EOS site report of the Joint Research Centre

EOS workshop April 2023

Armin Burger, Franck Eyraud

EOS at the Joint Research Centre

- Main storage backend of the JRC Big Data Analytics Platform (BDAP)

- Used as file storage by ~80 JRC projects

- Access via

  - BDAP services (FUSE client)

  - CIFS gateway (scientific PCs, read-only)

  - Cloud application

EOS set-up at JRC – Overview

- 1 production instance, 1 test instance

- 28.4 PiB gross capacity

- Access method currently via FUSE client only

- Replica 2 used throughout

EOS instance at JRC – Software

- EOS versions used
  - MGM: 4.8.91
  - FST: 4.8.91
  - Processing & service nodes: 5.x (some 4.x)

- QuarkDB : 0.4.2

- EOS FUSE client:
  - EOSX on all processing and service nodes except Nextcloud service

- Server OS: CentOS 7

EOS instance at JRC – Current status

- 60 FSTs with 1 or 2 JBODs connected
  - Disks per FST: 24, 48, 60, 84
  - Disk sizes: 6, 10, 14, 16 TB

- 28.4 PiB available gross space

- 24.1 PiB used

- 1.65 G files (2.4 G total created files)

- 241 M directories (393 M total created directories)

EOS storage space evolution planned in 2023

- Decommissioning phased-out FSTs

  - 7 year old hardware, 12 FSTs with 24*6TB

  => removing 1.7 PB gross

- Adding new 60-disk JBODs (16 or 18 TB disks)

  - 3 FSTs with 60*16TB JBODs in April/May

  - 6 FSTs with 60*18TB JBODs in summer

  => adding 9.4 PB gross

EOS activities in 2022/23

- Upgrade EOS MGM from 4.5 to 4.8
    - Worked fine
    - Main problems experienced were due to unexpected hardware behaviours
    - Solved various issues
        - No crashes any more during draining

- FUSE client migration to v5

EOS issues in 2022/23

- Continuing frequent crashes of FUSEX client 4.8
  - seems to be noticeably reduced by upgrading client to 5.x
  - still issue with FUSE cache growing too much and filling local partition, causing client crash

- FSTs with high number of files (84*16TB disks)
  - causing memory exhaustion due to LevelDB database in memory (on FST with only 128 GB RAM)
  - causes high latency on FST when accessing files from replica, writing could be interrupted
  - expected to be resolved with EOS v5 and deprecation of LevelDB, using extended file attributes

EOS issues in 2022/23 [2]

- Discovered default 512 GB file size limit for FUSEX client

- Crash of EOS due to "new-year bug"

- Upgrade to bug-fix version 4.8.100 caused continuous crashes of MGM => had to be downgraded back to v 4.8.91

Wishlist for the future

- Add to EOS permission system more flexibility and possibility of more fine-grained permission settings
  - for CIFS gateway also secondary groups taken into account
  - possible hierarchical permission evaluation (currently only permissions of actual directory count)

- Data encryption => seems to be natively supported in v5 via XRootD

- Native support of S3 protocol via XRootD HTTP interface

- Add some recommended configurations to documentation
  - for better logging of crashes and extraction of stacktraces

EOS planned activities in 2023

- Upgrade EOS to latest 5.x version: *high priority !*

- Evaluate possible virus-scan set-up based on parsing file access logs

- Evaluate encryption functionality for sensitive data

- Test native XRootD HTTP access mechanism as alternative to

  - FUSEX mount

  - MinIO gateway

- Evaluate token-based authentication as alternative to kerberos

Thank you