HSM, Data Movement, Tiering and more

LAD 2019
Ben Evans

bevans@cray.com
nrutman@cray.com
HSM – Current Status

```
IFS hsm_archive $foo
```

Diagram showing the flow from `Client` to `MDS` to `Copytool` with actions and status.
What’s Wrong with this Picture

1. lfs hsm_archive $foo
2. coordinator
3. Copytool
4. Copytool scaling DNE

Client
MDS

no programmatic interface
FIFO
no prioritization controls
no start indicator
Externalize the Coordinator LU-10986

Ifs hsm_archive $foo

Client → MDS → Coordinator → Copytool

- Prioritization schemes
- Preemption
- Can be global vs per MDS
Initiate a Request from a Copytool?

Ifs hsm_archive $foo

Client -> MDS -> coordinator

action

status

Begin, End

Copytool

Of course!

Work
Archive (New Interface)

• Can be handled entirely by the data mover
  • Only HSM xattr info needs to be set (ARCHIVED, etc.)
  • Need to check data versions to make sure nothing changed
• Failovers? Not a problem
• Locks? Not a problem
• Multiple backups? All userspace
• But wait, there’s more!
Restore (New Interface)

• Can be handled entirely by the data mover
  • Need to check that it can be restored (basically the ARCHIVED flag)
  • Restore data to a temp file, then do a layout swap
  • Remove RELEASED flag
  • Failovers? Not a problem
  • Locks? Not a problem (kind of)

• Implicit restores obviously need to be given to a copytool
  • MDS still takes layout lock
  • Reacquires lock after failover (lock replay)
New Picture

job scripts

policy engines

Copytool

status

Begin, End

Lfs hsm_archive $foo

Client

MDS

Coordinator

Work
“HSM” Migrate LU-9855

• Use current lfs migrate to create a temp file with new striping info
  • Add it to the HSM queue
  • Copytool copies data from the file to the temp file
  • Layout swap
  • Delete the old temp file
• No real coordinator changes needed (internal or external)
• Use your efficient parallel copytools instead of local single lfs thread
• Use your coordinator to prioritize and schedule
“HSM” Mirror Sync

• Use current lfs mirror to set up mirrored layout
  • Mirror sync adds it to the HSM queue
  • Copytool syncs mirrors
  • Copytool updates mirror status
• No real coordinator changes needed (internal or external)
• Use your efficient parallel copytools instead of local single lfs thread
• Use your coordinator to prioritize and schedule
THANK YOU

QUESTIONS?

bevans@cray.com

@cray_inc

cray.com

linkedin.com/company/cray-inc/