RedHawk 8.4 OFED for InfiniBand Cards Installation Instructions
Release Notes

May 18, 2022

READ ME BEFORE INSTALLING THIS PRODUCT

concurrent REAL-TIME

This page intentionally left blank

1. Introduction

This document provides instructions for installing a modified version of the Mellanox OFED version 5.6-1.0.3.3 for InfiniBand cards software onto an existing system installed with RedHawk Linux 8.4.x.

2. Requirements

- An x86_64 host system running RedHawk 8.4.x
- A *RedHawk OFED for InfiniBand* version 8.4 DVD
- A supported InfiniBand adapter card listed in the Release Notes on page 12 of the *Mellanox OFED Documentation*, available on the DVD.

3. Installation

3.1. Boot to RedHawk 8.4

Ensure that the system is booted with a RedHawk 8.4 kernel and invoke the following command to verify:

```
# uname -r
```

3.2. Prepare RedHawk kernel source

Issue the following commands as the root user to prepare the RedHawk kernel source tree for building kernel modules.

```
# cd /lib/modules/$(uname -r)/build
# ./ccur-config -c -n
```

3.3. Install required base distribution packages

Mount the *RedHawk OFED for InfiniBand* DVD and issue the following commands to install the various base distribution packages that are required by OFED:

```
# mkdir -p /tmp/mount
# mount mlnxofed-5.6-1.0.3.3-RedHawk-8.4-x86_64.iso /tmp/mount
# cd /tmp/mount
# ./install-required-base-packages-for-ofed
```

3.4. Install OFED software

In the same mounted directory, issue the following command to build and install all the OFED software:

```
# ./mlnxofedinstall --add-kernel-support --skip-repo
# dracut -f
```

The installation of the OFED software may alter the default boot kernel. It is recommended to verify the original kernel used for installing the OFED kernel is still the default boot kernel.

NOTE

*This installation script can run for as long as 30 minutes as it compiles
and installs kernel modules and several user-space packages and then
updates the firmware in the InfiniBand adapter card installed in the system.
The many warnings produced during installation can be safely ignored.*

When the installation is complete, <u>ignore</u> the suggestion to reload the `openibd` driver and instead immediately reboot the system with the same RedHawk kernel. After the reboot, proceed with the steps below.

3.5. Query board status

Invoke the following command to verify that the InfiniBand driver has loaded successfully upon reboot:

```
# ibstat
```

Output like the following should be displayed:

```
CA 'mlx5_0'
CA type: MT4119
Number of ports: 1
Firmware version: 16.23.1000
Hardware version: 0
Node GUID: 0x506b4b030029a2be
System image GUID: 0x506b4b030029a2be
Port 1:
State: Down
Physical state: Polling
Rate: 10
Base lid: 65535
LMC: 0
SM lid: 0
Capability mask: 0x2651e84a
Port GUID: 0x506b4b030029a2be
Link layer: InfiniBand# ibstat
```

Note that the port state will be shown as `Down` if the board is not currently connected with another InfiniBand adapter card or InfiniBand switch.

3.6. Simple InfiniBand configuration test

This section provides steps for verifying that connections between two directly connected InfiniBand cards can be established successfully. You may be able to skip these steps if your system is part of an established InfiniBand network at your site; contact your administrator.

1. On both systems, modify the `/etc/hosts` file to define two InfiniBand addresses. In this simple example, one system is named *server* and the other system is named *client*. Add entries like the following: note that the text added must be identical on both systems:

```
10.1.1.1 server-ib
10.1.1.2 client-ib
```

2. On both systems, create a `/etc/sysconfig/network-scripts/ifcfg-ib0` file with contents like the following:

```
NAME=ib0
DEVICE=ib0
BOOTPROTO=static
NM_CONTROLLED=yes
TYPE=InfiniBand
IPADDR=10.1.1.1          ← Use 10.1.1.1 for server and 10.1.1.2 for client
NETMASK=255.255.255.0
NETWORK=10.1.1.0
BROADCAST=10.1.1.255
ONBOOT=yes
```

3. On the server system, create an initial Subnet Manager configuration file:

```
# cd /etc/opensm
# /usr/sbin/opensm -c opensm.conf
```

4. On the server system, run the following command to determine the actual Port GUID number for the server system's installed InfiniBand adapter card. For example:

```
# ibstat -p
0x506b4b030029a2b3
```

5. On the server system, edit `/etc/opensm/opensm.conf` and change the initial `guid` entry on line 5 from `0x0000000000000000` to the number displayed above and save the file.

6. On the server system, Issue the following commands to enable, restart and display the status of the Subnet Manager:

```
# chkconfig --level 235 opensmd on
# systemctl enable opensmd
# systemctl restart opensmd
# systemctl -l status opensmd
```

   The output from the above commands should indicate that the Subnet Manager is both active and running.

