

Striped Directories Scope Statement

Introduction

The following scope statement applies to the Striped Directories Scope Statement project within the SFS-DEV-001 contract/SOW dates 08/01/2011.

Problem Statement

Today, Lustre filesystem routinely have thousands to tens of thousands of clients. As client numbers continue to increase, a single Metadata Server (MDS) for a single Lustre filesystem becomes a performance and scalability constraint. This restraint has been partially lifted with the completion of the [DNE1: Remote Directories](#) sub-project which spread sub-directories on Lustre onto independent MDS nodes and MDTs. A limitation still exists, however, for individual directories. Currently a single directory can only be a single MDT served by a single MDS. This restriction limits both the quantity and performance of files in any given directory.

Project Goals

DNE 2: Striped Directories enables multiple MDTs on multiple MDS nodes serving a single directory. The project goal include

- User tool to allow administration of striped directory.
- Show an increase in metadata throughput with doubling of MDT count when serving multiple clients on a single striped directory.
- Code landed on Lustre Master branch.
- Demonstrate that DNE reacts well and remains manageable in various error conditions, including failover of any MDT, recovery from complete server cluster failure and damage limitation after permanent failure of any MDT, including the root MDT.
- Demonstrate metadata performance isolation between subtrees located on separate MDSs.

In-Scope

- Migration tool will move individual inodes from one MDT to another MDT.
- LFSCK OI Scrub, LinkEA, FID-in-Direct agent and proxy parent interaction with DNE 1 verified.
- Asynchronous updates between MDTs.
- commit-on-share
- striped directory
- man page and manual documentation.
- test plan

Out of Scope

- Consistency checking. LFSCK Phase 3 will address consistency checking across MDTs.
- Accessibility of striped DNE directories with 1.8, 2.1 or 2.4 clients.
- Interoperability of DNE-enabled MDTs and non-FID-based 1.8/2.1 OSTs.
- Recovering from a permanent failure of the storage target when a remote directory has a parent directory on any MDT other than MDT0.
- Migration tool will not have automated MDT selection.

Project Constraints

- Wang Di is the only engineer with the correct expertise available for this work.
- All milestones that require a code demonstration will take place on the OpenSFS Test Cluster.

Key Deliverables

- Signed Milestone documents for project phases:
 - Solution Architecture.
 - High-Level Design.
 - Implementation & Test.

- Acceptance Testing (OpenSFS executed).
- Test Plan
- Source code that meets feature requirements and runs on OpenSFS Test Cluster.
- Source code for new test cases.
- DNE 2: Striped Directories code landed in the Lustre Master branch.

Key Milestones

Design Documentation Q2 2013

Code delivery: Q4 2013

Code landing: Q1 2014