www.cern.ch

Converging Storage Layers with Virtual CephFS Drives for EOS/CERNBox

Roberto Valverde Cameselle, CERN IT-ST
Dan van der Ster, CERN IT-ST
Andreas Peters, CERN IT-ST

26 January 2022

Outline

- Introduction & Motivation
- Previous Work
- Our Objectives
- Proof-of-Concept Testing
- Production Testing
- Discussions & Conclusions

Introduction & Motivation

- The CERNBox service is built on top of EOS Open Storage, CERN's highly scalable storage system initially developed for LHC physics analysis
    - **EOS** provides today **500 PB** of raw storage space
    - Data is persisted using **file based replication** (RW) or **Erasure Coding** (WORM) using XFS filesystems on disks
    - **Interactive use-cases** (mounted directly) require support for file **updates**
        - currently only supported with file replication
    - A file replication model has generic architectural and operational limitations

File Storage vs Object Storage

- Intrinsic limitations of file based storage with replication
  - IO performance is equal to that of a **single disk**
  - **Max file size** is the free space of the least full disk
    - in nearly full clusters, file appends can fail
  - File rebalancing and failure recovery time increases with file size used
    - problematic for very large (slow) and extremely small files (if many)

- Storing files in Object Storage
    - Each file is split into many **chunks**
    - **IO performance** scales with number of chunks / disks
    - File size is limited to the free space of the entire cluster
    - Data rebalancing and failure recovery is parallelized by chunks

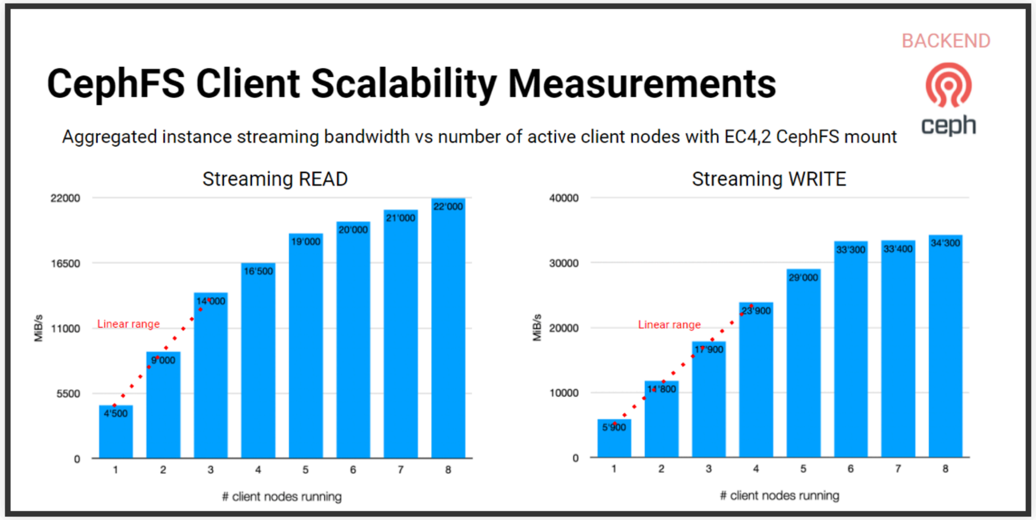
Virtualized Storage Services

- EOS provides a **separation of persistency** and a **(nearly) stateless** metadata service:
  - Metadata is stored in an HA backend (QuarkDB) and cached in the EOS manager daemon
  - The transition to this model has improved the service KPIs drastically

- By separating persistence from the **data** service we can have a fully virtualized EOS
    - Data **Availability**, **Durability**, and **Lifecycle mgmt** can be delegated to the storage backend
    - EOS IO daemons can be relocated between hosts as long as the storage backend provides concurrent access from several hosts

