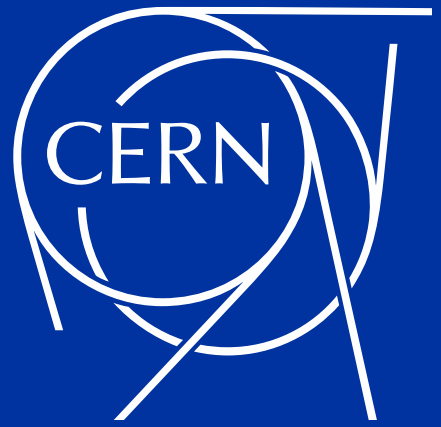


Converging Storage Layers with Virtual CephFS Drives for EOS/CERNBox

Roberto VALVERDE, Daniel VAN DER STER, Andreas PETERS

08/03/2022

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)

File Storage vs Object Storage (II)

- 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 **parallelised** by chunks

Virtualised 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

Virtualised Storage Services (II)

- By separating persistence from the **data** service we can have a fully virtualised 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

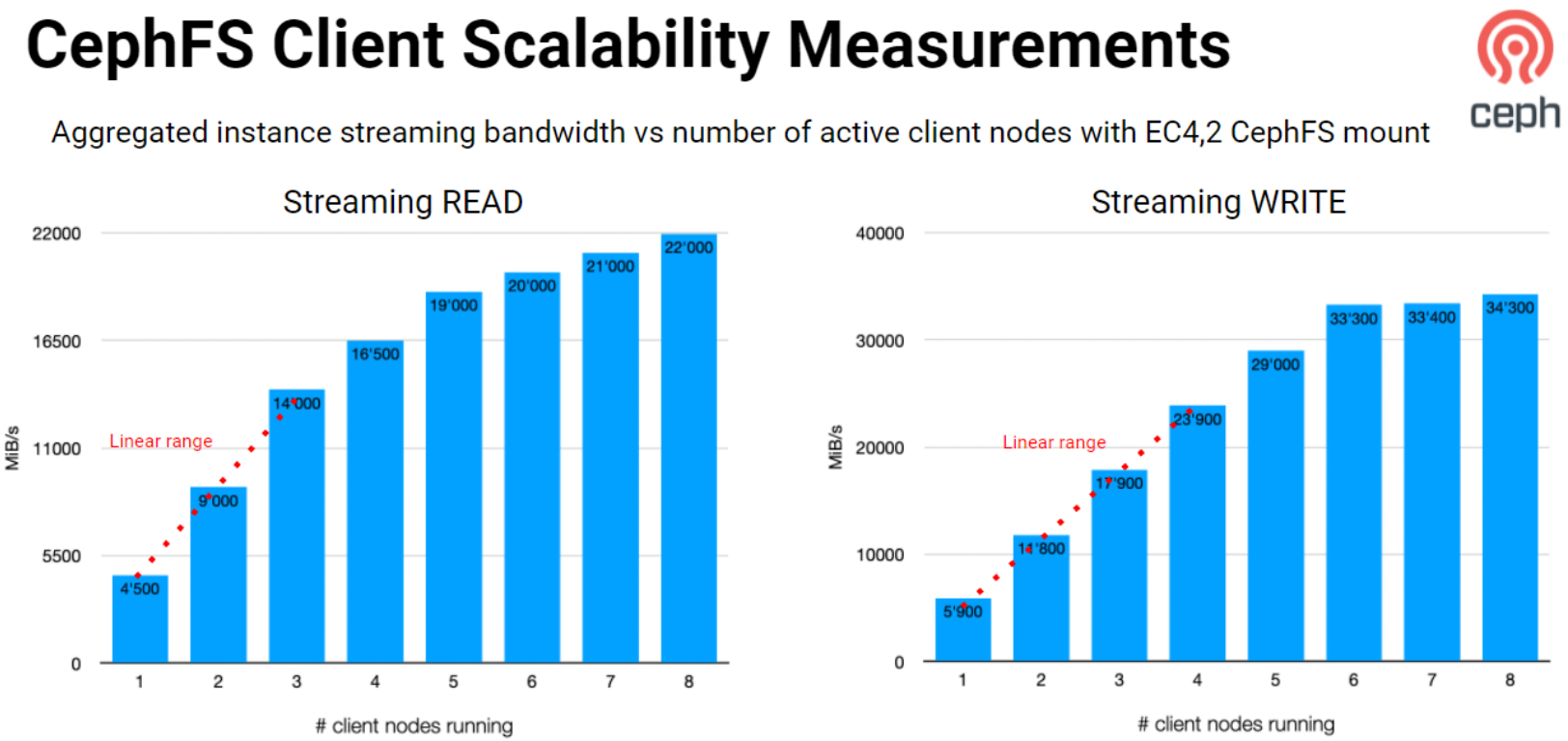
Previous Work

- 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 [Link]

Previous Work

- Benchmarking the CephFS kernel client.

