Fabric Computing That Works™

Installation Guide for Intel® Omni-Path Fabric Software - Debian Release

IntelOPA-IFS.DEB8-x86_64.10.2.0.0.158.1

**Prepared for Intel, Inc.
by System Fabric Works, Inc.**

December 15, 2016

Contact Information:

http://www.systemfabricworks.com

e-mail opasupport@systemfabricworks.com

1 Overview

This guide is a supplement to the standard documentation for Intel Omni-Path Fabric Software. These guides include the README for Intel® IFS 10.2.0.0.158, Intel® Omni-Path Architecture (OPA) Fabric Administrator's Guide and [Intel® Omni-Path Fabric SuiteFastFabric User Guide](). The focus of this guide is to document special procedures and considerations for installation and configuration of the software on Debian Linux.

2 Requirements

Note: The target OS is Debian 8.4. However, the release has also been tested with Debian 8.6.

This software release is intended to be installed on a Debian 8.4 host with no updates. The tested environment was installed from the 8.4 DVD set on a machine that was not connected to the Internet. If the machine is connected to the Internet during installation, the Debian installer will automatically upgrade the software to the current release.

It is recommended that all software components are installed from the DVD set for Debian 8.4. The only required images are 1-3. The additional DVDs can be added during installation, or the administrator may add them manually via the apt-cd command.

After installation, the only entries in /etc/apt/sources.list should begin with cdrom:. There should be no uncommented entries for http or ftp sources.

Debian 8.4 images may be downloaded from:

http://cdimage.debian.org/mirror/cdimage/archive/8.4.0/amd64/iso-dvd/

3 OPA Installation

Installation of OPA is performed via the INSTALL script provided within the software distribution. This script installs dependencies from the Debian distribution via apt-get, installs the OPA packages, creates configuration files for RDMA and OPA, rebuilds the initramfs and enables the opa service.

Note: You may wish to execute "INSTALL -h" first in order to see command-line options, such as "enable fabric manager".

1. Unpack the software distribution.
   ```
   tar xf IntelOPA-IFS.DEB8-x86_64.10.2.0.0.158.tgz
   ```
2. Change your working directory to the distribution folder.
   ```
   cd IntelOPA-IFS.DEB8-x86_64.10.2.0.0.158
   ```

3. Execute the installer as root
```
sudo ./INSTALL
```
4. Reboot the host.
```
sudo reboot
```

The installer may report an error after it installs the required packages from apt-get. If that occurs, execute INSTALL again and it should execute without errors. During installation, the Debian packaging rules will not overwrite existing configuration files from a previous install. The packager will output messages which indicate "Keeping old config file as default" and the packaged configuration file will be installed with a dpkg-dist suffix.

Once the installer completes, the administrator should reboot the host. After reboot, the hfi1 and ib modules should be loaded and the fabric can be tested according to procedures in the Administrator's Guide.

The packages installed are all located under packages in the distribution folder. This folder also includes Debian build artifacts and source tarballs.

Generated configuration files include:
- /etc/rdma/rdma.conf
- /etc/modules-load.d/rdma.conf
- /etc/security/limits.conf - modified. Look for "OPA Settings" in the file.
- /var/mpi-selector/data/

4 OPA Uninstallation

OPA is uninstalled via the INSTALL script in the software distribution. This script removes the installed OPA packages and removes generated configuration files. It does not remove the dependencies that were installed via apt-get.

1. Change your working directory to the distribution folder.
```
cd IntelOPA-IFS.DEB8-x86_64.10.2.0.0.158
```
Execute the installer as root, with the -u argument
```
sudo ./INSTALL -u
```

The installer does not modify /etc/security/limits.conf. The administrator can remove the OPA entries from this file before rebooting the host.

If you wish to remove all packages and their configurations (i.e. "purge"), execute
```
./INSTALL -u -p.
```

5 RDMA Module Loading

On the RedHat platform, RDMA kernel modules are usually loaded via /usr/libexec/rdma-init-kernel. This is a RedHat package, and the prescribed method for

loading these modules on Debian is via a configuration file in /etc/modules-load.d. The installer creates rdma.conf to load the RDMA modules at boot time.