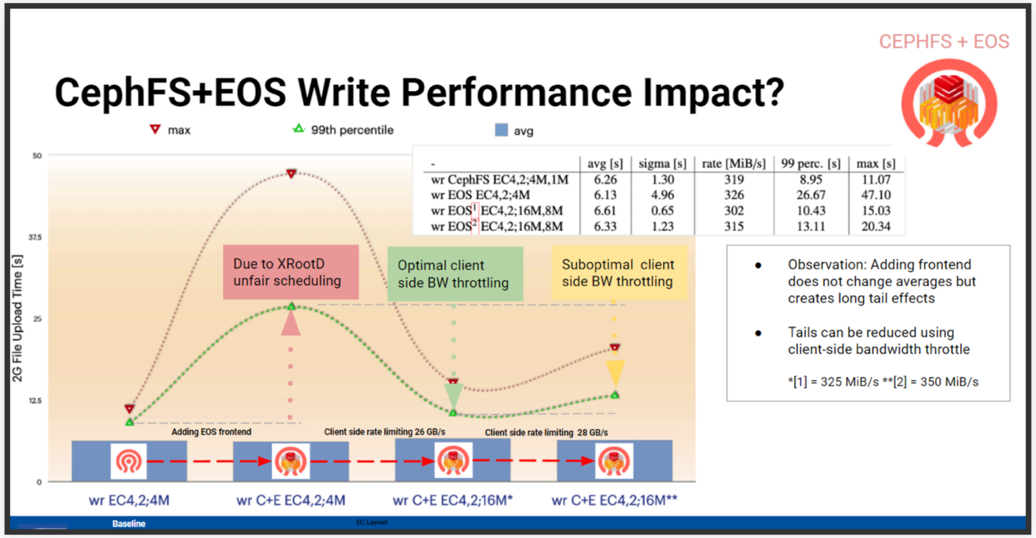
Previous Work

- At CHEP 2021 we evaluated a new approach to EOS storage:
  - CERN has many years of experience running CephFS for HPC and IT use-cases and has an active role in CEPH project
  - Replacing XFS with CephFS in the EOS storage back-end allows to benefit from Object Storage characteristics and keep EOS high-level functionality
- Evaluating CephFS Performance vs. Cost on High-Density Commodity Disk Servers



Benchmarking the CephFS kernel client.

On an 8-node 100Gig-E cluster it is capable of high throughput performance.



Layered EOS+CephFS introduced some long tail latencies in this high throughput test.

With tuned config it performed as well as the native CephFS backend.

Objectives

- Explore the benefits of a **combined EOS/CephFS solution as a CERNBox backend**
- Does it have an impact in **reliability**, **durability**, **availability**, **performance**?
- Would consolidating on one storage backend **save** on operations personnel or hardware?
- Can we enable **new use-cases** using this architecture?

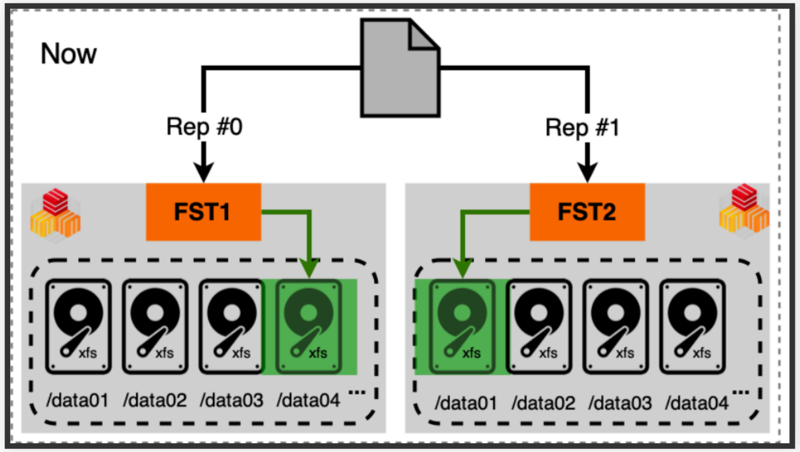
PoC Evaluation Criteria

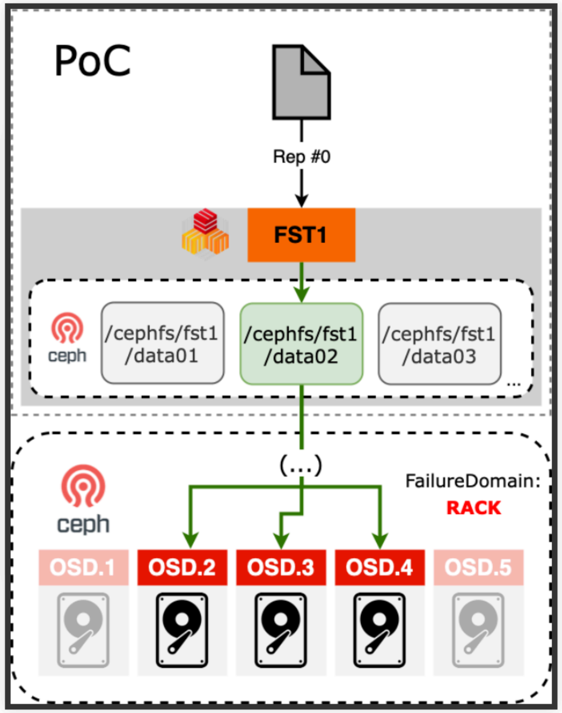
- **Reliability / Durability**
  - EOS consistency check (`fsck`) should confirm that data is **safely stored** on CephFS
- **Performance**
  - CephFS backend should not negatively impact performance (IOPS, throughput, latency)
- **Availability**
  - Frontend host failure should have **minimal impact** given the lack of a secondary EOS replica.
  - Understand how to dimension the frontends

