RFC5664 – pNFS OSD Layout ERRATA

Benny Halevy <bhalevy@panasas.com>
IETF NFSv4 W-G Interim Meeting
Feb. 18 2011

Motivation

- Implementation of RAID-5 scheme revealed mistakes and obscurity of definitions.
  - Typo level bugs in math
  - Undefined cases
    - e.g. interaction with group_depth
  - Missing context for RAID-6
    - e.g. what algorithm to use exactly.
    - Rotation scheme
  - Required clarification
    - e.g. does odm_group_width contain mirrors?
- No wire format changes!

What's to fix

- Section 3.4: rename PNFS_OSD_RAID_PQ enum value to PNFS_OSD_RAID_6_RS

- Section 5.3.3: Clarify mirroring-related math and placement of mirror components in olo_components array.

- Sections 5.3.1, 5.3.2, 5.4.4: Fix and clarify striping math and examples for RAID_0, 4, 5, and 6.
  - Missing parenthesis
  - Unclear device rotation method
  - Undefined case: rotation cycle across parity groups.

What's to fix (2)

- Section 5.4.4: PNFS_OSD_RAID_6_RS
  - Explicitly specify the Reed-Solomon algorithm parameters and polynomial
  - Add Xrefs
  - Specify method to compute parity rotation cycle length
- Section 5.4.5: RAID Usage and Implementation Notes
  - Add comment about possible future extensions for erasure codes with P>=2

How to proceed

- ERRATA to be sent to mailing list for review
- ... before submission to RFC Editor
- RFC5664bis?
  - Changes seem substantial enough to justify bis
- New Layout Type?
  - No need to