConnectX-4 Port Type Management/VPI Cards Configuration .............................................. 7

3  Modules Loading and Unloading ............................................................................................. 8

4  Important Packages and Their Installation ............................................................................ 9

5  SR-IOV Configuration .......................................................................................................... 11
   5.1  Setting up SR-IOV in ConnectX-3/ConnectX-3 Pro ........................................................... 11

6  Default RoCE Mode Setting .................................................................................................... 13
List of Tables

Table 1: Document Revision History ........................................................................................................ 5
## Document Revision History

*Table 1: Document Revision History*

<table>
<thead>
<tr>
<th>Revision</th>
<th>Date</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLES 12 SP2</td>
<td>December 12, 2016</td>
<td>Initial version of this document.</td>
</tr>
</tbody>
</table>
1. **Firmware Burning**

1. Identify the adapter card's PSID.

   ```bash
   # mstflint -d 81:00.0 q
   Image type:          FS2
   FW Version:          2.36.5000
   FW Release Date:     26.1.2016
   Rom Info:            type=PXE version=3.4.718 devid=4103
   Device ID:           4103
   Description:
   Node             Port1            Port2
   Sys image
   GUIDs:               e41d2d0300b3f590 e41d2d0300b3f591 e41d2d0300b3f592
   e41d2d0300b3f593
   MACs:                                     e41d2db3f591     e41d2db3f592
   VSD:
   PSID:                MT_1090111019
   ``

2. Download the firmware BIN file from the Mellanox website that matches your card's PSID:

   [www.mellanox.com](http://www.mellanox.com) ➔ Support/Education ➔ Support Downloader

3. Burn the firmware.

   ```bash
   # mstflint -d <lspci-device-id> -i <image-file> b
   ``

4. Reboot your machine after the firmware burning is completed.
2 Port Type Management

2.1 ConnectX-4 Port Type Management/VPI Cards Configuration

ConnectX®-4 ports can be individually configured to work as InfiniBand or Ethernet ports. By default both ConnectX®-4 VPI ports are initialized as InfiniBand ports. If you wish to change the port type, use the mstflint tool after the driver is loaded.

1. Install mstflint tools: Zypper install mstflint.

2. Check the PCI address.

   `lspci | grep Mellanox
   00:06.0 Infiniband controller: Mellanox Technologies MT27520 Family
   [ConnectX-3 Pro]`

3. Use mstconfig to change the link type as desired IB – for InfiniBand, ETH – for Ethernet.

   `mstconfig -d <device pci> s LINK_TYPE_P1/2=<ETH|IB|VPI>
   Example:
   mstconfig -d 82:00.1 s LINK_TYPE_P1=ETH`

4. Reboot your machine.
3 Modules Loading and Unloading

➢ To load and unload the modules, use the commands below:

- Loading the driver: modprobe <module name>
  ```
  modprobe mlx4_en
  ```

- Unloading the driver: modprobe -r <module name>
  ```
  modprobe -r mlx4_en
  ```
4 Important Packages and Their Installation

libibverbs: InfiniBand verbs library

<table>
<thead>
<tr>
<th>Package</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>libibverbs-devel</td>
<td>Development files for the libibverbs library</td>
</tr>
<tr>
<td>libibverbs-devel-static</td>
<td>Static libibverbs library</td>
</tr>
<tr>
<td>libibverbs-runtime</td>
<td>Tools for the InfiniBand Verbs library and manpages</td>
</tr>
<tr>
<td>libibverbs1</td>
<td>InfiniBand verbs library</td>
</tr>
<tr>
<td>libibverbs1-32bit</td>
<td>InfiniBand verbs library</td>
</tr>
</tbody>
</table>

librdmacm: RDMA cm library

<table>
<thead>
<tr>
<th>Package</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>librdmacm-devel</td>
<td>Development files for the librdmacm library</td>
</tr>
<tr>
<td>librdmacm-tools</td>
<td>Tools and example programs using the RDMA cm library</td>
</tr>
<tr>
<td>librdmacm1</td>
<td>RDMA cm runtime library</td>
</tr>
<tr>
<td>librdmacm1-32bit</td>
<td>RDMA cm runtime library</td>
</tr>
</tbody>
</table>

