

OFA IWG meeting – 1/29/2013

Attendees

First	Last	Company	12/04/12	12/18/12	01/15/13	01/29/13
Samantha	Jian-Pielak	Canonical		X	X	X
Tom	Reu	Chelsio		X		
Martin	Schlining	DataDirect Networks			X	X
Jeff	Kopko	Emulex				
Harry	Cropper	Intel		X	X	X
Mitko	Haralanov	Intel		X	X	X
Jess	Robel	Intel	X	X	X	X
Jim	Ryan	Intel		X		
Guy	Ergas	Mellanox		X		
Rupert	Dance	Software Forge	X	X		X
Glenn	Martin	UNH-IOL		X	X	X
Edward	Mossman	UNH-IOL	X	X	X	X
Bob	Noseworthy	UNH-IOL		X		X
Nate	Rubin	UNH-IOL	X		X	X

2013 Events

- [OFA Programming Course](#) 01/29 → 01/30/2013
- OFA Interop Logo GA Event 01/07/2013

April 2013 dates

- IBTA PF23 04/01 → 04/12/2013
- OFA April Interop Debug Event 04/08 → 04/12/2013

October 2013

- IBTA PF24 10/07 → 10/18/2013
- OFA Interop Debug Event 10/14 → 10/18/2013

OFA EWG Status

- OFED 3.5 RC5 Released 01/17/2013
- OFED 3.5 RC6 Released 01/29/2013
- OFED 3.5 GA 1-2 weeks after RC6

Test Plan updates

- 1) RoCE
 - a) IBM has joined the OFILG and intends to test its RoCE products at the April 2013 Interop Event. Blade Network is owned by IBM and we will approach them about including DCB switches as part of the RoCE testing.
- 2) iWARP – 40 Gb
 - a) The program needs some 10 and 40 GbE switches that can be used in the iWARP testing. We have been using a Fujitsu multi-port 10 GbE switch which Fujitsu has loaned to the group.
 - b) Harry - they do not have anything in the time frame we need it. They will consider it for October 2013.
- 3) Should we consider allowing switch vendors to get involvement and give logo. The suggestion was that we charge them a reduced fee of perhaps 50% because they don't get nearly as much out of it.
- 4) **Bob**: Extreme 40 GbE and a Fulcrum (Intel- now) 10 GbE DCB switch.

OFED 3.5 GA Logo – Review by Edward

- 1) He would like to start ASAP and so will begin as soon as it is released. At max it will be one week after the release of the GA
- 2) We need to verify that all firmware has been updated.
- 3) Mitko and Jess both feel all is ok on the Firmware side.
 - a) Mitko says from driver side all functionality is there but not on testing.

Requests to modify the OFILP – proposed by Jim Ryan and Intel

Qualified Pass

- 1) Allow a product on the Logo List as long as all issues and defects are very clearly outlined.

Arbitration Committee

- 1) Allow additional OFA members to participate in the Arbitration process. For example Oracle, IBM, HP and national labs who are member of the OFA.
- 2) OFILG members are disqualified from involvement in the arbitration process unless their products are involved and then they are only allowed to participate in claims involving their products.
- 3) The OFA IWG will continue to maintain a standing committee (currently Software Forge and UNH-IOL) but we can bring in additional OFA members based on the agreement of the arbitration parties involved.

Determining components to test

- 1) Allow more feedback from OFILG members regarding qualified products to be included in Interop Logo testing. The concern is to prevent the inclusion of “ringer” products that could cause another OFILG member product to fail. Potential guidelines for this enhancement:
 - a) A tested product must have a legitimate or current SKU and be available to the general public within 6 months of the event.

More Transparency and Openness

- 1) The goal is to make Interop issues more broadly known to the OFA community and the end users whenever.
 - a) **Bob:** this cannot conflict with confidentiality or NDA agreements in place. Vendor permission must be obtained directly.

OFILP expenses: goal is to reduce program costs – goal is to reduce from \$160K → \$140K

- 1) We cannot continue to do testing for companies that are not OFILG members
 - a) Microsoft
 - b) Oracle (RDS)
 - c) RedHat
 - d) SUSE
 - e) Ubuntu
- 2) We have recommended that we eliminate ULP Testing that is no longer supported or used by OFILG member companies
 - a) iSER
 - b) RDS (Oracle)

- c) SDP
- d) MVAPICH
- 3) Eliminate ULP testing for marginal protocols
 - a) NFSoRDMA
 - b) uDAPL (IBM is now an OFILG member and may want uDAPL support because of GPFS and DB2)
 - c) Vendor SM testing (only test OpenSM)

Open Items

Bob

- 1) What version of OFED do we use in April 2013? It seems like it will be OFED 3.5 GA. He would like to see more testing for RoCE. He wants the product to be on the Logo List.
- 2) Bob would like to see us add a new SM priority test to make sure that vendor SMs as well as the OpenSM yield when an SM is discovered with a higher priority or a lower GUID if the priority is the same. The test would just keep Vendor SM running and vary the OpenSM.
- 3) RoCE – DCB testing
 - a) Make sure we set priority on DCB device and load network with other priority and make sure that the RoCE protocol wins
 - b) Must include at least one DCB
 - c) Need to vet the fact that end point is doing a little bit of DCB. These will be Beta for April and Mandatory by October.
 - d) Intel bought Fulcrum and IBM bought Blade Networks
 - e) Emulex problem
 - i) They could be a **Qualified Pass** since their driver was not in OFED.
 - f) **Switch vendors**
 - i) Arista, Extreme and maybe Cisco.
 - g) Need 40 GbE switches for RoCE.
- 4) Bob wants to see us allow 2-5K for switch vendors. He does not want to include DCB testing because it cannibalize DCB consortium.

