



OFA Logo Program Developments



#OFADevWorkshop

Presented by Bob Noseworthy, Technical Sherpa

University of New Hampshire's InterOperability Laboratory
(UNH-IOL)

Outline

- Evolution and Benefits of the OFA Logo Program
 - Overview
 - Recent updates
- Using the OFA Logo List
- Feedback needed:
 - Continued evolution
 - Virtualization, NVMe, IPv6, Increased Distro involvement
 - Additional ULPs of Interest
- Next challenges

UNH-IOL / OFA



- UNH-IOL:
 - Host of the OpenFabrics Interoperability Logo Group
 - >25 Years providing Interop/Conformance Test
 - 2013: 12th Recipient of the IEEE-SA Corporate Award
- Your Speaker: Bob Noseworthy
 - Technical Sherpa / Chief Engineer
 - 18 Years in Ethernet conformance test
 - (10Mbps-100Gbps)
 - Expertise in IEEE 802.3, 802.1
 - Oversaw IOL's brief iWARP Consortium





Today's OFA Logo Program



OpenFabrics Interop Logo Group



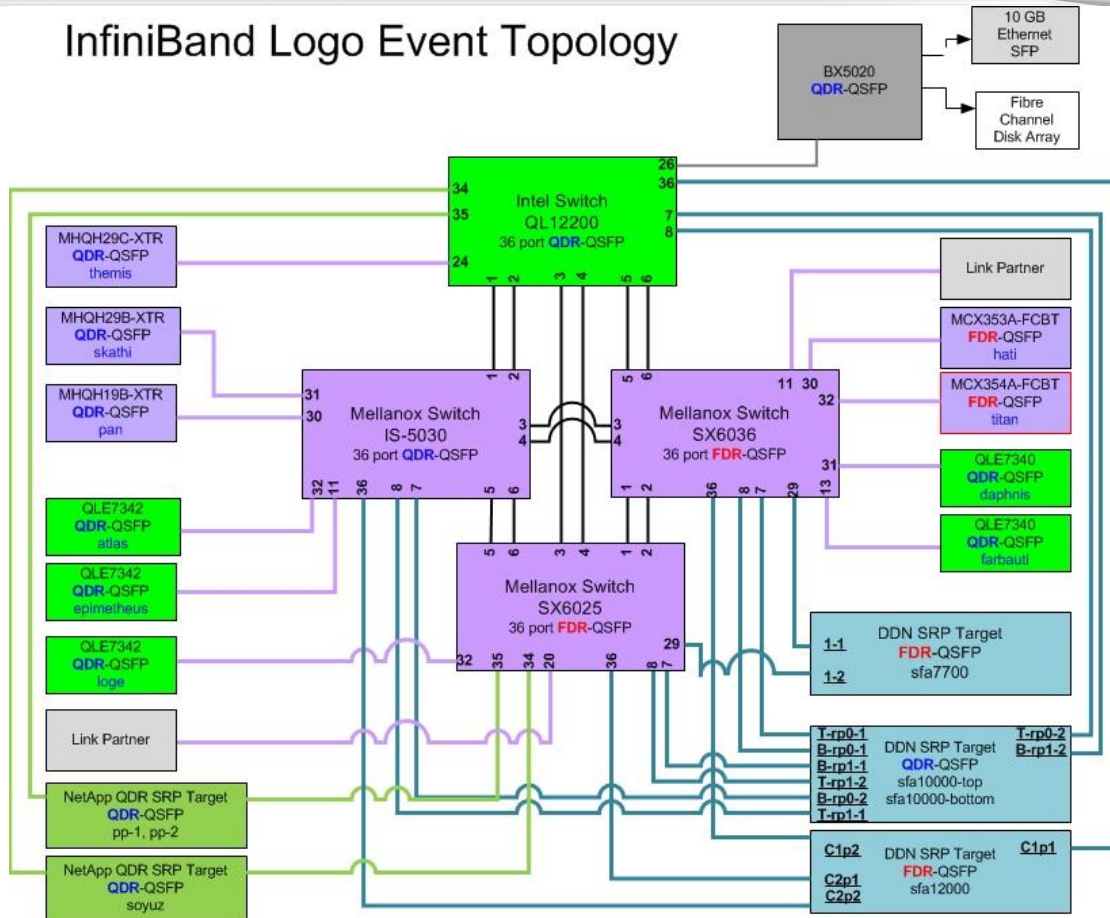
- OpenFabrics Interoperability Logo Group (OFILG)
 - **Purpose:** Validate OFED functionality, test ULPs and verify interoperability in a heterogeneous environment
 - **Current Members:**
 - Chelsio, DDN, Emulex, IBM, Intel, Mellanox and NetApp
- Validating IB, RoCE and iWARP
- Current Upper Layer Protocols Tested
 - Fabric Init, IPoIB, Link Init, NFSoRDMA, Open MPI, RDMA Utilities, RSocket, SM failover, SRP, uDAPL (& optional iSER and RDS)
- OFA Logo Test Plan defined by OFA-IWG
 - driven by OFA member contribution

