Intel® Enterprise Edition for Lustre* Software

Version 3.0.1.0
Release Notes

Copyright 2016, Intel® Corporation
World Wide Web: http://www.intel.com
August 17, 2016

Contents:

- Product Overview
  - Intel® Enterprise Edition for Lustre* Software
  - Intel® Manager for Lustre* Software
  - Operating System Support
  - Features

- New Features and Updates in this Release
- Browser Compatibility and Installation Instructions
- Known Issues
- Partner Installation Support
- Learn More
- Disclaimer and Legal Information

Product Overview

Intel® Enterprise Edition for Lustre* Software

Intel® Enterprise Edition for Lustre* Software, when integrated with Linux, aggregates a range of storage hardware into a single Lustre* file system that is well proven for delivering fast IO to applications across high-speed network fabrics, such as InfiniBand* and Ethernet.

Intel® EE for Lustre* software is a global, single-namespace file system architecture that allows parallel access by many clients to all the data in the file system across many servers and storage devices. Designed to take advantage of the reliability features of enterprise-class storage hardware, Intel® EE for Lustre* software supports availability features, including

redundant servers with storage failover. Metadata and data are stored on separate servers to allow each system to be optimized for the different workloads.

Intel® Manager for Lustre* Software

Intel® Manager for Lustre* software is purpose-built to simplify the deployment and management of Lustre* file systems. Intel® Manager for Lustre* Software reduces management complexity and costs, enabling storage administrators to exploit the performance and scalability of Lustre storage, and accelerate critical applications and work flows.

Intel® Manager for Lustre* Software greatly simplifies the creation and management of Lustre file systems, using either the graphical user interface (GUI) or a command line interface (CLI). The software dashboard lets you monitor one or more distributed Lustre file systems. Near real-time storage-monitoring lets you track Lustre file system usage, performance metrics, events, and errors at the Lustre level. Plug-ins provided by storage solution providers enable monitoring of hardware-level performance data, disk errors and faults, and other hardware-related information.

Operating System Support

Note: Before using the Red Hat or RHEL software referenced herein, please refer to Red Hat's website for more information, including without limitation, information regarding the mitigation of potential security vulnerabilities in the Red Hat software.

Operating Systems	Lustre Servers	Lustre Clients	Intel® Manager for Lustre* Software Support
RHEL/Centos 6.8	No	Yes[1]	No
RHEL/Centos 7.2	Yes	Yes	Yes
SLES* 11sp4[2]	Yes	Yes	No
SLES 12sp1[2]	No	Yes	No
Ubuntu* 14.04 LTS	No	Yes	No

[1] The el6 client is distributed in the ee-contrib-3.0.1.0.tar.gz tarball with this release
[2] SLES packages are contained in the ee-contrib-3.0.1.0.tar.gz tarball with this release.

Features

GUI-based creation and management of Lustre file systems

The Intel® Manager for Lustre* Software provides a powerful, yet easy-to-use GUI that enables rapid creation of Lustre file systems. The GUI supports easy configuration for high availability and expansion, and enables performance monitoring and management of multiple Lustre file systems.

Graphical charts display near real-time performance metrics

Fully configurable color charts display a variety of real-time performance metrics for single or multiple file systems, down to individual servers and targets, and reveal metrics including OST bandwidth, OST balance, read/write operations per OST (heat map), file system capacity, metadata operations, bandwidth, and various resource usage parameters.

Heat Map Charting with Job Statistics

The Heat Map, displayed on the Dashboard page, shows the level of read/write activity for each OST in all file systems. OSTs are displayed as rows, and consecutive time intervals are displayed as columns. One can monitor the level of read or write activity for a given OST over time. The job statistics feature displays the top five read and write job statistics for a given OST and time interval. This feature employs the jobstats feature in Lustre, and supports the creation of plug-ins to display user account, command line, job size, and job start/finish times.

