File Migration on Lustre

EOFS Workshop

Gabriele Iannetti

GSI Helmholtz Centre for Heavy Ion Research

22-02-2024

Table of Contents

Lustre Installation

- 3 MDS pairs with manual failover
- 112 OSS / 784 OST
- 22PB data / 874M files
- 609 Clients / 764 User
- Data can be archived to TSM tape library
- High number of small files 80% with size <= 32MB
- Very few files are striped
- No scratch file system!

Use Cases and Requirements

Use Cases:

- Adding new file server
- Removing old file server due to decommissioning
- Rebalancing uneven OST fill states for a balanced state

Requirements:

- Low impact on live system (single OST-to-OST migration)
- Consideration of source/target OST fill states

Lustre OST File Migration with Cyclone

Uses Cyclone - A distributed task driven framework, first presented at the Lustre Administrators and Developers Workshop 2017 (LAD'17)

CORE - Client/server architecture for distributing and executing tasks

USE CASE specific components:

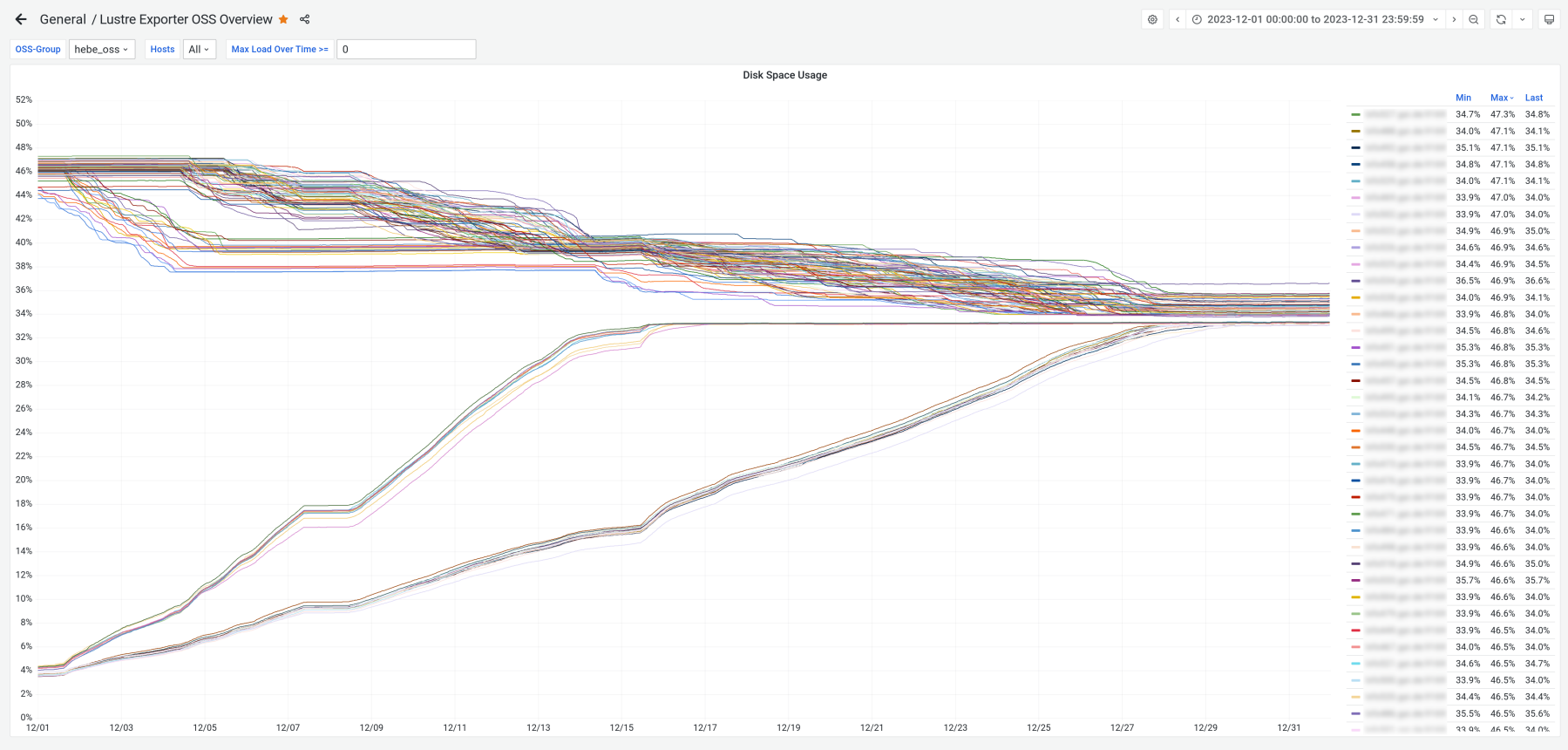
- LustreOstMigrationTaskGenerator (Impl. TaskGenerator Interface)
  - Checks OST states (fill level, available) for source/target OSTs7
  - Loads input files containing information about files to be migrated
  - Uses FIFO approach for each source OST
  - Decides to generate a LustreOstMigrateTask
- LustreOstMigrateTask (Impl. TaskInterface)
  - Migrates file (e.g. non-blocking, can ignore if already migrated)
  - Striped files bypass single OST-to-OST migration
  - Logs result

Workflow

1. Set 'max_create_count=2000' for new OSTs / 'max_create_count=0' for OSTs to be removed
2. Create unload files for each source OST by dumping OST information from the Robinhood Policy Engine[1]
3. Transform unload files to input files for Cyclone
4. Shuffle input files for filling new/existing OSTs
5. Pass input files to Cyclone[2]

[1] https://github.com/cea-hpc/robinhood
[2] https://github.com/GSI-HPC/cyclone-distributed-task-driven-framework

Example File Migration



- 8 Clients with each 22 worker processes (176 worker)
- 609 Source OSTs
- 174 Target OSTs
- 7.76PB data
- 144M files

Thank you!