OFILG (2)

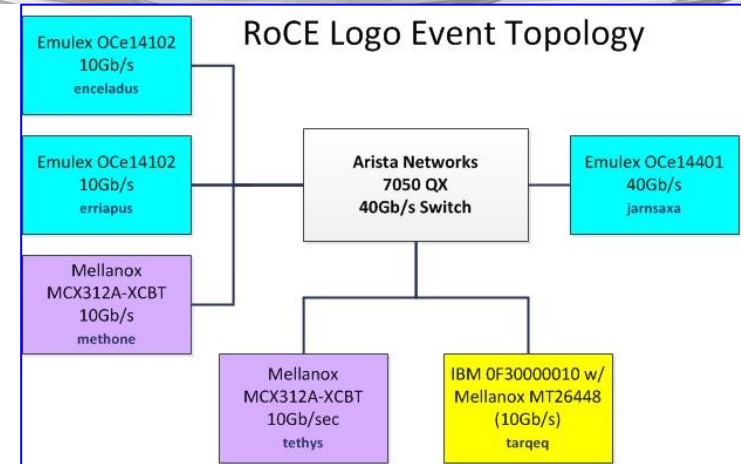
- OFA Cluster hosted at UNH-IOL
 - **Servers:** iWARP 12 hosts, InfiniBand 18 hosts, RoCE 15 hosts
 - **InfiniBand HW :** 12 HCAs, 4 switches, 5 SRP targets, 1 gateway
 - **iWARP HW:** 9 RNICs, 1 switch
 - **RoCE HW:** 6 RCA, 1 switch
- OFED versions Tested
 - 1.5.x, 3.5.x, 3.12
- PXE Boot environment available
 - Centos 6.x
 - Past Distros: RHEL 5.x and 6.x, SLES 11, Ubuntu 10.04 and 12.04
 - OFED 1.4.x, 1.5.x, 3.5.x, 3.12
- Highly scripted test environment
 - January 2014 Logo Event Tests executed – approximately 8,049

OFILG – Recent test topologies

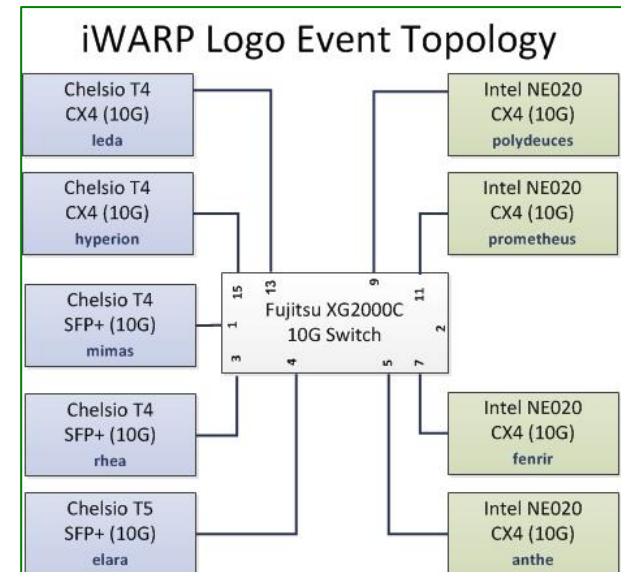
InfiniBand Logo Event Topology



RoCE Logo Event Topology



iWARP Logo Event Topology



OFA Logo Program (1/3)



- iol.unh.edu/ofilp
- P&P doc (Policy & Procedures)
- What can be on the Logo List?