OFED Release Notes

- 1) We discussed the need to update the release notes for OFED 3.5 GA. Rupert and the EWG have been working on updating the document. See the [updated release notes](#) following the ARs
- 2) We added all of the new hardware supported by OFED 3.5
- 3) We eliminated all of the notations regarding changes in the versions except for the changes between OFED 1.5.4.1 and OFED 3.5

Action Requests (ARs)

Rupert Dance

1/29/2013

- 1) Check with Mellanox regarding the state of their firmware updates and the readiness to do GA testing
- 2) Discuss with the EWG and determine what are next OFED release
- 3) Get Brad Benton of IBM and Emulex to join next meeting

12/04/2012

- 1) Log SRP bug with RedHat

Older

- 1) Update Logo Program to include OFA Policy regarding OFILG Membership and granting of the Logo
- 2) Check with EWG and XWG about the distribution of OFED Binaries
- 3) Create Logo with version or change Logo Guidelines – [see 4/3/2012 minutes](#)
- 4) Incorporate ICR into Logo Program

UNH-IOL

12/04/2012 - Nate

- 1) The RoCE test plan currently does not reflect the additional tests that we performed (uDAPL and MPI).

10/9/2012

- 1) **Nate:** verify that rping works for RoCE as well as iWARP
- 2) IOL Check the OpenMPI commands for RoCE

5/1/2012

- 1) Make a separate webpage page to show how Distros have been installed. Include a sanitized kick start file and the update logs.

Open Fabrics Enterprise Distribution (OFED)
Version 3.5
Release Notes
February 2013

=====
Table of Contents
=====

1. Overview, which includes:
 - OFED Distribution Rev 3.5 Contents
 - Supported Platforms and Operating Systems
 - Supported HCA and RNIC Adapter Cards and Firmware Versions
 - Tested Switch Platforms
 - Third party Test Packages
 - OFED sources
2. Change log
3. Known Issues

=====
1. Overview
=====

These are the release notes of OpenFabrics Enterprise Distribution (OFED) release 3.5. The OFED software package is composed of several software modules, and is intended for use on a computer cluster constructed as an InfiniBand Fabric, an iWARP Network or a RoCE Fabric.

Note: If you plan to upgrade the OFED package on your cluster, please upgrade all of its nodes to this new version.

1.1 OFED 3.5 Contents

The OFED package contains the following components:

- OpenFabrics core and ULPs:
 - IB HCA drivers (mthca, mlx4, qib, ehca)
 - iWARP RNIC driver (cxgb3, cxgb4, nes)
 - RoCE drivers (mlx4)
 - ib core
 - Upper Layer Protocols: IPoIB, SRP Initiator and target, iSER Initiator and target, RDS, uDAPL, qlgc_vnic and NFS-RDMA*.
- OpenFabrics utilities:
 - OpenSM (OSM): InfiniBand Subnet Manager
 - Diagnostic tools
 - Performance tests
- Extra packages:
 - infinipath-psm: Performance-Scaled Messaging API, an accelerated interface to Intel(R) HCAs
- Sources of all software modules (under conditions mentioned in the modules')

- LICENSE files)
- Documentation

1.2 Supported Platforms and Operating Systems

- o CPU architectures:
 - x86_64
 - x86
 - ppc64
- o Linux Operating Systems:
 - RedHat EL6.2 2.6.32-220.el6
 - RedHat EL6.3 2.6.32-279.el6
 - SLES11 SP2 3.0.13-0.27-default
 - kernel.org 3.5 *

* Minimal QA for these versions.

1.3 HCAs and RNICs Supported

This release supports IB HCAs by IBM, Intel and Mellanox Technologies, iWARP RNICs by Chelsio Communications and Intel and RoCE adapters by IBM and Mellanox.