Auto-configured high-availability clustering for server pairs

Pacemaker and Corosync are configured automatically when the system design follows configuration guidance. This removes the need for manually installing HA configuration files on storage servers and simplifies high-availability configuration.

PDU configuration and server outlet assignments support automatic failover

The PDU tab lets you configure and manage power distribution units. You can add a PDU and assign specific PDU outlets to specific servers. When you associate PDU failover outlets with servers using this tool, STONITH is automatically configured.

IPMI and BMC Configuration

As an alternative to PDU configuration, support for Intelligent Platform Management Interface and baseboard management controllers support server monitoring, high-availability configuration, and failover.

Support for Intel® Xeon Phi™ Coprocessor Clients

Intel® EE for Lustre* client software can be installed and configured to run on Intel® Xeon Phi™ Coprocessor clients. This means that the Intel® Xeon Phi™ Coprocessor clients can directly mount Lustre.

Simplified ISO-less installation and automated deployment streamlines installation

The installation strategy removes the need to manually install Intel® EE for Lustre* software on each server. Rather, Intel® EE for Lustre* software is quickly installed on the manager server, and from there required packages are automatically deployed to all storage servers. Storage servers and the manager server can run the same standard operating system as the rest of your estate.

Support for OpenZFS in Monitor Mode

Intel® EE for Lustre* software supports ZFS as a back-end file system replacement for ldiskfs. While Intel® Manager for Lustre* software is able to configure and manage high-availability Lustre storage solutions, Intel® EE for Lustre* software can only discover and *monitor* ZFS file systems. Full, high-availability *management* of ZFS file systems using Intel® Manager for Lustre* Software is not supported.

Update OpenZFS/SPL version to 0.6.5

The OpenZFS software has been updated to the latest stable version from the ZFSonLinux project. This brings with it several performance and stability improvements. One of the most significant additions is the ability to set the dataset record size larger than 128KB, improving throughput IO for large files.

Intel® EE for Lustre* Software ZFS Snapshots

The OpenZFS file system provides integrated support for snapshots, a data protection feature that enables an operator to checkpoint a file system volume.

In Intel® EE for Lustre* software, version 3.0.0.0, Intel® has developed a mechanism in Lustre* that is capable of leveraging ZFS to take a coordinated snapshot of an entire Lustre* file system, if all of the storage targets in the file system are formatted using ZFS.

Hierarchical Storage Management

This release of Intel® EE for Lustre* software incorporates support for hierarchical storage management (HSM). HSM provides tiered storage by allowing data to be moved between the file system and secondary, archival storage. The HSM framework included with this release of Intel® EE for Lustre* software includes the POSIX copytool implementation, the worker agent, and the MDS Coordinator. HSM agents and copytool instances are configured via the Intel® Manager for Lustre* GUI.

Robinhood Policy Engine

The Robinhood policy engine, version 2.5.5, is included with Intel® EE for Lustre* software, and can be used to manage policies for an HSM solution. Intel® Manager for Lustre* software performs the provisioning of the Robinhood agent server, which is performed via the manager GUI. Robinhood can be used with the HSM capabilities described above to automate HSM archiving and report generation. Robinhood also can be used by administrators for gathering information about your Lustre file system, such as the ranked list of the oldest files.

Automated Provisioning of Custom Lustre Service Nodes

This feature allows users to create custom profiles for new Lustre client types and, based on a given profile, deploy and install custom code to provide new services. The HSM copytool (above) is deployed in this way.