OFA Logo Program (2/3)



- What can be on the Logo List? Additions of v1.16/v1.17 in bold
 - InfiniBand HCA
 - Ethernet R-NIC
 - **RoCE Adapter (RCA)**
 - InfiniBand Switch with Subnet Manager
 - InfiniBand Switch with no Subnet Manager
 - Ethernet Switches
 - **Ethernet DCB & Fabric Switches**
 - SRP Target/Server over InfiniBand
 - iSER Target/Server (over InfiniBand or iWARP **or RoCE**)
 - NFS-RDMA Client/Server (over InfiniBand or iWARP **or RoCE**)
 - Gateways (InfiniBand-to-Ethernet, InfiniBand-to-FibreChannel)

OFA Logo Program (3/3)



- What can be on the Logo List? Additions of v1.16/v1.17 in bold
 - Server Systems using InfiniBand HCA and running OFA software
 - Server Systems using InfiniBand HCA and running non-OFA software
 - such as for Sun Solaris, Apple Mac, HPUX, IBM AIX and other operating systems
 - Server Systems using Ethernet R-NIC and running OFA software
 - Server Systems using Ethernet R-NIC and running non-OFA software
 - such as for Sun Solaris, Apple Mac, HPUX, IBM AIX and other operating systems
 - Server Systems **using RoCE Adapters** and running OFA software
 - Server Systems **using RoCE Adapters** and running non-OFA software
 - such as for Sun Solaris, Apple Mac, HPUX, IBM AIX and other operating systems

Recent additions - RoCE

- First RCAs added May 2013
- End stations tested
 - RCA providers
 - System providers
- Anticipate return of iSER targets and additional system provider participation this year
- Bridges / Ethernet Fabrics
 - More on this in upcoming slides

The OFA Logo List (iol.unh.edu/ofilglist)



Example
from
2014 list:

InfiniBand Switches With Managers

Features Tested:

Link Initialization, Fabric Initialization, IPoIB, SM Failover, SRP, RDMA Operations, RDMA Stress, MPI

Manufacturer	Model Number	Ports	Speed	HW	FW	Report
Intel	12200-CH01	36	QDR	3	7.2.0.1.1	Report
Mellanox	SX-6036	36	FDR	X2	9.1.9470	Report
Mellanox	IS-5030	36	QDR	X2	7.4.220	Report

InfiniBand Switches Without Managers

Features Tested:

Link Initialization, Fabric Initialization, IPoIB, SM Failover, SRP, RDMA Operations, RDMA Stress, MPI

Manufacturer	Model Number	Ports	Speed	HW	FW	Report
Mellanox	SX-6025	36	FDR	X2	9.2.0	Report

SRP Targets


Features Tested:

IB Link Initialization, IB Fabric Initialization, SM Failover, and SRP

Manufacturer	Model Number	Ports	Speed	HW	FW	Report
DataDirect Networks	SFA10000	8	QDR	NA	2.1.1.2.18319	Report
DataDirect Networks	SFA12000	4	FDR	NA	2.1.1.2.18319	Report
DataDirect Networks	SFA7700	4	FDR	NA	2.1.1.2.18319	Report
Netapp	Pikes Peak (E5400)	4	QDR	1.0	07.86	Report
Netapp	Soyuz (E5500)	4	QDR	1.0	07.86	Report

OFA Logo List Reports

- <http://iol.unh.edu/ofilglist>
- Most recent Logo List includes summary report of all testing



OpenFabrics Alliance
Interoperability Logo Group (OFILG)
OFED 3.5-2 Interoperability Report

UNH-IOL – 121 Technology Drive, Suite 2 – Durham, NH 03824 - +1-603-862-0090
OpenFabrics Interoperability Logo Group (OFILG) – ofalab@iol.unh.edu

To Whom It May Concern

Date: 25 March 2014
Report Revision: 1.0
OFED Version Tested: 3.5-2
Operating System on Compute Nodes: Scientific Linux 6.4

Enclosed is an interoperability overview of RDMA capable devices tested during the January 2014 Logo Event. Nineteen different devices spanning six device classes and three RDMA transports were tested during this Event. Of the nineteen devices tested during this Logo Event, fifteen were found to pass the mandatory interoperability tests required for the OpenFabrics Interoperability Logo. This document is meant to highlight the performance of member companies' products while alerting the industry to problems that were uncovered with OFED 3.5-2.

