HDFS to EOS migration - Purdue site report

7th EOS workshop at CERN

April 25, 2023

Stefan Piperov

spiperov@purdue.edu

PURDUE UNIVERSITY | Research Computing INFORMATION TECHNOLOGY

Motivation (refresher)

- In 2021 USCMS asked us to migrate to Erasure-Coding capable systems, in order to improve density and therefore cost of storage.

- We 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.

- A year later - how does all that look?

PURDUE
UNIVERSITY

Where did we start

- **HDFS storage system, ~10PB in size, continuously grown/updated**
  - Every year we'd add a bunch of storage nodes, and retire to oldest ones (Typically: 5-years in HW warranty, +3-years outside)

- **A couple of management servers (name-nodes)**

- **68x storage servers with**
  - Varying disk sizes - from 3TB to 14TB
  - Varying disk counts - from 12 to 36 (plus 60 or 102 in the JBODs)
  - Vastly differing CPU/RAM configurations:
    - 32GB RAM minimum (a *lot* more for the JBODs)
    - between 6 and 16 CPU cores.
  - 10GbE NICs.

- **'Ragged Array' does not even start to approximate the layout…**

PURDUE UNIVERSITY

Optimize the layout

- **How does one go from a ragged array to (approx.) rectangular one?**
- **We considered multiple namespaces**
  - one with 36-disk servers only,
  - another with 102-disk JBODs
  - But something felt off…

- **And then - BAM!**

14 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

Our current deployment

- **3x QDB servers, 1 shared with the MGM**
  - 2x AMD 7662 CPUs (128C/256T), 512GB RAM
  - 2x 480GB SSD (OS), 1.6TB Enterprise NVMe (DB)
  - each runs 3x QDBs (yes, a total of 9 in the QDB cluster!)
    - QDBs had different NVMe partition sizes on purpose. That helped us catch a problem in time, as intended.
  - Single MGM - no HA/redundancy
- we believe these 3 nodes are somewhat oversized HW-wise - will replace with less powerful machines this year.

- Mix of 36-disk storage servers, and JBODs connected to more powerful storage servers.
- Everything uses 10Gb Ethernet; only the largest storage servers have 2x10Gb bonded NICs.

PURDUE UNIVERSITY

Going into production

- Very slow start
  - We decided to go directly for EOS5 (to avoid early upgrades)
  - Rocky Linux 8 (we wanted even 9, but...)
  - Had no local expertise with EOS
  - A lot of difficulties with the early 5.0.X versions.
  - Documentation available for EOS4 only.

- Things were further delayed by having to move to a new Data Center

- Required a lot of discussions with developers at CERN
  - Slow in the beginning, but overall great in the end.
  - Several bug-fixes as a result
  - ‘Support by developers’ vs ‘support by community’

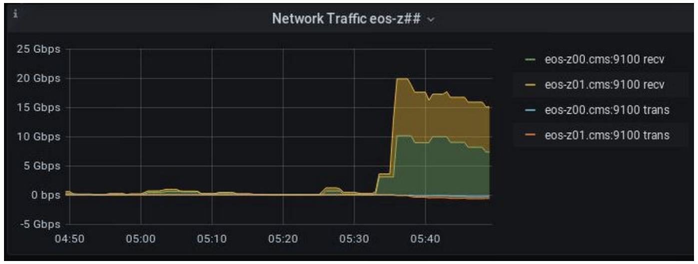
- Several test-deployments (and months) later, we reached production level stability and performance with the first EOS 5.1.x versions

Observations

- No complete example of deploying an EOS site in Documentation.
  - A quick minimal example
  - Many sections describing various aspects of installing an EOS system
  - But lack of overall picture (from a site's point of view, not developer's).
- Many vital sub-systems are disabled by default (e.g. balancer)
- Too many things are called 'default'
  - We identified some bugs just by systematically avoiding the name 'default' throughout our deployment process.
- Some config parameters are not clear about units

- Things that didn't work quite as expected in the end
  - Symbolic links (maps before that)
    - xrdfs cannot follow them, breaking the Rucio Consistency Checks
  - Authentication mechanisms, identities still give us trouble.
    - "Everything works just fine with Kerberos"

- Things that worked well
  - Everything else!
  - Performance is great, even with our old storage servers
  - Balancer does fantastic job!



