Demonstration Milestone for OI Scrub

Overview

This document describes demonstration of the Inode Iterator and OI Scrub (LFSCK Phase 1). The Inode Iterator and OI Scrub code is functionally complete. The purpose of this Milestone to verify that the code performs acceptably in a production-like environment. In addition to completing all the Test Scenarios (recorded in the OI Scrub Solution Architecture), OI Scrub Performance will be measured as described below.

Performance Baseline

Inode Iterator and OI Scrub is novel development work. No baseline measurement exists. The following test will be performed to create a baseline measurement.

1. performance of urgent OI rebuild in case of missing OI file.
2. performance of urgent OI repair after backup/restore (if the results are identical to (1) at a large scale this test will be omitted from future testing).
3. performance impact of urgent OI scrub under load.
4. performance impact of background OI scrub under load.

Test Methodology

The intention of this phase of testing is to demonstrate the Inode Iterator and OI Scrub behaves in accordance with typical real-world usage scenarios on a realistic environment (defined above). Completing these tests will provide a baseline for future performance enhancements. To perform the tests, the following tools are required:

Test filesystem configuration.

- 2 Clients, each with 16 threads.
- Create 10 000 000 files.
- Run OI Scrub over all the files.
- Where time permits, repeat with 20 000 000, 30 000 000, 40 000 000 and 50 000 000 files.

Load Generation.

Load will be generated on the MDS using `mds-survey`.

Hardware platform

Demonstration will take place on the Whamcloud cluster Toro. The hardware specification include:

- 2 Client Nodes, 3 Server Nodes (1 x MDS, 2 x OSS)
- MDS node are of type 'Fat Intel Node'
  - 2 x Intel Xeon(R) X5650 2.67GHz Six-core Processor (2-HT each core)
  - 16GB DDR3 1333MHz Memory
  - 2PT 40Gb/s 4X QSFP InfiniBand adapter card (Mellanox MT26428)
  - 1 QDR IB port on motherboard
  - SSD as external journal device (INTEL SSDSA2CW120G3), SATA II Enterprise Hard Drive as MDT (single disk, WDC WD2502ABYS-02B7A0),
- OSS Nodes are of type: 'Fat AMD Node'
  - 2 x AMD Opteron 6128 2.0GHz Eight-Core Processor

        - 16GB DDR3 1333MHz Memory
        - 2PT 40Gb/s 4X QSFP InfiniBand adapter card (Mellanox MT26428)
        - 1 QDR IB port on motherboard
        - 3 x 1TB SATA II Enterprise Hard Drive (single disk, WDC WD1003FBYX-01Y7B0)
- Client Nodes are of type: 'Type 1 Thin Clients'
- 2 x Quad-Core Intel E5507 2.26G/4MB/800
    - Mellanox ConnectX 6 QDR Infiniband 40Gbps Controller (MTS3600Q-1BNC)
    - 12GB DDR3 1333MHz E/R memory
    - 3.5" 250GB SATA II RAID Enterprise Hard Drive (SATA II 3.0Gb/s, 16MB Buffer, 7200 RPM)
- Infiniband between all the nodes: MTS3600Q managed QDR switch with 36 ports.

Test results

All agreed tests have been completed and the follow results are now available.

OI Scrub completes all acceptance tests.