The test suite referenced in this report is available at the IOL website. Release 1.49 (2013-Nov-5) was used.

<https://iol.unh.edu/ofatestplan>

The following table outlines the overall status of OFED 3.5-2 Interoperability across iWARP, RoCE, and InfiniBand transports as per the Test Plan referenced above.

Test Procedures	iWARP	RoCE	InfiniBand
11.1: Link Initialization	No OFED Issues Discovered	No OFED Issues Discovered	No OFED Issues Discovered
11.2: IB Fabric Initialization	Not Applicable	Not Applicable	No OFED Issues Discovered
11.3: iPoB Connected Mode	Not Applicable	Not Applicable	No OFED Issues Discovered
11.4: iPoB Datagram Mode	Not Applicable	Not Applicable	No OFED Issues Discovered
11.5: SM Failover and Handover	Not Applicable	Not Applicable	No OFED Issues Discovered
11.6: SRP	Not Applicable	Not Available	No OFED Issues Discovered
13.1: TI iSER	Not Available	Not Available	Not Available
13.2: TI iVFS-over-RDMA	Not Available	OFED Issue Bug 2449	No OFED Issues Discovered
13.4: TI uDAPL	No OFED Issues Discovered	No OFED Issues Discovered	No OFED Issues Discovered
13.5: TI RDMA Basic Interop	OFED Issue Bug 2457	OFED Issue Bug 2457	OFED Issue Bug 2457
13.6: TI RDMA Stress	No OFED Issues Discovered	Not Tested	No OFED Issues Discovered
13.8: TI MPI – Open MPI	No OFED Issues Discovered	No OFED Issues Discovered	No OFED Issues Discovered

OFILG OFED 3.5-2 Interoperability Report

Executive Summary

The goal of the OpenFabrics Interoperability Logo Program is to give assurance to customers, in the form of the OpenFabrics Interoperability Logo, that products on the [Logo List](#) using the OpenFabrics Enterprise Distribution will be interoperable with each other. The grant of the OpenFabrics Interoperability Logo signifies that a product has successfully passed all Mandatory tests during a test event pair and further signifies that the product is interoperable with all other products on the Logo List. By examining the Logo List, customers can be assured that products selected from the List will be compatible with each other, saving deployment time and resources. Each individual product report makes note of any issues that were discovered with the product while executing the test plan.

The Logo Program tests the multi-vendor interoperability of products that employ Remote Direct Memory Access (RDMA) using the OpenFabrics Enterprise Distribution (OFED). The Logo Program is split into two test event pairs:

- The Interop Debug Event
- The Interop GA Event

The Interop Debug Event is held at the UNH-IOL in early April and again in early October. The goal of the Interop Debug Event is to have representatives from OFILG member companies present while testing is being performed so that any and all issues discovered with either the member company's product(s) or OFED at large are reported and resolved in a timely manner. Upon completion of the Interop Debug Event, our members have thirty days to submit bug fixes and firmware updates to enable their products to interoperate with link partners from other companies.

Upon release of the General Availability (GA) release of OFED, the Interop GA Event starts. The goal of the Interop GA Event is to verify multi-vendor interoperability of our members' products using the latest available firmware and the GA release of OFED. The requirements for the OpenFabrics Interoperability Logo are defined in the [Test Plan](#) and [Logo Program](#). Additional general information about the program can be found at our [website](#).

Grant of the OpenFabrics Interoperability Logo signifies that a product has undergone 3-5 months of testing and validation, and that the member companies involved are committed to the multi-vendor interoperability of their respective RDMA technologies and the OpenFabrics Enterprise Distribution. During the January 2014 Interop Logo Event, 8,409 individual command line tests were performed, in addition to the testing done in the months after the Interop Debug Event leading up to the GA Event.

University of New Hampshire
InterOperability Laboratory (UNH-IOL) 2 OpenFabrics Interoperability Logo Group
Report Revision: 1.0

Using the OFA Logo List

- How do you use it?
- Do you also use the IBTA Integrator's List?
 - For Cables?
 - For Equipment (HCAs, Switches)?
- Is the product you are evaluating on the list?
- What limits your usage?
 - Validation Depth?
 - Validation Breadth?
 - More ULPs covered, different Distros validated
 - Participating Vendors/Product variety?



Tomorrow's OFA Logo Program?



