User’s Manual

For

OpenCPI OFED Data Transfer Driver

Author: John Miller   
Lyndeborough Object Technologies, LLC

Last modified: 2/19/2011

Version .0

Prepared For Mercury Federal Systems, OpenCPI Phase II project.

Revision History 3

1 Purpose 4

2 Introduction 5

2.1 OpenCPI Data Plane Overview 5

2.2 Terms and Definitions 5

3 OFED Installation 6

3.1 Prerequisites 6

3.2 Where to find OFED 6

3.3 Using the OFED installation script 6

4 Linking the OFED driver into an Application 7

5 Configuring the OFED driver 8

5.1 User resource limitations 8

5.2 XML Schema format 8

5.3 System Level Configuration Parameters 9

5.4 IB-Verbs Driver Specific Configuration Parameters 9

6 Appendices 11

6.1 OpenCPI OFED Test Application 11

6.2 Reference Documents 12

# Revision History

Version Date By Notes

0.00 2/19/11 John Miller Creation.

0.01 3/9/11 John Miller Added review comments.

# Purpose

The purpose of this document is to provide all of the necessary information required to install, configure and use the OpenCPI OFED driver in an RCC container on a Linux system. It is assumed that the author is familiar with the OFED architecture.

# Introduction

## OpenCPI Data Plane Overview

The OpenCPI data plane provides a high performance heterogeneous data path between distributed RCC and HDL containers. The data plane is organized in two layers, the transport and the data transfer layers. The transport layer is fabric agnostic and provides buffer and protocol management. The data transfer layer is a RDMA driver level module that allows new drivers to be plugged into OpenCPI enabling data exchange over a variety of fabrics. OpenCPI provides several intrinsic drivers with the standard distribution including Host Memory, network sockets and PCI bus.

## Terms and Definitions

IB-Verbs - InfiniBand Verbs

OpenCPI - Open Component Portability Infrastructure

OFED - Open Fabrics Enterprise Distribution

OFEDTD - OpenCPI OFED data transfer driver

RCC - Resource Constrained C Language

PIO - Programmed Input/Output

SMB - Shared memory block. Note that the term SMB refers to a block of memory that is shared by a endpoint to support RDMA, it does not imply PIO accessibility.

# OFED Installation

The OpenFabrics Enterprise Distribution must be installed on any OpenCPI system that will be used to run or build OpenCPI applications that will utilized OFED enabled networks. OFED installation is not required to compile and build the OpenCPI OFED driver library.

This section is not intended to be a comprehensive OFED installation guide, the OFED web site provides all of the up to date information required to install their software distribution.

The OpenCPI OFED driver has only been tested with the OFED version 1.5.3.

## Prerequisites

The OpenFabrics Enterprise Distribution is a standalone distribution that does not require any addition software prior to installation. However there are several linux package dependencies that are required and the OFED installation script will identify each dependency during the installation process.

OFED supports a limited sub-set of linux OS and kernel versions. Check the OFED web site to determine if your configuration is supported.

## Where to find OFED

The standard OFED installation can be found on the OpenFabrics download page.

[www.openfabrics.org/OFA-Linux.html](http://www.openfabrics.org/OFA-Linux.html)

Follow the directions on the OFED web site to download the OFED release.

## Using the OFED installation script

Once an OFED release has been downloaded and installed on your computer, type the following command from the OFED installation directory

* ./install.pl <CR>

Follow the scripts instructions to complete the OFED installation.

# Linking the OFED driver into an Application

It is necessary for an OpenCPI application that uses the RCC Container to include any required optional data plane drivers during link time. Therefore, to include the OpenCPI OFED data plane driver the following commands must be included when the application gets linked.

$(OCPI\_BASE)/core/dataplane/rdma\_drivers/$(OCPI\_OUT\_DIR)/DtOfedXfer.obj \

-L $(OCPI\_BASE)/libs -lrdma\_driver\_interface -locpios

# Configuring the OFED driver

The IB-Verbs API provides several programmable parameters that allow fine-tuning for specific driver implementations. All of the configurable parameters have default values for the general case. This section identifies the System and IB-Verbs parameters that can be modified by providing the appropriate XML data to the OpenCPI system.

## User resource limitations

The OpenCPI IB-Verbs driver uses locked memory. Therefore the user must ensure that the application is running in an environment in which the amount of “lockable memory” is greater than or equal to the memory that the driver is configured to allocate.

## XML Schema format

The OFEDDT configuration data is XML based and has the following schema.

