

A vibrant, colorful nebula or galaxy core with swirling patterns of blue, cyan, magenta, and orange, set against a dark space background with distant stars. The text "FUEL YOUR INSIGHT" is overlaid on the lower half of the image.

FUEL YOUR INSIGHT



SC16

FUEL YOUR INSIGHT

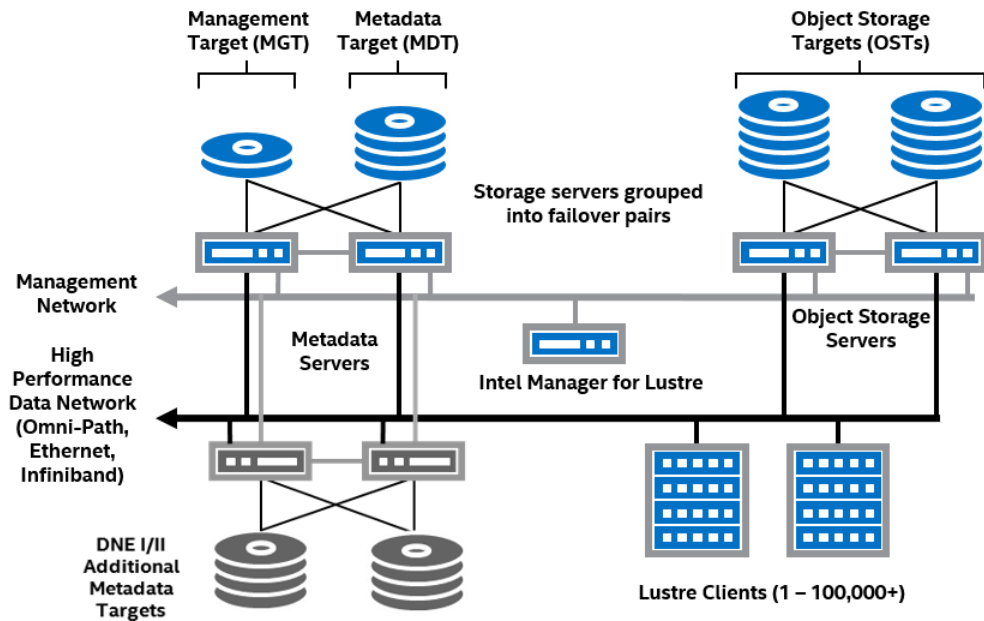
Lustre* & Intel® Omni-Path Fabric

Keith Mannthey, Lustre Solutions Architect
Intel High Performance Data Division

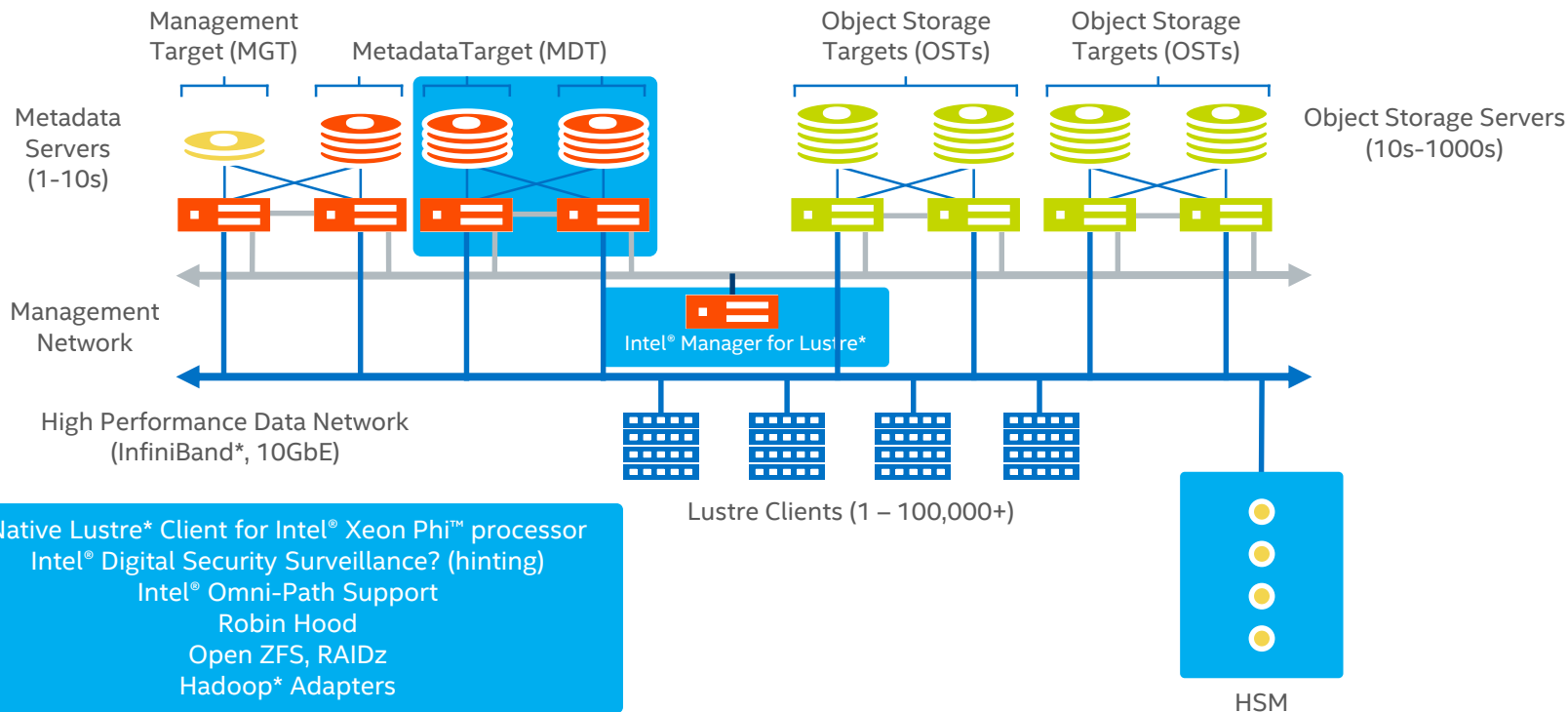
Agenda

- Lnet Introduction
- Intel® Omni-Path Architecture (OPA)
- Lnet Intel® OPA Considerations
- Lnet Intel® OPA Integration into Existing Fabrics