Logo Program Next Steps

- Input needed
 - Set priorities
 - End User Stories / Needs
- Any OFA Member may participate in the OFA's Interoperability Working Group (OFA-IWG):
 - e.g:
 - End Users;
 - Application Developers;
 - Purchasers and Evaluators;
 - share your needs and desires with the OFA-IWG
- Only OFILG Members can receive logos for their products
- *The following slides are my speculation and are **NOT** active work within the OFA-IWG, if you support some of the ideas, please get involved !!!*

RoCE Bridge Testing

- Ethernet Bridges tested
 - None formally, yet play a critical role in CE
 - Emulex provided an Arista Networks 40GbE switch for current RoCE testing
- Put the “Fabric” in Ethernet Fabric validation
- Ethernet Fabrics are real
 - Congestion Avoidance (Data Center Bridging)
 - Well-beyond Multiple Spanning Tree
 - Proprietary examples: Qfabric, MLAG, FabricPath
 - Standards-based: IETF TRILL, IEEE SPB

Virtualization

- Support for hardware pass-thru for RDMA
- Which Hypervisor solutions?
 - Must be OFILG member driven
- Which OSes within the VM
 - Again: OFILG member driven
- Virtualization providers encouraged to directly participate in OFILG

Distro/OS provider Involvement



- Today using Scientific Linux, or Centos
- Distro/OSes tested in previous Interop events:
 - Canonical's Ubuntu
 - Microsoft (WinOfed)
 - RedHat
 - SUSE
- Distro/OS providers encouraged to directly participate in OFILG
- Distro/OS choice is at the direction of OFILG and OFA-IWG members
 - (as is the entirety of the OFA Logo Test Plan)

Speed

- IB at FDR / EDR
- Ethernet at 40 / 100 Gbps
 - 4x10Gbps or 10x10Gbps
 - Soon 4x25Gbps
- Testing at Higher Speeds / lower latencies critical for validation of some ULPs
- Requires
 - PCIe gen3 / gen4
 - Continuously challenging OFA test-cluster resources
 - Credit to AMD, HP, Intel for their contributions to the cluster over the years to support ever faster technologies.

Speed begets Performance

- Should OFA Logo Program expand to include Performance requirements?
 - Discussed at length during formation process
 - Currently Logo Program validates functionality and interoperability, but does not restrict based on performance
 - The Market is currently left to make that determination
 - Should a “minimum bar” be set for some of the validation
 - e.g: A fixed number or some percentage below a moving average of performance numbers from past events

- Further pushing the envelop of speed
- UNH-IOL NVMe Consortium
 - Working in concert with the NVM Express Organization
- Integrator's List
 - <https://www.iol.unh.edu/services/testing/NVMe/integratorslist.php>
- Validation of storage protocols / solutions at NVMe speeds
 - eg: EMC XtremeIO, Fusion-IO, DDN Storage Fusion Xcelerator (SFX) solutions, etc.

- UNH-IOL IPv6 Consortium / IPv6 Forum
- IPv6 Ready Logo List
 - <https://www.iol.unh.edu/services/testing/ipv6/logoholders.php>
- USGv6 Certification
 - <https://www.iol.unh.edu/services/testing/ipv6/usgv6tested.php>
- Not done today:
 - Validation of IPv6 connectivity in an RDMA OFED environment
 - IPoIB
 - IPoCE
 - iWARP-o-v6 ?

Still more growth options

- Scalability
 - Today, only functional testing is done (homogeneously and heterogeneously)
- Improved Fail-Over validation
 - HCA to HCA Failover / RCA-to-RCA Failover
 - requested via OFA RDMA Programming & System Admin Classes
- Additional ULPs
 - Lustre
- Or...?
 - Contribute your ideas to the OFA-IWG discussion

Interop vs Conformance

- OFILG focus:
 - Interoperability validation of OFED and ULPs with underlying RDMA transports
 - Conformance validation is currently not required to participate in the logo testing.

