Purdue EOS status report

8th EOS workshop at CERN

March 15, 2024

Stefan Piperov

spiperov@purdue.edu

PURDUE UNIVERSITY

Research Computing
INFORMATION TECHNOLOGY

Motivation (refresher)

- In 2021 US-CMS asked Tier-2 sites to migrate to storage systems capable of Erasure-Coding, in order to improve density and therefore cost of storage.

- We at Purdue chose EOS over CEPH and HDFS3 for the following advantages:
  - Main storage system at CERN - proven performance, reliability;
  - Developed and maintained by CERN - implies availability and support for the lifetime of LHC and the CMS experiment;
  - Native XRootD storage;
  - Better Erasure Coding - incl. “Dynamic EC” and higher strip size.

- 2 years later - how does all that look?

Our setup

- 3x QDB servers, 1 shared with the MGM
  - 2x SSD (OS), 1.6TB Enterprise NVMe (DB)
  - each runs 3x QDBs (yes, a total of 9 in the QDB cluster!)
  - Single MGM - no HA/redundancy

- Still a mix of 36-disk storage servers, and JBODs connected to more powerful storage servers.
  - Retiring the smaller 36-disk boxes, adding only 102-disk JBODs
- Everything uses either bonded 2x10Gb Ethernet, or 2x25Gb NICs

- Namespace lauout

15 disks

102 groups

JBOD1 102 disks	JBOD2 102 disks	3x 36-disk servers	...	3x 36-disk servers	...	JBOD3 102 disks	...	JBOD5 60 disks	(future JBODs)
FST1 34x16TB	FST1 34x16TB	FST 34x3TB	...	FST 34x6TB	...	FST1 34x14TB	...	FST 60x14TB	
FST2 34x16TB	FST2 34x16TB	FST 34x3TB	...	FST 34x6TB	...	FST2 34x14TB	...		
FST3 34x16TB	FST3 34x16TB	FST 34x3TB	...	FST 34x6TB	...	FST3 34x14TB	...	(empty)	

PURDUE UNIVERSITY

Observations, Problems

- Replacing failing disks (we have some very old ones!) works well.
  - We are glad we adopted *unique* mount-point naming convention, based on disk's WWN ID - e.g.:
    - /data/d042_g2_258508dac
- LRU mechanism depends on 'default' namespace being present
  - We didn't have one :)
  - Regular cleaning of empty directories was not working
  - replication→EC conversion wouldn't have worked either
  - "Too many things in EOS have 'default' value"
- UNIX auth is not working correctly with FUSEX mounting.
  - UID limited to 128k ?
  - Still working with Devs on this
  - "It just works with Kerberos"...
- Failing HDDs tend to take the whole FST down :(
  - We had multiple cases where (multiple) failing disks in one server causes it to reboot periodically
  - We have narrowed it down to memory exhaustion (OOM killer).
  - Is the xrootd process trying to buffer too aggressively the data from the failing I/O operations on the disk(s)? To the point where it consumes all the RAM in the node.

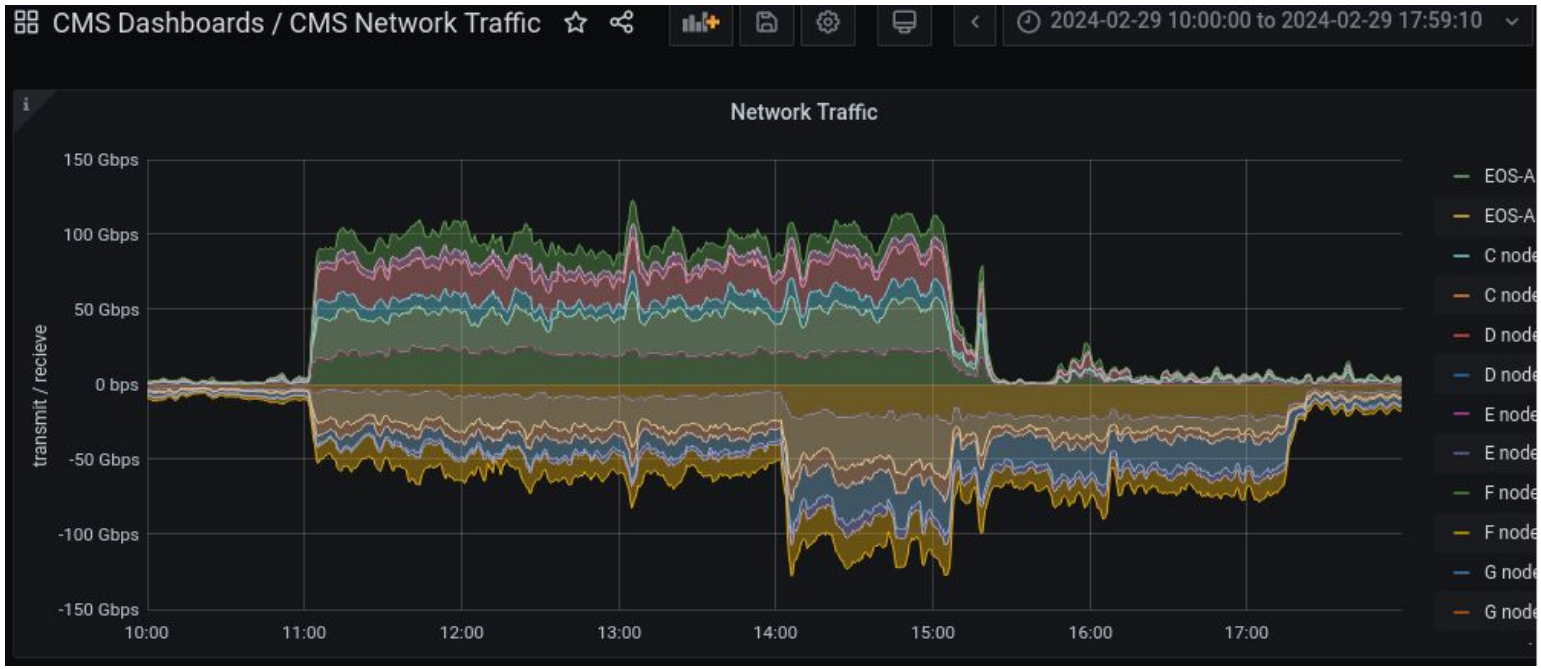
Current status and Plans

- EOS is *the* Production Storage System at Purdue
  - Current capacity: 14 PB
  - Used: 11 PB
- CMS Production and Analysis jobs are running without problems.
- FUSE mounted (read-only) in Front-end machines and Analysis Facility pods.

- Plans
  - Update to latest EOS5 version (perennial)
  - Provide R/W FUSE-mount access for users
  - Switch to Erasure Coding for CMS data
  - Retire older storage servers (perennial)
  - Deploy new JBODs (perennial)

Conclusions

- 2 years in production EOS storage system is serving us well!
- Performance is adequate - e.g. during latest DC'24 we were receiving data from FermiLab at the rate of 75Gb/s and storing (replication x2) in EOS.
  - Limit most likely imposed by FTS (only 300 parallel transfers)
- Great collaboration with the developers at CERN - Thank you!
- We are still to reap the benefits of Erasure Coding



PURDUE UNIVERSITY