Apache Hadoop* adapter software

Intel® EE for Lustre* software is supported by the Apache Hadoop* adapter software, however the adapter software is a separate download. This Hadoop adapter for Lustre is compatible with the Apache Hadoop software, versions 2.3 and 2.5 as of this writing. Hadoop software allows users who run MapReduce jobs to bypass storing data in HDFS, and store the MapReduce output directly to Lustre instead. This allows the analytical processes direct access to scientific output instead of transferring data from the compute cluster storage system to another file system. Optimizations have also been made to the shuffle step in MapReduce to take advantage of Lustre's high-speed network access to data. Many workloads will see an overall reduction in end-to-end processing time by using the Hadoop adapter with the Intel® EE for Lustre* software file system. For more information, see *Installing Hadoop, the Hadoop Adapter for Intel® EE for Lustre* Software, and the Job Scheduler Integration.*

The following table shows version compatibility with Apache distributions. This information may not be complete. Please contact your Intel® support representative for current information.

Hadoop Adapter Version	Target Hadoop Distributions	Java Version (Recommended)
hadoop-lustre-plugin-3.0.0	hadoop-2.3.0-apache hadoop-2.5.0-apache	Java Runtime Environment 1.7.0+

HPC Job Scheduler integration with MapReduce

Intel® EE for Lustre* software works with the HPC *job scheduler integration* with MapReduce; however the job scheduler integration is a separate download. The HPC job scheduler integration supports Apache Hadoop. This adapter for job schedulers allows you to integrate common resource schedulers into your cluster. You have the choice of installing the SLURM (Simple Linux Utility for Resource Management) job scheduler integration or the PBS (portable batch system) job scheduler integration. An integration guide is available*: Installing Hadoop, the Hadoop Adapter for Intel® EE for Lustre* Software, and the Job Scheduler Integration.*

Hadoop commonly uses Yarn to manage MapReduce jobs. However, virtually all HPC systems use a *job scheduler*, for example, SLURM. However, having two job schedulers, e.g., SLURM and Yarn, in a single system can cause problems. The HPC Job Scheduler integration with

MapReduce replaces YARN with an interface to the main resource manager for the system. This allows MapReduce applications to be run as normal HPC jobs. The following table shows version compatibility with Apache distributions and SLURM and PBS.

Target Hadoop Distribution	Hadoop Adapter for Lustre	HPC Scheduler Version	Job Scheduler Versions	Java Version (Recommended)
hadoop-2.5.0-apache	hadoop-lustre-plugin-3.0.0	hadoop-hpc-scheduler-3.0.0	SLURM 2.5.6 or PBS 2.5.6	Java Runtime Environment 1.7.0+

Apache Hive compatibility

Hive is a data warehouse infrastructure built on top of Hadoop for providing data summarization, query, and analysis. Intel® has tested the Hadoop adapter for Lustre provided with Intel® EE for Lustre* software for compatibility with Apache Hive version 2.5.

Apache Hbase compatibility

HBase is a non-relational, distributed database modeled after Google's BigTable and written in Java*. Hbase runs on top of HDFS (Hadoop Distributed File System). Intel® has tested the Hadoop adapter for Lustre provided with Intel® EE for Lustre* software for compatibility with Apache Hbase version 2.5.

Lustre 2.7.x

This release of Intel® EE for Lustre* software is based on the Intel® Foundation Edition for Lustre* 2.7 release tree, representing a major update to the underlying Lustre* version for the Intel® Enterprise Edition for Lustre* software (as of version 3.0.0.0).

Online Lustre File System Consistency Checks (LFSCK)

LFSCK is an administrative tool that was first introduced in Lustre* software release 2.3 for checking and repairing attributes specific to a mounted Lustre* file system. LFSCK is similar in concept to an offline FSCK repair tool for a local file system, but LFSCK is implemented to run as part of the Lustre* file system while the file system is mounted and in use.

LFSCK allows consistency checking and repair by the Lustre software without downtime, and can be run on the largest Lustre* file systems with negligible disruption to normal operations.

Distributed Namespace

Distributed Namespace (DNE) allows the Lustre metadata to be distributed across multiple metadata servers. Intel® EE for Lustre* software supports DNE1 (as of release 2.3.0.0), which supports the use of multiple MDTs. This enables the size of the Lustre namespace and metadata throughput to be scaled with the number of OSSs. This featured is supported at the Intel® Manager for Lustre* GUI.

