

LUG13

Lustre and HSM

Ashley Pittman

- ▶ Lustre HSM is coming
 - Unfortunately not in 2.4
- ▶ Multiple vendors are working on this and we're hopeful that it'll be soon.
 - Once it's in master/released version then we anticipate a large takeup of the new features.

A quick over-view of the simple and not-so-simple options for using HSM.

Copy files to backend to protect against data loss or accidental deletion.

Copy files to POSIX backend in such a way as they can be viewed through some other filesystem.

Copy files to backend and stub in Lustre to increase capacity or use cheaper storage for long-term archive.

Copy files to cloud via Object/S3 type interface.

Allows new access models to data but requires re-tooling

Policy-based mirroring to remote sites

Copy files to remote Lustre filesystem.

Is this any different to mirroring to POSIX? Is it viable over long-distance, non highly available links.

HSM Stubs allow “pull” based model from external data sources

Lustre could be used to mirror external data, creating stubs in Lustre as remote metadata created any only importing contents onto Lustre on read.

Having two or more Lustre filesystems and having access to the same data globally, modifiable by all simultaneously.

Use of HSM allows inspection of data at archive time

- ▶ Dedupe
- ▶ Compression
- ▶ Search
- ▶ Transcoding?

- ▶ What are people doing today?
 - Perception is few people are using it currently but a lot of sites are interested in it.

- ▶ How would people like to use it?

- ▶ Can HSM change/improve the end-user workflow?

- ▶ Are more HSM features required in Lustre itself?
 - Or will the emphasis be more on add-on tools?