OPENFABRICS ALLIANCE

10TH ANNUAL INTERNATIONAL DEVELOPER WORKSHOP

OFED for Linux Status and Plans

Authors:
Robert J Woodruff; Rupert Dance

Date: 4/1/2014

#OFADevWorkshop

Agenda

- General Goals and Charter
- EWG – OFED for Linux status update
  - OFED-3.5-x
  - OFED-3.12
- Process Review and Changes
- OFED H/W and S/W validation
- OFED Roadmap
- Open Discussion Items
- How to contribute

EWG General Goals and Charter

- The charter of the EWG working group is to provide enterprise ready distributions of the Open Fabrics code for Linux
  - Includes providing backports to support several Linux kernel versions and Linux distributions
  - Includes comprehensive testing, validation, and hardening of the code
  - Includes software packaging, release notes, and software installers to allow for easy installation
  - Includes processes for bug tracking and problem resolution
  - Provides platform for experimental OFA technologies
  - Used for interoperability testing and OFA logo program

EWG – OFED Linux Status Update

- Production releases done since last year:
  - OFED 3.5-1
  - OFED 3.5-2
  - OFED 3.12 RC 1
- Experimental feature releases done since last year
  - OFED-3.5-MIC
  - OFED-3.5-1-MIC
  - OFED-3.5-2-MIC

OFED 3.5-1

- Released in September 2013
- Main new features:
  - Added support for RHEL EL 6.4

OFED 3.5-2

- Released in December 2013
- Main new features:
  - Added support for SLES 11 SP 3
  - Removed packages
  - - compat-dapl (older uDAPL 1.0)
  - Updated packages: dapl-2.0.39, ibacm-1.0.8, infiniband-diags-1.6.2, infinipath-psm-3.1-4, libcxgb4-1.3.1, libibmad-1.3.10, libibverbs-1.1.7, libmlx4-1.0.5, libmthca-1.0.6, librdmacm-1.0.17.1, mstflint-3.0-0.6, opensm-3.3.16, perftest-2.0-0.58

OFED 3.12 RC 1

- Based on kernel.org 3.12 kernel
- OSes supported
  - RHEL EL 6.4, RHEL EL 6.5
  - SLES 11 SP 3
  - Kernel.org 3.12
- Main new features:
  - Mellanox Connect-IB (mlx5) support
  - Emulex RoCE NIC (ocrdma) support
  - Updated user-space packages
    - dapl-2.0.40, ibsim-0.6, infiniband-diags-1.6.4, infinipath-psm-3.2-2_ga8c3e3e_open, libcxgb4-1.3.2, libibmad-1.3.11, libibumad-1.3.9 libipathverbs-1.2.1, libnes-1.1.4, librdmacm-1.0.18, mstflint-3.5.0 opensm-3.3.17, perftest-2.0-0.80.g54c73c6, srptools-1.0.2

OFED 3.5-x MIC

- Experimental branch releases to support running OFED on Intel® Xeon Phi ™
  - OFED-3.5-MIC (released December 2013)
    - For Intel® Xeon Phi™ MPSS 2.1
  - OFED-3.5-1 beta (based on OFED-3.5-1)
    - Since OFED-3.5-2 was released, OFED-3.5-1 was discontinued to upgrade to OFED-3.5-2 as base.
  - OFED-3.5-2 (currently under development)
    - Support for Intel® MPSS 3.1.x and later on RHEL EL 6.3,6.4, 6.5 and SLES 11 SP 2 and SLES 11 SP3
    - Added support for Mellanox Connect-IB (mlx5)
    - Added support for kernel mode Intel® Xeon Phi™ clients, e.g. IPoIB and Lustre from on the Intel® Xeon Phi™
    - Performance enhancements for Intel® Truescale HCAs

OFA Process Changes

- Starting with OFED-3.5, the code is based on upstream kernel and user-space packages.
  - Only fixes accepted upstream are included
  - Moved to new compat-rdma method for backports
  - MPI no longer included
  - Changes were designed to allow better alignment with distro and upstream code base to reduce fragmentation
- Companies can provide experimental features based on OFED.
  - e.g. OFED-3.5-x-MIC

OFA Process Changes

- As a result of new process changes, OFED and Linux distribution releases are now much closer aligned.
  - e.g. – Next SLES release (SLE12) will be based on kernel.org 3.12 and upstream user-space packages
  - RHEL EL 7 based on 3.10 kernel and upstream user-space packages
- Experimental branch release features
  - Original process was to allow vendors to make branch releases for new non-upstream experimental features, e.g. OFED-3.5-x-MIC
  - However, feedback is that is causing more fragmentation of OFED, not less, which was the original goal of the process changes.