A consequence of this technique is that the settings in /etc/rdma/rdma.conf are not used for the RDMA modules (e.g. "IPOIB_LOAD=no does not prevent ib_ipoib from loading). However, the opa service does use this file for its configuration.

The administrator may edit /etc/rdma/rdma.conf to enable/disable modules or provide module parameters.

6   IPoIB Interface Configuration

An IPoIB interface is configured using the standard Debian network interface configuration techniques. An overview is provided at https://pkg-ofed.alioth.debian.org/howto/infiniband-howto-5.html

To bring up the ib0 interface temporarily, execute the following as root, using the appropriate CIDR:
```
  ifconfig ib0 10.20.30.2/24
```

The ib0 device should then be configured with an IP address.
```
  ifconfig ib0
  ping 10.20.30.2
```

To bring up the ib0 interface automatically at boot time, add the following to /etc/network/interfaces:

```
auto ib0
iface ib0 inet static
  address 10.20.30.2
  netmask 255.255.255.0
  broadcast 10.20.30.255
```

If connected mode is desired, and/or an MTU should be set, use a configuration that resembles the following:

```
auto ib0
iface ib0 inet static
  address 10.20.30.2
  netmask 255.255.255.0
  broadcast 10.20.30.255
  post-up echo connected > /sys/class/net/ib0/mode && ifconfig ib0 mtu 65520
```

The post-up directive in that configuration sets ib0's IPoIB mode to "connected" and sets the MTU to 65520. The directive may be modified to meet your requirements.

Test the configuration:
```
  ifup ib0
```

```
ifconfig ib0 #view and verify IP address
ping 10.20.30.2
```

The ib0 interface should come up automatically and be configured with the IP address when the host is rebooted.

7 Use of PSM and /dev/ipath, MPI

If an MPI implementation has not been compiled with PSM2 support, and PSM is desired instead of ibverbs, please review section 4 of:

http://www.intel.com/content/dam/support/us/en/documents/network-and-i-o/fabric-products/Intel_OP_Fabric_Host_Software_UG_H76470_v4_0.pdf

In order to create /dev/ipath, which is required for PSM support:
1. cp /lib/udev/rules.d/40-psm-compat.rules /etc/udev/rules.d
2. Edit /etc/udev/rules.d/40-psm-compat.rules. Change hfi1 in the last line to the name of the hfi1 device you want to use for ipath. For example:

```
KERNEL=="hfi1_0", SYMLINK+="ipath"
```

3. Reboot the host.

8 openmpi and Debian

There are two issues when running openmpi on Debian:
1. openmpi does not specify absolute paths for certain utilities that it executes via ssh (e.g. orted).
2. Debian's .bashrc terminates early when executed in a non-interactive shell (e.g. a remote ssh command).

The consequence of this is that mpirun will report an error "bash: orted: command not found". In order to correct this:
1. Use mpi-selector-menu to select openmpi on all nodes.
2. Add ". /etc/profile.d/mpi-selector.sh" to the beginning of $HOME/.bashrc on all nodes.

9 Fabric Manager Service

The Fabric Manager installed by opa-fm is not enabled by default. This service is managed via the systemctl utility.

To start the Fabric Manager:

```
systemctl start opafm
```

To stop the Fabric Manager:
```
systemctl stop opafm
```

To enable the Fabric Manager on boot:
```
systemctl enable opafm
```

10 Build mpi_apps

In order to build mpi_apps, which are required for certain OPA utilities (e.g. opacabletest), the administrator should install the necessary development tools and execute the make process.

```
# sudo bash
# apt-get install build-essential gfortran
# ln -s /usr/bin/make /usr/bin/gmake
# cd /opt/opa/src/mpi_apps
# make
```

After the build completes, these final messages should appear on the console:

```
build base sample applications
Built subset of sample applications
Built sample applications
```

11 References

- Intel® Omni-Path Architecture (OPA) Fabric Administrator's Guide
- Intel® Omni-Path Fabric SuiteFastFabric User Guide
- Infiniband HOWTO: Setting up a basic infiniband network
- Infiniband HOWTO: IP over Infiniband (IPoIB)
- Intel® Omni-Path Fabric Suite FabricManager User Guide