DNE II Striped Directories Support (Preview)

Striped directories support (Distributed Name Space, phase 2) is available in Intel® EE for Lustre* software, as of version 3.0.0.0, as a technology preview. Striped directories allow operators to shard directory entries across multiple metadata storage targets, providing both namespace and metadata performance scalability.

Single Client Metadata Concurrency

Also referred to as “multi-slot last_rcvd”, this update to the metadata communications interface between client and server allows multiple metadata RPCs to be in flight in parallel, per-client for both read and write transactions. Prior to this release, any client RPCs that modified file system metadata (for example, creates or unlinks), were sent serially to the server. With this update, this restriction is removed.

Differentiated Storage Services

Differentiated Storage Services (DSS) allows I/O data to be classified, sometimes referred to as “hinting". These hints pass seamlessly through Intel® EE Edition for Lustre* software, at which point data can be tiered and intelligently cached by the storage system. This enables a more efficient use of cache space, and decreases the likelihood of critical data being evicted when the cache fills. Intel® is working directly with storage and cache vendors to enable DSS hinting in Lustre appliances, and to provide optimized performance to Intel® EE Edition for Lustre* software deployments with a mix of SSD and traditional storage.

Support for Intel® Omni-Path Architecture

Intel® Omni-Path fabric support is available for Intel® EE for Lustre* software systems running RHEL 7.2. Intel® OPA driver support requires RHEL 7.1 or newer, and so is not available for RHEL 6.x based systems.

LNet Configuration

This feature assists in configuring LNet for a given server’s network interface by setting the LNet network ID for that port. This feature requires a single LNet. You can configure multiple LNets (i.e., with the use of routers), however in this release, additional LNets cannot be configured from the GUI.

Dynamic LNet Configuration

Dynamic LNet configuration (DLC) is a powerful extension of the LNet software to simplify system administration tasks for Lustre networking. DLC allows an operator to make changes to LNet (for example, network interfaces can be added and removed, or parameters changed,) without requiring that the kernel modules be removed and reloaded. Parameters can be altered while LNet is still running, meaning that tuning and optimization can be conducted while Lustre* is still running on the target node. Dynamic LNet configuration also applies to LNet routers, so that routes can be added, removed and updated without affecting other Lustre network traffic.

Kerberos Network Authentication and Encryption

Kerberos provides a means for authentication and authorization of participants on a computer network, as well as providing secure communications through authentication. This functionality has been applied to Intel® EE for Lustre* software for the purposes of establishing trust between Lustre* servers and clients, and optionally, supporting encrypted network communications.

SELINUX* Lustre* Client Support

SELINUX* is a mature access-control platform for Linux* systems, was originally developed by the NSA, and is available in RHEL to enforce access control policies including Multi-Level Security (MLS). The implementation is currently restricted to enforcing SELINUX* policies on the Lustre* client.

New Features and Updates in this Release

This update incorporates revised Lustre packages that resolve the following tickets:

- EL6 Server support is removed from this release. The el6 client is supported and distributed in the ee-contrib-3.0.1.0.tar.gz tarball.
- SLES packages are now packaged in the ee-contrib-3.0.1.0.tar.gz tarball with this release.
- Ubuntu packages are now packaged in the ee-contrib-3.0.1.0.tar.gz tarball with this release.

Browser Compatibility and Installation Instructions

Intel® Manager for Lustre* Software is supported on:

- Google Chrome browser. Use the most current stable version.
- Mozilla Foundation Firefox browser. Use the most current stable version.

The Intel® Manager for Lustre* Software GUI is not supported on Internet Explorer*.

For information on installing or updating Intel® Enterprise Edition for Lustre* software and its components, including system requirements for the manager server and storage servers, please see the *Intel® Enterprise Edition for Lustre* Software Installation Guide*.

