

Layout Enhancement Scope Statement

signed-off

This document was signed-off by OpenSFS on 10th October 2013.

Introduction

The following scope statement applies to the Layout Enhancement project within the Technical Proposal by High Performance Data Division of Intel for OpenSFS Contract SFS-DEV-003 signed Friday 23rd August, 2013.

Problem Statement

The Lustre* software layouts represent how file data is distributed over OSTs. Only simple RAID0 (striping) layouts are supported today, and enhancements are required to implement future features such as Data on MDS, Data Replication, live Data Migration, and RAID1/5/6 or erasure coding. The Layout Enhancement (LE) project is therefore a prerequisite for the Data on MDS and Replication projects proposed here.

Project Goals

- Deliver a solution architecture.
- Deliver a high-level design.
- Identify the behaviour of inter-operating components.

In Scope

- Design an extension to the Lustre file system layout to accommodate the following features:
 - N-way stripe replication. This is a requirement for the Data Replication project. It generalizes RAID-0+1 to allow multiple replicas of the data in a Lustre file. This will ensure that data can remain available in the face of multiple OST failures - e.g. multiple short- or long-term OST failures.
 - RAID5/6 and erasure codes. These anticipate future developments to enable more space-efficient replication techniques.
 - Layout extents. This allows different layouts in different extents of a Lustre file. It can be used to enable wider striping as files grow in size, to prevent inconsistent `-ENOSPC` errors as individual OSTs become full, and to enable incremental migration, replication, and HSM restore.
- Determine the additional requirements for the CRUSH algorithmic layout.
- Extend the layout locking protocol to ensure layout changes can be effected while a file is actively being accessed.
- Determine client-side changes required to allow layout changes to be requested within the I/O stack.
- Determine a new protocol strategy for handling large layout representations that address the issue with the current strategy, which results in grossly inefficient network buffer utilization for extremely large layouts.
- Discuss compatibility and version interoperability.
- Address requirements to support rolling upgrades.

Out of Scope

- Code will not be delivered.

Project Constraints

John Hammond is the strongly preferred engineer for this project and John Hammond is also the strongly preferred engineer for the Dead Code Removal project.

Key Deliverables

- Solution Architecture.
- High-level design document.

- Implementation assessment.

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