Lnet Introduction



Lnet Introduction



Lnet (continued)

- **Networks are given unique names**
 - o2ib0, tcp0, tcp1
- **Lustre Network Identifier (NID) defines interfaces**
 - 10.1.145.16@o2ib0
- **Includes native support for multiple networks**
 - Accomplished via the Lustre Network Driver (LND)
 - Infiniband via o2ib verbs interface, with RDMA support
 - Ethernet via TCP/IP interface
- **Lustre -> Network RPC API -> LNet -> LND -> Linux Driver**

Lnet

- **Designed to meet the needs of large-scale computing clusters**
 - Optimized for very large node counts, high throughput
- **Works with most networks, supports RDMA**
 - Omni-Path, Ethernet, Infiniband, XC/XT/XE, ELAN, Myrinet*, etc.
- **LNet is independent of the Lustre file system**
 - Abstracts network details from Lustre
 - Implemented as a set of kernel modules

Intel® Omni-Path Architecture

Building on the industry's best technologies

- Highly leverage existing Aries and Intel® True Scale fabric
- Adds innovative new features and capabilities to improve performance, reliability, and QoS
- Re-use of existing OpenFabrics Alliance* software

Robust product offerings and ecosystem

- End-to-end Intel product line
- Strong ecosystem with 70+ Fabric Builders members

HFI Adapters

Single port
x8 and x16



x16
Adapter
(100 Gb/s)



x8 Adapter
(58 Gb/s)

Edge Switches

1U Form Factor
24 and 48 port



48-port
Edge Switch



24-port
Edge Switch

Director Switches

QSFP-based
192 and 768 port



768-port
Director Switch
(20U chassis)



192-port
Director Switch
(7U chassis)

Silicon

OEM custom designs
HFI and Switch ASICs



HFI silicon
Up to 2 ports
(50 GB/s
total b/w)



Switch silicon
up to 48 ports
(1200 GB/s
total b/w)

Software

Open Source
Host Software and
Fabric Manager



Cables

Third Party Vendors
Passive Copper
Active Optical



Lnet Intel® OPA Considerations

- **Base OS Support**
 - Red Hat Enterprise Linux 7
 - SUSE Linux Enterprise Server 12
- **Lustre 2.7+ Required for OS Support Server Side**
- **Intel Fabric Suite (IFS) delivers OPA UPDATES**
 - IFS updates Base OS OFED components only as required
 - Enables the use of other in Kernel drivers concurrent with OPA

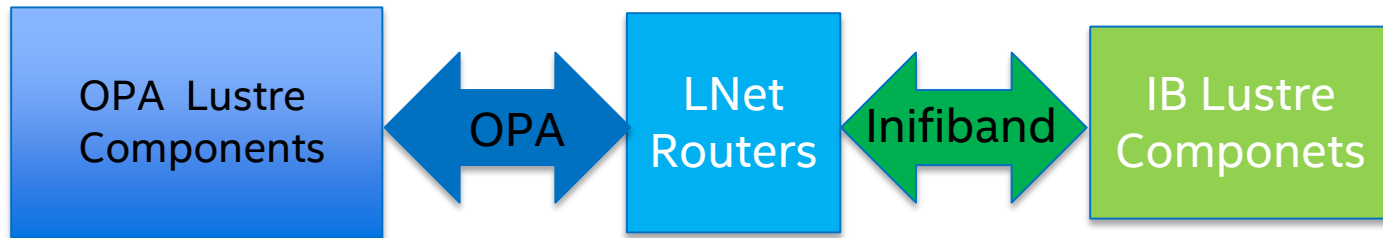
Lnet Intel® OPA Considerations (continued)

- **Intel® OPA gen 1 Supports RDMA Verbs in OFED**
 - Lnet uses the existing LND Infiniband Driver
- **Only LND and Driver TUNING required for operation**
 - Automated LND settings at Lnet install time

Integration into Existing Fabrics

Challenge : Intel® OPA is not directly compatible with Infiniband. Intel® OPA cannot plug into IB Switch

Solution: Lnet Router



Lnet Routers Overview

- Lnet Routers Lustre Software + Standard HW
 - Use off the shelf Hardware
 - Software is apart of standard Lnet/Lustre
 - Clustered Deployment Recommended
- Supported Configurations with Redhat* 7.2 and Lustre 2.7+
 - Intel® OPA -> Ethernet
 - Intel® OPA -> FDR (use in kernel drivers for IB)
 - Intel® OPA -> EDR (use in kernel drivers for IB)
 - See *Intel® Enterprise Edition for Lustre* software Configuration For LNet Routers*

Wrap up

- Lustre Intel® OPA is in production Today
- Lustre Routers provided Flexible Deployment Options
- Learn More
 - www.intel.com/Lustre



@IntelHPC | intel.com/SC16