Note: If you plan to update the Linux operating system to 7.2, update the operating system first. Then update the Intel® EE for Lustre* software.

Known Issues

The following table lists those tickets that the partner and customer should be made aware of.

Ticket Number	Description
LU-7410	If downgrading back to the original EE 2.4.x release after having upgraded to an EE 3.x release, you must remount the MDT and OSTs with "abort_recovery" option, and unmount them with "-f" option to avoid getting an "unsupported incompat filesystem features" error.
HYD-1423	UI breaks when there are storage resources in the DB whose plugins are unavailable.
HYD-1500	File system client count freezes when the file system has stopped - so may show connected clients to a stopped file system.
HYD-1643	Changes in size of block devices are not detected.
HYD-1891	Removed storage devices still appear in Intel® Manager for Lustre* GUI until the agent or manager is restarted. Recommendation: If you remove a storage device from a server, you will need to restart the agent or Intel® Manager for Lustre* Software.
HYD-3232	The software makes a best attempt to configure iptables on the manager. However, if your system has a drastically different format than the default el6 iptables configuration, this configuration may fail. This will surface as an error during install. Recommendation: Use a fresh el6 iptables configuration when installing Intel® Manager for Lustre* software, or repair the configuration manually.
HYD-3280	Corosync blocks 10.10.10.x LNet addresses. After "add server" or "create file system" ping stopped working on the LNet interface. It seems corosync consumes the entire 10.x.x.x addr space. Recommendation: File system running Intel® EE for Lustre* software cannot be installed on Class A subnets (the /8 subnets, in CIDR notation). It is recommended that /16 or /24 subnets be used (Class B and C, respectively).
HYD-3811	Failover that occurs while the manager software is down is not detected when the manager software comes back up. Restarting the agent resolves this issue.

Ticket Number	Description
HYD-3937	Scanning ZFS FS with suspended pool results in error. Recommendation: In cases such that one or more of the ZFS pools in your ZFS based file system are in an suspended, degraded, or offline state, the software may not detect the file system. Any attempt may result in the software displaying an error message. Make sure your ZFS pools and file system are functioning properly before attempting to scan.
HYD-4248	On rare occasions, the Manager GUI will sometimes display a 502 proxy error. Recommendation: Reload the page using the browser reload button. No data or actions are lost because of this.
HYD-4250	In some situations, the GUI will display an Updates Required alert even immediately after an install or update. Recommendation: This alert should be cleared by repeating the update process.
HYD-4639	Executing failover and/or failback commands for multiple targets in an HA pair, in parallel, will leave location constraints configured in Pacemaker that Intel® Manager for Lustre* software expects to be removed. This will cause the removal of targets (and by extension the removal of a file system) to fail, and may cause future failover and/or failback actions to fail. Recommendation: At the GUI and the CLI, and when using the REST API, avoid executing failover and/or failback commands in parallel for multiple targets in an HA pair. Rather, execute those commands in serial, preferably with a pause of two seconds after issuing the preceding command.
HYD-5751	Storage plugin UpperBoundCondition always triggered. A storage plugin using the alert_conditions.UpperBoundCondition always has active alerts, no matter the value entered.

Ticket Number	Description
HYD-5802	When a server is removed that has file system targets still active on it, it will fail with an error. Recommendation: If the remove has not yet been performed, remove any file system targets running on the server prior to removing the server. If the remove has already been done and ended in error, you can recover by using the Server Detail page to start pacemaker, and then remove the file system targets and the server.
HYD-5866	Server removal can occasionally fail with the error "Error: Unable to update any nodes". Recommendation: Simply remove the server again, and it will successfully remove the server.

HYD-5887	Stopping the copytool at the Intel® Manager for Lustre* dashboard may hang on StopCopytoolStep.
	