PoC Testing Results

- We ran a microtest suite against the PoC over a 3 month period.
- Three configs: EOS dual replica, EOS single replica, CephFS

Replica Layout





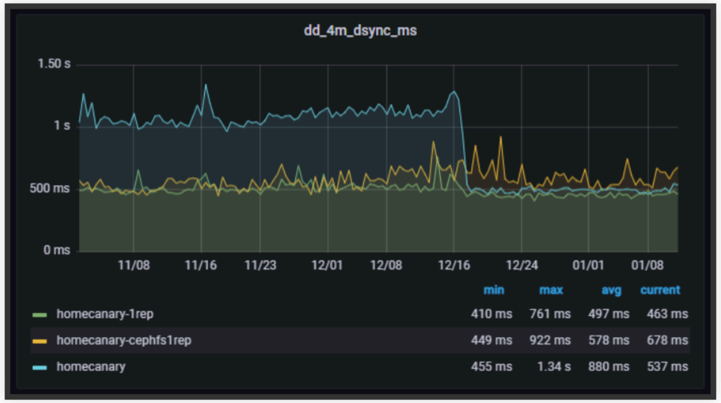
Reliability / Durability

- `fsck` confirmed that adding a CephFS backend did not introduce any data durability issues
- We found an unrelated replication issue EOS-5045

Performance

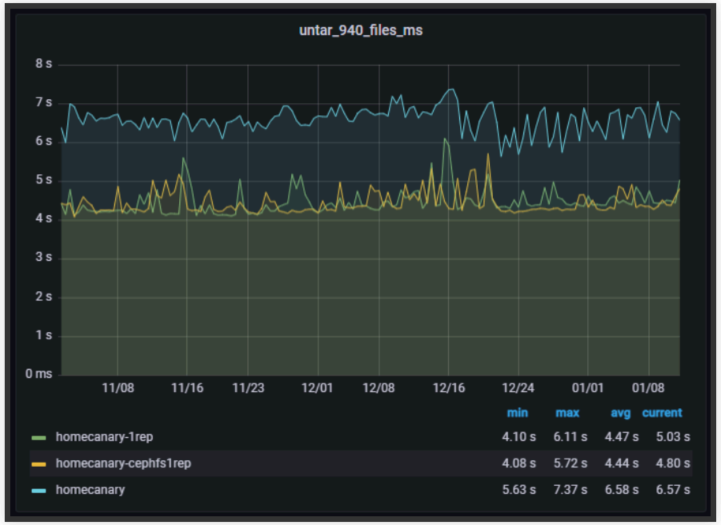
- Previous work confirmed that EOS+CephFS can achieve multi-GBps throughputs, but didn't measure interactive workloads

Example microtest: Time to write 4MB O_DSYNC:



Single replica performance is similar.
2x replica had a perf issue which was fixed on Dec 17.

Example microtest: Time to untar a small archive (~1000 files)



Single replica performance is similar.

Availability

- **Data is unavailable** when a frontend virtual FST is down (e.g rebooting or broken)
  - The virtual disk is just a path in the shared `/cephfs`
  - `eos fs mv` can be used to reassign that virtual FST to another frontend
- This impacts how many EOS virtual FSTs per frontend box

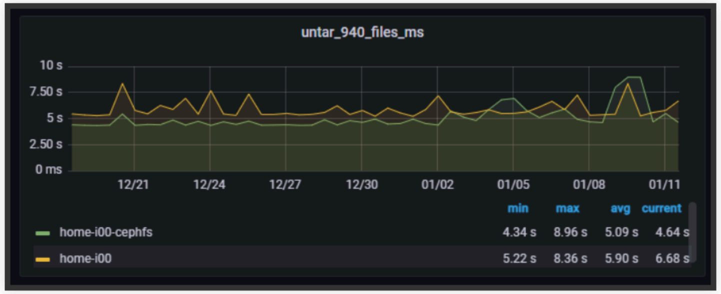
- When a frontend fails, we need to **redistribute** its virtual disks to the other remaining frontends.
- Operationally it is best if we can use as many other frontends in parallel
    - Ex 1: with 1 virtual FST – that single FST is taken over by one other box, whose load now doubles.
    - Ex 2: with 10 virtual FSTs – a single frontend failure can be taken over by 10 other boxes, whose load increases by only 10%.
- We choose to use 12 virtual FSTs per frontend box.
- Another approach would be to have idle standby frontends, but this wastes resources.

