

12th ANNUAL WORKSHOP 2016

INTEL OMNI-PATH ® FABRIC OPEN SOURCE OVERVIEW

John Fleck

Intel

[April 6th, 2016]



Introduction



Highly Leverages existing Aries and Intel[®] True Scale technologies





Open Source software and supports standards like the OpenFabrics Alliance*



Robust Ecosystem of trusted computing partners and providers

INTEL OMNI-PATH ® ARCHITECTURE AND THE OFA COMMUNITY

Open Sourcing software is not just throwing code on GitHub

- Phase 1 Project
 - Define the project
 - Build the contribution base
- Phase 2 Product
 - Definition and packaging
 - Identify customer pain points
 - · Work to enhance the project
- Phase 3 Proliferation
 - Established presence
 - Community development
- What does the community want?
- How can we collaborate with members?

INTEL OMNI-PATH ® ARCHITECTURE OPEN SOURCE APPROACH

Leverage: OpenFabrics Alliance Infrastructure

- Off-the-shelf application compatibility
- Extensive set of upper layer protocols
- Integrate scalable PSM for HPC

Access: Open Source Key Elements

- Push new software to kernel.org
- Push Intel® Omni-Path management elements to github

Channels: Integrate into Linux Distributions

Work with Red Hat and SuSE

OPEN SOURCE IMPLEMENTATION



INTEL OMNI-PATH ® ARCHITECTURE ECOSYSTEM



Fabric Management

- Traditional SM responsibilities
- Extended TrueScale
- SM, SA, PM, PA, FE
- https://github.com/01org/opa-fm



OPA Tools

User space tools

- Fabric bring up
- Fabric debug

RPM Packages for...

- Management node
- Compute node
- Address resolution
- MPI benchmarks
- https://github.com/01org/opa-ff
- <u>https://github.com/01org/opa-mpi-apps</u>



Three RPMs built from this package

Basic Tools RPM

sma and pma query tools

Fast Fabric RPM

- opatop
- opareport

Address resolution

- Plugin for ibacm
- Library to share path records with PSM

MPI APPs

Benchmarks and fabric verification

PSM2

- PSM library for OPA
- High performance MPI data path
- Supports many ranks
- https://github.com/01org/opa-psm2



Verbs Plugin

- libibverbs plugin
- Enables our verbs interface
- Requires no changes to libibverbs
- <u>https://github.com/01org/opa-libhfi1verbs</u>



INTEL OMNI-PATH ® ARCHITECTURE OPA FM GUI

- FM GUI
- Java based application
- Supports Fabric Monitoring
- Interfaces the Fabric Manager via the FE
- Provide Windows and Linux packages
- https://github.com/01org/opa-fmgui



OFA Community Elements

OpenMPI

- Pushed changes for PSM2 support
- Accepted
- Fully supported in 1.10.2 release

Libibumad

Added support for user space RMPP

Mvapich2

• PSM2 support found in the 2.2rc1 release track.

openshmem / GASNet

- Pushed changes to support PSM2
- Found in GASNet 1.26.0 release

ibacm

- Worked to get a plugin architecture
- Currently have a plugin as part of FF tools
- Added kernel netlink support

Kernel Elements

Iinux-firmware

• Four pieces of firmware

RDMA Core Kernel Patches

- Fully accepted 2K MAD changes in 4.2 kernel
- Enhancements to IPoIB for scalability
- Added netlink interface to ibacm
- Allows ULPs to access the path record cache
- Accepted in the 4.3 kernel
- Added rdmavt, common software before

HFI1 Driver

- Currently in staging
- Patches accepted in 4.6 kernel

INTEL OMNI-PATH ® ARCHITECTURE HOSTED REPOSITORIES

Current Locations for Intel Omni-Path ® Architecture software

FM and Tools centric

- <u>https://github.com/01org/opa-fm</u>
- <u>https://github.com/01org/opa-ff</u>
- <u>https://github.com/01org/opa-mpi-apps</u>

PSM2 library

<u>https://github.com/01org/opa-psm2</u>

Verbs plugin

https://github.com/01org/opa-libhfi1verbs

Fabric Manager GUI

https://github.com/01org/opa-fmgui

INTEL OMNI-PATH ® ARCHITECTURE MAIL LIST

- Collaborate with the OPA project via the following mail list
- intel-opa@lists.01.org
- Sign up at the following link
- https://lists.01.org/mailman/listinfo/intel-opa
- Kernel changes use the linux-rdma mail list
- linux-rdma@vger.kernel.org

LEGAL DISCLAIMERS

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF INTEL PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT.

All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest Intel product specifications and roadmaps.

Any forecasts of goods and services needed for Intel's operations are provided for discussion purposes only. Intel will have no liability to make any purchase in connection with forecasts published in this document.

Cost reduction scenarios described are intended as examples of how a given Intel- based product, in the specified circumstances and configurations, may affect future costs and provide cost savings. Circumstances will vary. Intel does not guarantee any costs or cost reduction.

Results have been estimated or simulated using internal Intel analysis or architecture simulation or modeling, and provided to you for informational purposes. Any differences in your system hardware, software or configuration may affect your actual performance.

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer or learn more at intel.com.

Intel processor numbers are not a measure of performance. Processor numbers differentiate features within each processor family, not across different processor families: Go to: <u>Learn About Intel® Processor Numbers</u>

The products described in this document may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Copies of documents which have an order number and are referenced in this document, or other Intel literature, may be obtained by calling 1-800-548-4725, or go to: <u>http://www.intel.com/design/literature.htm</u>

Intel, Intel Xeon, Intel Xeon Phi[™] are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States or other countries.

Copyright © 2016, Intel Corporation

OPTIMIZATION NOTICE

Optimization Notice

Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel.

Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice.

Notice revision #20110804



12th ANNUAL WORKSHOP 2016

THANK YOU

John Fleck

Intel