7. Reboot both systems, again ensuring that the same RedHawk kernels are booted.

8. Login after reboot and issue the following commands on both systems to verify that the InfiniBand network is configured correctly and that both hosts are active and connected:

```
# ifconfig ib0
# ibhosts
# iblinkinfo
```

9. Finally, verify that you can ping and ssh from both systems to the other system:

   - ping client-ib from *server*

- ping server-ib from *client*
- ssh into client-ib from *server*
- ssh into server-ib from *client*

This concludes the simple InfiniBand configuration and connection verification.

Please contact Concurrent Real-Time technical support if you had any problems during this installation (http://www.concurrent-rt.com/support or 1-800-245-6453).

3.7. IP over InfiniBand Throughput Testing

This section provides steps for verifying the connection between two directly connected InfiniBand cards can achieve advertised throughput. You may be able to skip these steps if your system is part of an established InfiniBand network at your site; contact your administrator. Note this test measures throughput and these speeds will never achieve theoretical packet delivery (advertised bandwidth).

1. On both systems install the network traffic measurement tool iperf:

```
# yum install epel-release
# yum install iperf
# systemctl stop firewalld.service
```

2. On the server run the following:

```
# iperf -s -P 50
```

3. On the client run the following:

```
# iperf -c 10.1.1.1 -P 50
```

Output like the following should appear on the server when the iperf connections have closed:

```
[SUM] 0.00-10.03 sec     31.2 GBytes   26.7 Gbits/sec
```

Note this test is extremely CPU bound. The '-P' option specifies a specific number of parallel threads for sending packets. InfiniBand cards capable of high throughput may require a larger argument than 50. As this test is dependent upon the number of threads to achieve up to 100Gb/s, other hardware resources may bottleneck this test's ability to achieve the advertised bandwidth.

4. Porting Notes

This section details all the modifications to the original Mellanox OFED 5.6-1.0.3.3 software that were required in order to make it compatible with RedHawk 8.4. All the modifications described here have already been made to the software provided on the *Mellanox OFED for RedHawk 8.4* Installation DVD; these details are only being provided for informational use.

NOTE

None of these steps are required to use the Mellanox OFED software provided on the Mellanox OFED for RedHawk 8.4 Installation DVD.

This section is only provided for transparency and completeness.

Perform all the following steps as the root user to re-create the modified tarball.

4.1. Prepare RedHawk kernel source tree

Invoke the following commands to prepare the RedHawk kernel source tree for building out-of-tree device drivers:

```
# cd /lib/modules/$(uname -r)/source
# ./ccur-config -c -n
```

4.2. Download Mellanox OFED tarball

Download the Mellanox OFED 5.6-1.0.3.3 Linux tarball from nvidia.com using the following URL:

https://network.nvidia.com/products/infiniband-drivers/linux/mlnx_ofed/

4.3. Expand Mellanox OFED tarball

Invoke the following commands to expand the Mellanox OFED tarball:

```
# tar xzf MLNX_OFED_LINUX-5.6-1.0.3.3-rhel8.4-x86_64.tgz
# cd MLNX_OFED_LINUX-5.6-1.0.3.3-rhel8.4-x86_64
```

4.4. Expand the Mellanox OFED source tarball

Invoke the following commands to expand the Mellanox OFED source tarball:

```
# cd src/
# tar xzf MLNX_OFED_SRC-5.6-1.0.3.3.tgz
# cd MLNX_OFED_SRC-5.6-1.0.3.3
```

4.5. Modify install.pl script

Edit the install.pl file and insert the following lines starting at line 954:

```
if ($DISTRO =~ m/RHEL8/) {
    $rpmbuild_flags .= " --nodebuginfo";
}
```

4.6. Pack up modified Mellanox OFED source tarball

Invoke the following commands to pack up the modified Mellanox OFED source tarball:

```
# cd ..
# tar czf MLNX_OFED_SRC-5.6-1.0.3.3.tgz MLNX_OFED_SRC-5.6-1.0.3.3
# rm -rf MLNX_OFED_SRC-5.6-1.0.3.3
```

4.7. Install modified Mellanox OFED product software

1. Before proceeding with the installation, make sure that the following RPM packages are already installed on your system:

- kernel-rpm-macros
- python36-devel
- tcsh
- gcc-gfortran
- tk
- kernel-modules-extra
- rpm-build
- gdb-headless

2. Invoke the following commands to build and install the modified Mellanox OFED product software:

```
# cd ..
# ./mlnxofedinstall --add-kernel-support --skip-repo
```

Note that the software installation can take 30 minutes or more to complete, however logs of all activity can be found under the following directory on the system:

```
/tmp/MLNX_OFED_LINUX*$(uname -r)*/
```

Once the entire installation process is complete then all Mellanox OFED RPM packages and modules have successfully been installed on the system.