Interop v Conformance (2)

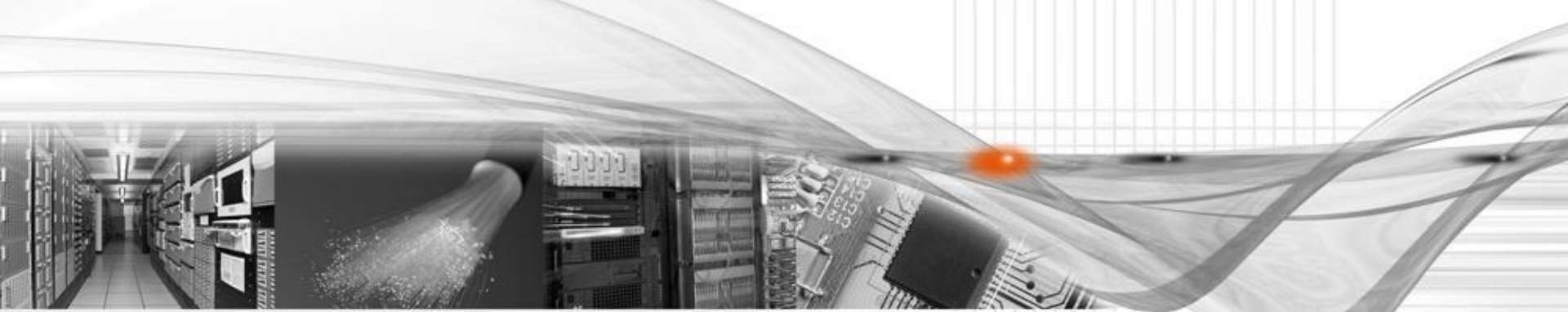
	Interop	Conformance	Logo / I.L./Cert.	Notes
IB	OFA	IBTA	OFA Logo, IBTA I.L.	
iWARP	OFA	None	OFA Logo	Past UNH-IOL Conformance test
RoCE	OFA	None	OFA Logo	Future IBTA I.L. ?
DCB	UNH-IOL	UNH-IOL	None	OFA Logo:Add subset?
TRILL	UNH-IOL	UNH-IOL	None	OFA Logo:Add subset?
SPB	UNH-IOL	UNH-IOL	None	
IPv6	UNH-IOL	UNH-IOL	IPv6 Ready / USGv6	OFA Logo:Add subset?
NVMe	UNH-IOL	UNH-IOL	UNH-IOL I.L.	OFA Logo:Add subset?
PCIe	UNH-IOL	PCI-SIG/ UNH	PCI-SIG Workshops	
40/100GE	UNH-IOL	UNH-IOL	None	
Cables (IB)	None	IBTA	IBTA I.L.	
Cables (10-100GE)	UNH-IOL	UNH-IOL	None	

Next Steps: Much to discuss

- Far more to consider than this brief presentation and discussion can touch on
- Continue the discussion with the:
Open Fabrics Interoperability Working Group
(OFA-IWG)
- Survey
 - https://docs.google.com/forms/d/1qwGuWe4rsFvYM6nyJVwa4q12_bWDiLIfRFi3p0EGV_w/viewform
 - Same link as above: <http://tinyurl.com/2014OFADev>

How to get involved...

- OFA Interoperability Working Group (IWG)
 - Every other Tuesday, 1pm ET
 - **Tuesday, April 08, 2014, 10:00 AM US Pacific Time**
916-356-2663, 8-356-2663, Bridge: 1, Passcode: 7004471
- Join the OFA-IWG Mailing List
 - <https://www.openfabrics.org/index.php/working-groups/wg-mail-subscription.html?view=form>
- Links:
 - *OFA Interop Program Overview:*
<https://www.openfabrics.org/index.php/ofedofaw.html>
https://www.openfabrics.org/images/docs/LinkedDocs/UNH_IOL_OFA_11-11_MR_FINAL.pdf
 - *OFA Logo List:* <http://iol.unh.edu/ofilglist>
 - *OFA Logo Program:* <http://iol.unh.edu/ofilp>
 - *OFA Test Plan:* <http://iol.unh.edu/ofatestplan>



Thank You

- Contact:

Bob Noseworthy (ren@iol.unh.edu)

UNH-IOL OFA Group (ofalab@iol.unh.edu)



And in other news: Training



Next OFA Training courses @ UNH-IOL

- May 19-20th: Writing Application Programs for RDMA using OFA Software
 - <https://www.openfabrics.org/index.php/resources/training/training-offerings.html>
- May 21-22rd : Infiniband Fabric Administration
 - <https://www.openfabrics.org/index.php/resources/training/infiniband-fabric-administration.html>
- On-site training can also be arranged
 - please contact Rupert Dance at rsdance@soft-forge.com