libibcm: Userspace InfiniBand Connection Management API

<table>
<thead>
<tr>
<th>Package</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>libibcm-devel</td>
<td>Development files for the libibcm library</td>
</tr>
<tr>
<td>libibcm-devel-static</td>
<td>Libibcm static library</td>
</tr>
<tr>
<td>libibcm1</td>
<td>InfiniBand Connection Management runtime library</td>
</tr>
<tr>
<td>libibcm1-32bit</td>
<td>InfiniBand Connection Management runtime library</td>
</tr>
</tbody>
</table>

libibmad: Low layer InfiniBand diagnostic and management programs

<table>
<thead>
<tr>
<th>Package</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>libibmad-devel</td>
<td>Development files for the libibmad library</td>
</tr>
<tr>
<td>libibmad-devel-static</td>
<td>Static libibmad library</td>
</tr>
<tr>
<td>libibmad5</td>
<td>Libibmad runtime library</td>
</tr>
<tr>
<td>libibmad5-32bit</td>
<td>Libibmad runtime library</td>
</tr>
</tbody>
</table>

libibumad: Low layer InfiniBand diagnostic and management programs

<table>
<thead>
<tr>
<th>Package</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>libibumad-devel</td>
<td>Development files for the libibumad library</td>
</tr>
<tr>
<td>libibumad-devel-static</td>
<td>Static libibmad library</td>
</tr>
<tr>
<td>libibumad5</td>
<td>Libibmad runtime library</td>
</tr>
<tr>
<td>libibumad5-32bit</td>
<td>Libibmad runtime library</td>
</tr>
</tbody>
</table>

libmlx4: Mellanox ConnectX InfiniBand HCA User space Driver

<table>
<thead>
<tr>
<th>Package</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>libmlx4-rdmav2</td>
<td></td>
</tr>
<tr>
<td>libmlx4-rdmav2-32bit</td>
<td></td>
</tr>
</tbody>
</table>

libmlx5: Mellanox Connect-IB InfiniBand HCA User space Driver

<table>
<thead>
<tr>
<th>Package</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>libmlx5-rdmav2</td>
<td></td>
</tr>
<tr>
<td>libmlx5-rdmav2-32bit</td>
<td></td>
</tr>
</tbody>
</table>

opensm: InfiniBand Subnet Manager

<table>
<thead>
<tr>
<th>Package</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>opensm</td>
<td>InfiniBand Subnet Manager</td>
</tr>
<tr>
<td>opensm-devel</td>
<td>Development files for OpenSM</td>
</tr>
<tr>
<td>opensm-devel-static</td>
<td>Static libraries for OpenSM</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Package</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>opensm-libs3</td>
<td>OpenSM runtime libraries</td>
</tr>
<tr>
<td>opensm-libs3-32bit</td>
<td>OpenSM runtime libraries</td>
</tr>
</tbody>
</table>

**Ibutils: OpenIB Mellanox InfiniBand Diagnostic Tools**

ibutils

**infiniband-diags: OpenFabrics Alliance InfiniBand Diagnostic Tools**

infiniband-diags

**infiniband-diags-devel**

**infiniband-diags-devel-static**

**srptools: Tools for SRP/IB**

srptools

**perftest: IB Performance tests**

perftest

**mstflint: Mellanox Firmware Burning and Diagnostics Tools**

mstflint

mstflint-devel

- To install the packages above run:

  ```bash
  #zypper -n install libibverbs* librdmacm* libibcm* libibmad* libibumad* libmlx4* libmlx5* opensm ibutils infiniband-diags srptools perftest mstflint
  ```
5 SR-IOV Configuration

5.1 Setting up SR-IOV in ConnectX-3/ConnectX-3 Pro

1. Download mstflint tools: zypper install mstflint

2. Check the device’s PCI.

   ```
   lspci | grep mellanox
   ```

3. Check if SR-IOV is enabled in the firmware.

   ```
   mstconfig -d <device pci> q
   ```

   Example:

   ```
   # mstconfig -d 81:00.0 q
   Device #1:
   ------------
   Device type: ConnectX3Pro
   PCI device: 81:00.0
   Configurations: Current
   SRIOV_EN          True(1)
   NUM_OF_VFS        0
   LINK_TYPE_P1      VPI(3)
   LINK_TYPE_P2      VPI(3)
   LOG_BAR_SIZE      3
   BOOT_PKEY_P1      0
   BOOT_PKEY_P2      0
   BOOT_OPTION_ROM_EN_P1 True(1)
   BOOT_VLAN_EN_P1   False(0)
   BOOT_RETRY_CNT_P1 0
   LEGACY_BOOT_PROTOCOL_P1 PXE(1)
   BOOT_VLAN_P1      1
   BOOT_OPTION_ROM_EN_P2 True(1)
   BOOT_VLAN_EN_P2   False(0)
   BOOT_RETRY_CNT_P2 0
   LEGACY_BOOT_PROTOCOL_P2 PXE(1)
   BOOT_VLAN_P2      1
   IP_VER_P1         IPv4(0)
   IP_VER_P2         IPv4(0)
   ```

4. Check SRIOV_EN and NUM_OF_VFS configurations.

5. Enable SR-IOV:

   ```
   mstconfig -d <device pci> s SRIOV_EN=<False|True>
   ```

6. Configure the needed number of VFs

   ```
   mstconfig -d <device pci> s NUM_OF_VFS=<NUM>
   ```

**NOTE:** This file will be generated only if IOMMU is set in the grub.conf file (by adding “intel_iommu=on” to /boot/grub/grub.conf file).

7. **[mlx4 devices only]** Edit the file /etc/modprobe.d/mlx4.conf:

   ```
   options mlx4_core num_vfs=[needed num of VFs] port_type_array=[1/2 for IB/ETH],[ 1/2 for IB/ETH]
   ```

   Example:
options mlx4_core num_vfs=8 port_type_array=1,1

8. **[mlx5 devices only]** Write to the sysfs file the number of needed VFs.

   `echo [num_vfs] > /sys/class/infiniband/mlx5_0/device/sriov_numvfs`

   **Example:**

   `echo 8 > /sys/class/infiniband/mlx5_0/device/sriov_numvfs`

9. Reboot the driver.

10. Load the driver and verify that the VFs were created.

    `lspci | grep mellanox`

    **Example:**

    ```
    dev-r-vrt-214:~ # lspci | grep nox
    82:00.0 Ethernet controller: Mellanox Technologies MT27700 Family [ConnectX-4]
    82:00.1 Ethernet controller: Mellanox Technologies MT27700 Family [ConnectX-4]
    82:00.2 Ethernet controller: Mellanox Technologies MT27700 Family [ConnectX-4 Virtual Function]
    82:00.3 Ethernet controller: Mellanox Technologies MT27700 Family [ConnectX-4 Virtual Function]
    82:00.4 Ethernet controller: Mellanox Technologies MT27700 Family [ConnectX-4 Virtual Function]
    82:00.5 Ethernet controller: Mellanox Technologies MT27700 Family [ConnectX-4 Virtual Function]
    ```

For further information, refer to section Setting Up SR-IOV MLNX_OFED User Manual.
6 Default RoCE Mode Setting

1. Create a directory for the mlx4/mlx5 device.

```
mkdir -p /sys/kernel/config/rdma_cm/mlx4_0/
```

2. Validate what is the used RoCE mode in the default_roce_mode config file.

```
# cat /sys/kernel/config/rdma_cm/mlx4_0/ports/1/default_roce_mode
IB/RoCE v1
```

3. Change the default RoCE mode,

- For RoCE v1: IB/RoCE v1
- For RoCE v2: RoCE v2

```
# echo "RoCE v2" >
/sys/kernel/config/rdma_cm/mlx4_0/ports/1/default_roce_mode
# cat /sys/kernel/config/rdma_cm/mlx4_0/ports/1/default_roce_mode
RoCE v2

# echo "IB/RoCE v1" >
/sys/kernel/config/rdma_cm/mlx4_0/ports/1/default_roce_mode
# cat /sys/kernel/config/rdma_cm/mlx4_0/ports/1/default_roce_mode
IB/RoCE v1
```