	Recommendation: If this happens, enter the following two commands at the command line, in the order presented, to manually stop the copytool:
	
	`service chroma-copytool-monitor-<copytool-number> stop`
	
	`service chroma-copytool-<copytool-number> stop`

LDEV-337	When upgrading between versions of Intel EE for Lustre* software, if both the starting and ending version contain the exact same kernel version, the kernel will not be upgraded. In this case, you need to yum reinstall the kernel with the kernel from the newer version Intel® EE for Lustre* software on each Lustre server.

Partner Installation Support

For resellers experiencing problems during the initial install and validation of their configuration, please go to https://jira.hpdd.intel.com/browse/IEEL and create a support ticket. When reporting a problem, please be as specific as possible about its nature (installation, configuration, etc.), and include information about the hardware and any underlying software used in your configuration. Note that tickets filed at the above site are confidential and not publicly viewable.

Reasonable commercial efforts will be made to respond promptly to a support ticket; however, response times will vary. If you encounter any issues regarding your request, please contact your account manager.

Learn More

For related documentation and to learn more about Lustre, Intel® Enterprise Edition for Lustre* Software, and enterprise-class support and services, visit the Intel® Lustre* Partner Portal: http://www.intel.com/content/www/us/en/software/intel-solutions-for-lustre-software-reseller-portal.html

Disclaimer and Legal Information

Copyright 2016 Intel Corporation. All Rights Reserved.

Portions of the software contained or described herein and all documents related to said software (referred to herein as “Material”) are owned by Intel Corporation or its supplier or licensors. “

Title to the Material remains with Intel Corporation or its suppliers and licensors. The Material contains trade secrets and proprietary and confidential information of Intel or its suppliers and licensors. The Material is protected by worldwide copyright and trade secret laws and treaty provisions. No part of the Material may be used, copied, reproduced, modified, published, uploaded, posted, transmitted, distributed, or disclosed in any way without Intel’s prior express written permission. (Continued)

No license under any patent, copyright, trade secret or other intellectual property right is granted to or conferred upon you by disclosure or delivery of the Materials, either expressly, by implication, inducement, estoppel or otherwise. Any license under such intellectual property rights must be express and approved by Intel in writing.

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.

A "Mission Critical Application" is any application in which failure of the Intel Product could result, directly or indirectly, in personal injury or death. SHOULD YOU PURCHASE OR USE INTEL'S PRODUCTS FOR ANY SUCH MISSION CRITICAL APPLICATION, YOU SHALL INDEMNIFY AND HOLD INTEL AND ITS SUBSIDIARIES, SUBCONTRACTORS AND AFFILIATES, AND THE DIRECTORS, OFFICERS, AND EMPLOYEES OF EACH, HARMLESS AGAINST ALL CLAIMS COSTS, DAMAGES, AND EXPENSES AND REASONABLE ATTORNEYS' FEES ARISING OUT OF, DIRECTLY OR INDIRECTLY, ANY CLAIM OF PRODUCT LIABILITY, PERSONAL INJURY, OR DEATH ARISING IN ANY WAY OUT OF SUCH MISSION CRITICAL APPLICATION, WHETHER OR NOT INTEL OR ITS SUBCONTRACTOR WAS NEGLIGENT IN THE DESIGN, MANUFACTURE, OR WARNING OF THE INTEL PRODUCT OR ANY OF ITS PARTS.

Before using any third party software referenced herein, please refer to the third party software provider’s website for more information, including without limitation, information regarding the mitigation of potential security vulnerabilities in the third party software.

Intel may make changes to specifications and product descriptions at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked "reserved" or "undefined". Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information here is subject to change without notice. Do not finalize a design with this information.

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.

Contact your local Intel sales office or your distributor to obtain the latest specifications and before placing your product order. 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: http://www.intel.com/design/literature.htm.

Intel and the Intel logo are trademarks of Intel Corporation in the U.S. and/or other countries.

* Other names and brands may be claimed as the property of others.