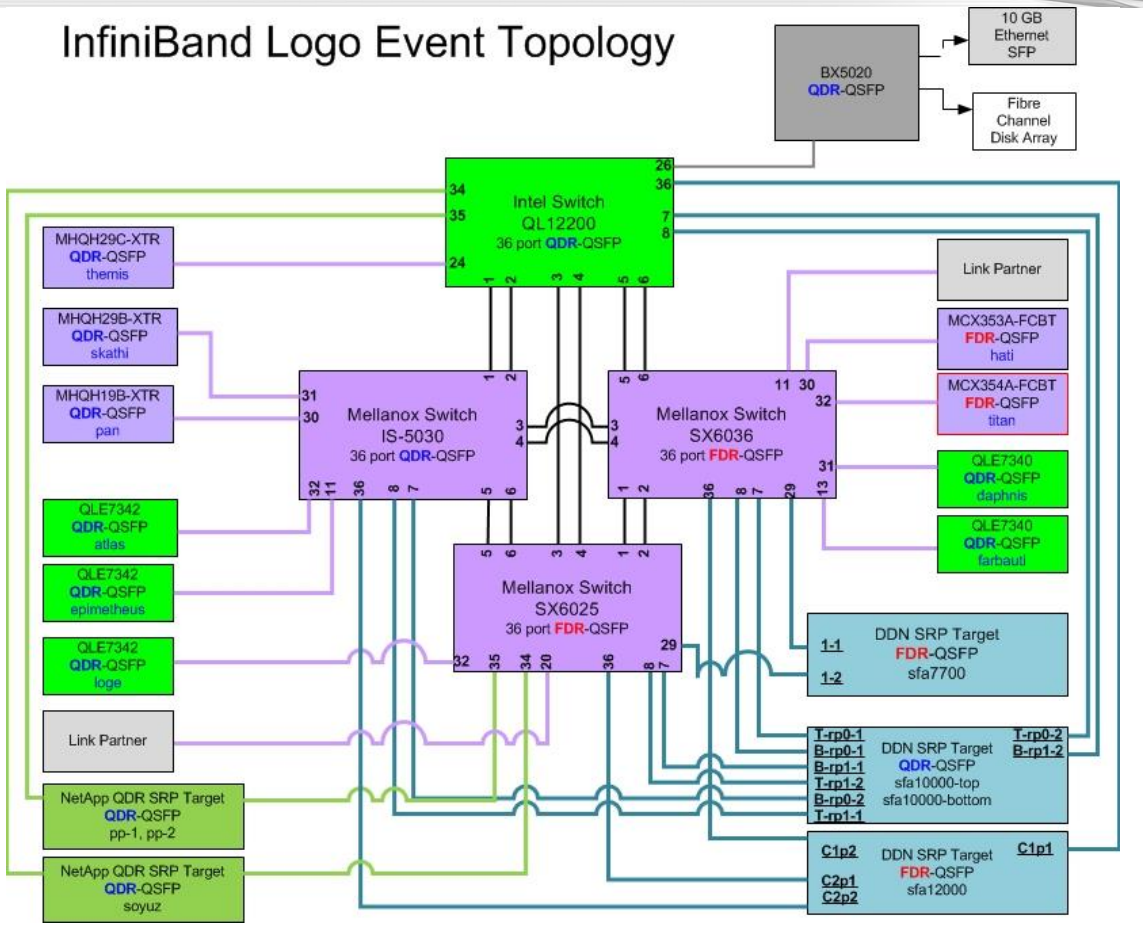
OFA Process Changes

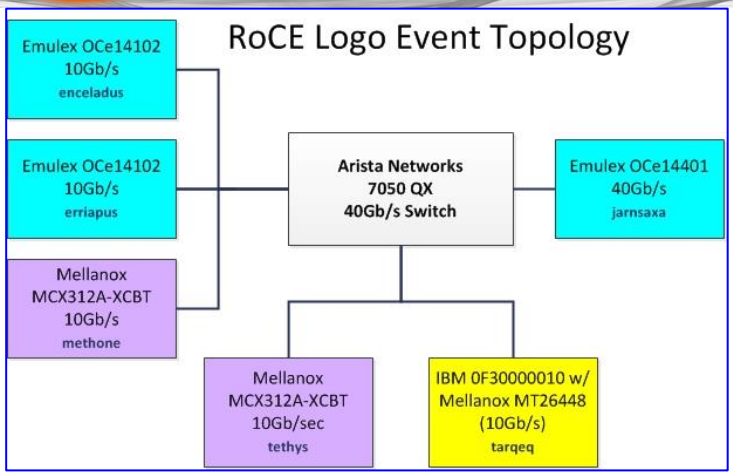
- Experimental branch release features (cont.)
  - Thus we are modifying the process to allow new experimental features be included in base OFED
  - Reduce OFED fragmentation by having a single OFED with both production and experimental code
  - Experimental code not enabled by default and clearly marked as a non-upstream feature
  - Similar to what kernel.org has for experimental code
  - Will start this new process after OFED-3.12

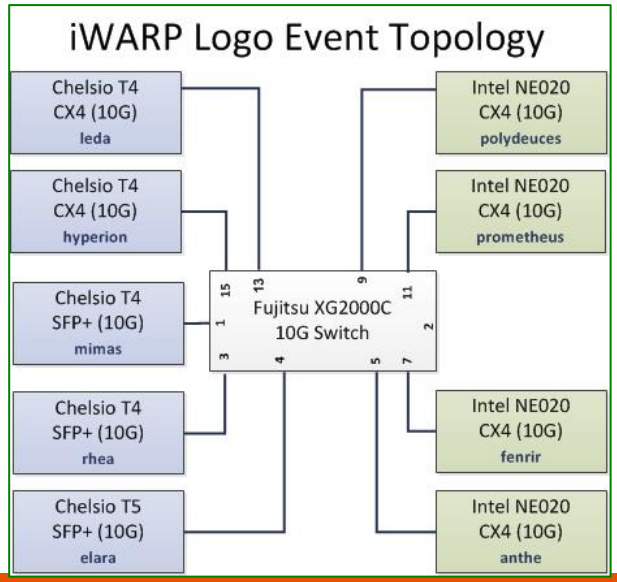
OFED – SW & HW validation

- OpenFabrics Interoperability Logo Group (OFILG)
  - **Purpose**: validate OFED functionality, test ULPs and verify interoperability in a heterogeneous environment
  - **Members**: Chelsio, DDN, Emulex, IBM, Intel, Mellanox and NetApp
- OFA Cluster 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
  - 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
- Protocols Tested
  - Fabric Init, IPoIB, Link Init, NFSoRDMA, Open MPI, RDMA Utilities, RSockets, SM failover, SRP, uDAPL
  - Tests executed – approximately 8,049

OFED – test topologies







OFED Roadmap.

- OFED-3.12-1
  - Add Intel® Xeon Phi™ support as an experimental feature