<?xml version="1.0" encoding="UTF-8" ?>  
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">

<xs:complexType name="OpenCPIOFEDDeviceDrivers">

<xs:sequence>

<xs:element name="device" type="xs:string" minOccurs=”0”

maxOccurs=”unbounded”>

<!-- Device id Name of the linux device -->

<xs:attribute name="id" type="xs:string" use="required"/>

<!-- Device port -->

<xs:attribute name="port" type="xs:positiveInteger" use="required"/>

<!-- SMB size -->

<xs:attribute name="SMBSize" type="xs:positiveInteger"

use="optional"/>

<!-- IBV\_QP\_MIN\_RNR\_TIMER -->

<xs:attribute name="min\_rnr\_timer" type="xs:positiveInteger" use="optional"/>

<!-- IBV\_QP\_RETRY\_COUNT -->

<xs:attribute name="retry\_cnt" type="xs:positiveInteger" use="optional"/>

<!-- IBV\_QP\_TIMEOUT -->

<xs:attribute name="timeout" type="xs:positiveInteger" use="optional"/>

<!-- IBV\_RNR\_RETRY -->

<xs:attribute name="rnr\_retry" type="xs:positiveInteger" use="optional"/>

</xs:element>

</xs:sequence>

</xs:complexType>

</xs:schema>

## System Level Configuration Parameters

These are the Data Transfer system level attributes that affect all the installed drivers. These parameters can be overridden in the device specific attribute section of the XML.

SMBSize - Transfer driver Shared Memory Block size in bytes.

## IB-Verbs Driver Specific Configuration Parameters

This section describes the OFED IB-Verbs specific attributes that can be modified from the driver’s XML configuration file. A brief description is provided with each attribute along with the IB-Verbs attribute flag that it is associated with. Please refer to your device specific documentation to determine the optimal setting for each of these attributes.

Note that the OpenCPI OFED driver uses a Reliable Channel (RC) for its implementation and several of the following parameters affect the RC behavior.

|  |  |
| --- | --- |
| Attribute name | Description |
| port | IB-Verbs port associated with the device |
| hopLimit | Maximum number of routers a global packet is permitted to cross before it is discarded. |
| IBV\_QP\_MIN\_RNR\_TIMER | Minimum “Receiver Not Ready” timer |
| IBV\_QP\_RETRY\_COUNT | Number of retries for Ack timeout |
| IBV\_QP\_TIMEOUT | Local Ack timeout |
| IBV\_RNR\_RETRY | Retry on RNR timeout if 1 |
|  |  |

# Appendices

## OpenCPI OFED Test Application

An OpenCPI RCC test application can be found in

$(OpenCPI\_BASE)/core/container/ctests/bin/src/RTxTest\_main.cxx. This test can be used to determine if your OFED drivers are operating properly. To run this test, make sure the environment variable named “OpenCPI\_HAVE\_IBVERBS” is set to 1 and build the OpenCPI code tree. The executable can be found in $(OpenCPI\_BASE)/core/container/ctests/$(OpenCPI\_OUT\_DIR). Execute the following command

* ./RtxTest --sd

The --sd switch tells the test to just print out the available endpoints that have been discovered by the installed drivers. An example output from this command is shown below.

List of supported endpoints:

OpenCPI-ofed-rdma://rxe0:1:0.0:0:0:0:0:3145728.1.10

OpenCPI-socket-rdma://localhost.localdomain;40001:3145728.2.20

OpenCPI-pci-pio://0.0:3145728.3.10

OpenCPI-smb-pio://pioXfer189380:3145728.4.10

The example output above shows that an OFED device was discovered and the associated endpoint is displayed in the first line of the output. Now we can run the test again and select the OFED endpoint using the following command.

* ./RTxTest --pi=0

The --pi switch tells the test to use protocol endpoint string index 0 which corresponds to the OFED device. This part of the test is now running and is acting as a server waiting for the client part of the test to connect.

Now we can run the client part of the test on a remote machine that has an OFED network connection to the device that was selected for the server test, which is now listening for a client connection.

(Note that this test requires a network socket connection between the remote machines to establish the RCC connections. This is only a requirement for the test; a network connection is not required by the OFED RDMA driver. )

To start up the client part of the test you must select the protocol endpoint string like we did with the server. Make sure that the selected OFED device is on the same network as the device that was selected for the server. In addition we will tell the client where to find the server with the --h switch.

* ./RTxTest --pi=1 --h=10.0.1.5

If the test is running properly the following messages will appear on in the window that the server is running in.

Buffer 0 data integrity test passed

Buffer 500 data integrity test passed

Buffer 1000 data integrity test passed

These messages should continue until the test has been manually terminated. If this test fails to run, refer to the OFED documentation and your devices documentation to ensure that your network is operational.

## Reference Documents

Document Link

OFED Overview <http://www.openfabrics.org/OFED-Overview.html>

OFED API Documentation <http://www.openfabrics.org/OFA-Linux.html>

InfiniBand NetWork Architecture Author: Tom Shanley