PURDUE UNIVERSITY

Migrating data from HDFS to EOS

- **Iterative process**
  - EOS instance started with only 2 JBODs and 9 of the smallest storage servers (3TB disks) decommissioned already from HDFS
  - Moved one data-tier (`/store/data`) to EOS
  - Decommissioned more nodes from HDFS (slow!)
  - Move those to EOS
  - Copy next data-tier
  - wash, rinse, repeat…
- **Dedicated ‘transfer’ nodes**
  - 15 machines which were fuse-mounting both HDFS and EOS, and had both HDFS and EOS installed.
  - Files/datasets to be copied were distributed among those in a quasi-uniform manner
  - We gave up on FUSE mount (HDFS side was very slow), and wrote our own scripts to do native HDFS reading and XRootD writing.
  - 4GB/s speed (peak); ~2GB/s average; (btw: single FST does ~5Gb/s)
  - 5PB moved in 30 days.

PURDUE
UNIVERSITY

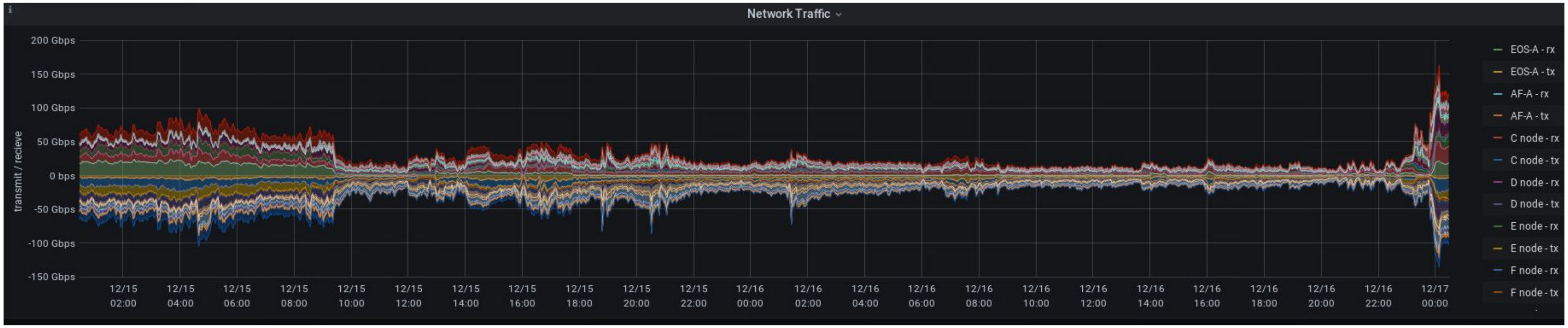
Current status and Plans

- EOS is the Production Storage System at Purdue since start of 2023!
  - Current capacity: 12.5 PB
  - Used: 8.5 PB
- CMS Production and Analysis jobs are running without problems.
- Users are starting to love the improved performance over HDFS
- FUSE mounted (read-only) in Front-end machines and Analysis Facility pods.

- Plans
  - Update to latest EOS5 version (perennial)
  - Add Kerberos authentication (Purdue's 'BoilerAD')
    - (too bad we cannot directly use CERN's instead!)
  - Switch to Erasure Coding for CMS data
    - Keep user files replicated; increase replication to 3x
  - Get rid of symlinks
  - Retire old(est) storage servers (8+ years)
  - Deploy new JBODs

Conclusions

- A year later we do have an EOS storage system in production!
- Clear need for better documentation for v5
  - esp with the growing popularity of EOS!
- Parts of the initial plan (2022) have changed
  - Bare-metal deployment instead of k8s
  - Storage groups layout (big improvement)
- Good performance and stability so far
- Growing users' appreciation ("appetites comes with eating")
- Great collaboration with the developers at CERN - Thank you!
- We are still to reap the benefits of Erasure Coding



PURDUE UNIVERSITY