InfiniBand Adapters

- o IBM HCAs:
 - GX Dual-port SDR 4x IB HCA
 - GX Dual-port SDR 12x IB HCA
 - GX Dual-port DDR 4x IB HCA
 - GX Dual-port DDR 12x IB HCA
- o Intel (formerly QLogic) HCAs:
 - Intel(R) True Scale DDR PCIe x8 and x16 HCAs
 - Intel(R) True Scale QDR PCIe x8 Gen2 HCAs
- o Mellanox Technologies HCAs (SDR, DDR and QDR Modes are Supported):
 - ConnectX(R) and ConnectX EN (fw-25408 Rev 2.9.1000)
 - ConnectX-2 (fw-ConnectX2 Rev 2.9.1000)
 - ConnectX-3 (fw-ConnectX3 Rev 2.11.0500) (FDR and FDR10 Modes are Supported)

For official firmware versions please see:

http://www.mellanox.com/content/pages.php?pg=firmware_download

iWARP Adapters

- o Chelsio RNICs:
 - S310/S320 10GbE Storage Accelerators
 - R310/R320 10GbE iWARP Adapters
- o Intel RNICs:
 - NE020 10Gb iWARP Adapter

RoCE Adapters

- o IBM
 - IBM Flex System EN4132 2-port 10 GbE RoCE
 - IBM EL27 PCIe LP 2-Port 10GbE RoCE SFP+ adapter
 - IBM EC28 PCIe 2-Port 10GbE RoCE SFP+ adapter

- o Mellanox
 - ConnectX-2 EN (fw-ConnectX2 Rev 2.9.1200)
 - ConnectX-3 EN (fw-ConnectX3 Rev 2.11.0500)

1.4 Switches Supported

This release was tested with switches and gateways provided by the following companies:

InfiniBand Switches

- o Flextronics
 - F-X430044
- o Intel (formerly QLogic)
 - 12200
- o Mellanox
 - IS-5030
 - SX6025
 - SX6036

iWARP Switches

- o Fujitsu
 - XG2000C 10Gb Ethernet Switch

RoCE Switches

- o Arista
- o BLADE Network Technologies (BNT)
- o Mellanox
 - SX1036
 - SX1024
 - SX1016

1.5 Third Party Packages

The following third party packages have been tested with OFED 3.5:

- MVAPICH2 - 1.8
- Intel MPI - 4.1
- Open MPI - 1.6/1.6.1

1.6 OFED Sources

All sources are located under [git://git.openfabrics.org/](http://git.openfabrics.org/)

Linux:

URL: [git://git.openfabrics.org/compat-rdma/linux-3.5.git](http://git.openfabrics.org/compat-rdma/linux-3.5.git)

Branch: for-ofed-3.5

- Linux kernel sub-tree that includes files relevant for the OFED project only. Based on v3.5. Used to shorten git clone time.
- Note: the regular Linux git tree can be used as well.

compat:

URL: git://git.openfabrics.org/compat-rdma/compat.git

Branch: ofed

- Based on compat project (<https://github.com/mcgrof/compat>). The compat module provides functionality introduced in newer kernels to older kernels through a set of header files and exported symbols. See <https://github.com/mcgrof/compat/wiki> for details.
- Used to replace kernel_addons in the previous OFED kernel tree.

compat-rdma:

URL: git://git.openfabrics.org/compat-rdma/compat-rdma.git

Branch: master

User level Sources are downloaded from <http://www.openfabrics.org/downloads/> as written in the BUILD_ID

The kernel sources are based on Linux 3.5 mainline kernel. Its patches are included in the OFED sources directory. For details see HOWTO.build_ofed.

The list of maintainers is available under: <http://www.openfabrics.org/downloads/MAINTAINERS>

=====

2. Change log

=====

OFED-3.5 Main Changes from OFED 1.5.4.1 and 1.5.4

1. Based on linux-3.5 kernel
2. kernel-ib RPM renamed to compat-rdma
3. Removed MPI packages and tests
4. Removed SDP kernel module and libsdp, sdpnetstat
5. XRC is supported in kernel only
6. Removed support for the OSes:
 - RHEL 5.X
 - SLES 10 SPx
4. Updated packages: .

=====

3. Known Issues

=====
The following is a list of general limitations and known issues of the various components of the OFED 3.5-rc2 release.

1. When upgrading from an earlier OFED version, the installation script does not stop the earlier OFED version prior to uninstalling it.
Workaround: Stop the old OFED stack (/etc/init.d/openibd stop) before upgrading to OFED 3.5 or reboot the server after OFED installation.
2. Memory registration by the user is limited according to administrator setting. See "Pinning (Locking) User Memory Pages" in OFED_tips.txt for system configuration.
3. Fork support from kernel 2.6.12 and above is available provided that applications do not use threads. fork() is supported as long as the parent process does not run before the child exits or calls exec(). The former can be achieved by calling wait(childpid), and the latter can be achieved by application specific means. The Posix system() call is supported.
4. The qib driver is supported only on 64-bit platforms.
5. IPoIB: brctl utilities do not work on IPoIB interfaces. The reason for that is that these utilities support devices of type Ethernet only.
6. In case uninstall is failing, check the error log and remove the remaining RPMs manually using 'rpm -e <rpms list>'.
The former can be achieved by calling wait(childpid), and the latter can be achieved by application specific means. The Posix system() call is supported.
7. On SLES11.x, set allow_unsupported_modules parameter to 1 in file: /etc/modprobe.d/unsupported-modules. Without this the modules will not load.
8. RDS is supported only on kernel.org 3.5 and not on any distro OS.

Note: See the release notes of each component for additional issues.