Prod Testing Environment

- **EOSHOME-i00** is a production CERNBox instance hosting several thousand users.
- We added a new "CephFS" space:
  - Two virtual FST hosts (CentOS Stream 8, 64G)
- Backed by our large shared production CephFS.
  - Also used by OpenShift, HPC, and many other CERN services.

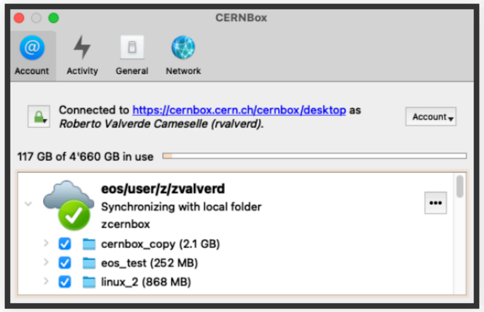
Prod Testing Results

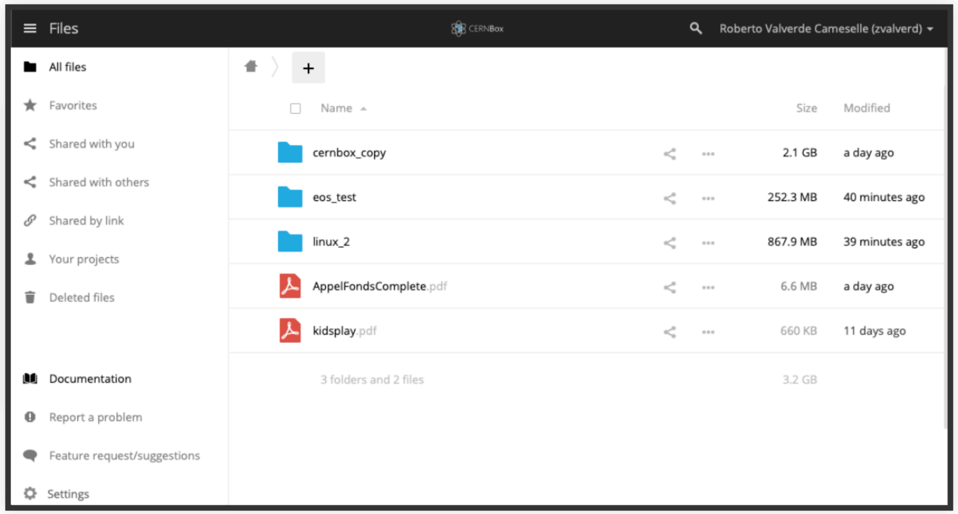
We enabled the same microtest suite in Dec 2021.



The results roughly match what we observed on the PoC.

- I also moved my home directory onto the CephFS-backed space.





Discussion & Conclusions

- Replacing XFS disks with CephFS completes the **storage virtualisation** of EOS
  - We expect significant **increase in KPIs**, similar to the EOS metadata -> QuarkDB transition
- CephFS backend is based on object storage
  - **Fewer limitations** related to performance, file size, and failure recovery
- This brings a much more **flexible architecture**
  - Delegate **reliability**, **durability**, **lifecycle mgmt** to Ceph (and e.g. Kubernetes)

Discussion & Conclusions (cont'd)

- What about **cost**?
  - At the **multi-PB** scale, CephFS read-write erasure coding should bring substantial savings
  - May also save on operations personnel by consolidating on our **existing** Ceph infrastructure and lifecycle processes
- Still lots to do:
  - Need experience with real CERNBox user workloads
  - Explore options to automate the EOS storage daemons, e.g. with Kubernetes persisent volumes

THE END

Any Questions?

Extra slides

PoC Testing Environment

- EOS Namespace Server:
  - 3x physical boxes w/ Xeon Silver 4216, 384GB RAM
- Baseline EOS Diskserver:
  - 3x physical boxes w/ Xeon E5-2650, 64GB RAM, 10Gig-E, 24x 3TB HDDs
- New “CephFS” EOS Virtual Diskservers:
  - 2x virtual boxes w/ 10 cores, 60GB RAM, 10Gig-E
- Backend CephFS cluster:
  - 3x OSD disk boxes w/ Xeon Silver 4216, 192 GB RAM, 48x 12TB HDD, 4x 1TB NVMe
  - 1x virtual CephFS metadata server

PoC Testing Environment

- This gives us two EOS “spaces”:
  - `default`: the traditional EOS storage, for baseline testing
  - `cephfs`: data stored in a CephFS backend
- We configured three paths in EOS for testing:
  - `/homecanary` -> `default` space with 2 replicas
  - `/homecanary-1rep` -> `default` space with single replica
  - `/homecanary-cephfs1rep` -> `cephfs` space with single replica