  - Add support for new distros, i.e. RHEL EL 7
  - Updated user-space packages and cherry picked bug-fixes from later upstream kernel.org
  - Release timing depends on RHEL GA release

Open Discussion Items

- When should we do the next major OFED release ?
  - What kernel should we base it on ? 3.16, 3.17 ?
  - What new features are being developed that would drive us to rebase to a new kernel.org kernel ?
    - OFIWG libfabric 1.0 release, when will that be ready to include in OFED ?
    - Are there other new features being developed that will need never kernel.org kernel support ?

If You Want to Help….

- Developing code:
  - Including back-ports in Linux
  - Reviewing code submitted to Linux kernel/libs
- Doing QA and testing
- Performance tuning
- Sending patches and comments to the mailing lists:
  - **OFED for Linux: ewg@lists.openfabrics.org**
  - **General Linux development: linux-rdma@vger.kernel.org**
  - **Maintainers and git trees:** http://www.openfabrics.org/downloads/MAINTAINERS
- Participate in EWG meetings
- Opening bugs in Bugzilla (http://bugs.openfabrics.org/)
  - When opening a new bug you can choose OpenFabrics Linux

Thank You

OPENFABRICS ALLIANCE

10TH ANNUAL INTERNATIONAL DEVELOPER WORKSHOP