



Restart/Continue pNFS Metastripe



Metastripe: pNFS for Metadata

- Symmetry
- First Eisler proposal 2008?
- Has been expired for a while (2011)
 - Eisler slides from 2010 last visible effort
 - NFSv4 discussion indicates there is (was) interest in seeing further work on metastripe



Rationale

- Metastripe can provide 1-to-1 mapping to MDS architecture concepts that were
 - In CephFS already
 - Or that were added in CohortFS



CohortFS Metastripe Edits

- We like original metastripe
 - Can take advantage of all defined striping modes
 - Will implement
 - Are exploring some conceptual changes
 - <http://cohortfs.com/sites/default/files/draft-mbenjamin-pnfs-alt-metastripe.txt>



Cosmetic

- Terminology that we've started using internally
 - Inode striping (“filehandle-only”)
 - Dentry striping (everything else)



Substantive

- Reduce layout state overhead
 - Avoid requirement to take metastripe layouts on regular files
 - Propose *stripe hints* (attribute)
 - Coarse-grained layout (file-system)
 - Deviceid hint (attribute)



- Simplify metadata layout itself
 - Remove layout filehandles (try to)
 - Simplified device model
 - Define one device structure and layout device presentation
 - always use it
 - Keep “offsets” opaque



Inode striping

- Filesystem “singleton” layout
 - Holds device list and stripe pattern
 - Permits clients to request stripe hint
 - Recommended attribute



Dentry striping

- All other metastripe (on directories)
 - Essentially just like original metastripe
 - Name-based operations (CREATE)
 - Directory enumeration (in parallel)



Recommended Attributes

- meta_stripe_deviceid (deviceid4)
 - Inode stripe hint
- meta_stripe_count (uint32_t)
 - Help clients decide whether to take a directory metadata layout



PREADDIR

- Like READDIR, but takes a stateid
 - Consequence of removing layout filehandles



Q/A