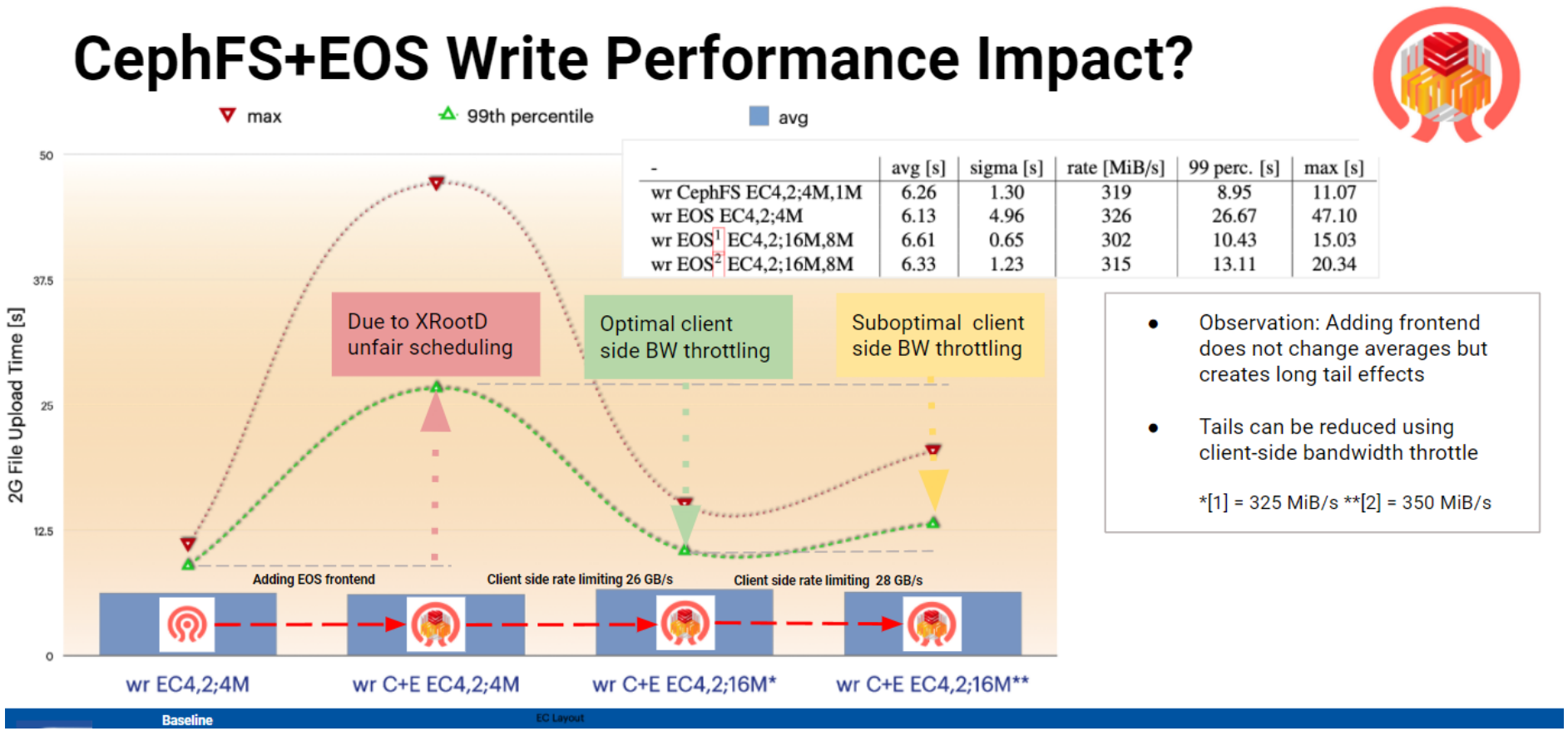
BACKEND



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

Previous Work

CEPHFS + EOS



-	avg [s]	sigma [s]	rate [MiB/s]	99 perc. [s]	max [s]
wr CephFS EC4,2;4M,1M	6.26	1.30	319	8.95	11.07
wr EOS EC4,2;4M	6.13	4.96	326	26.67	47.10
wr EOS[1] EC4,2;16M,8M	6.61	0.65	302	10.43	15.03
wr EOS[2] EC4,2;16M,8M	6.33	1.23	315	13.11	20.34

- Observation: Adding frontend does not change averages but creates long tail effects
- Tails can be reduced using client-side bandwidth throttle

*[1] = 325 MiB/s **[2] = 350 MiB/s

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

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?

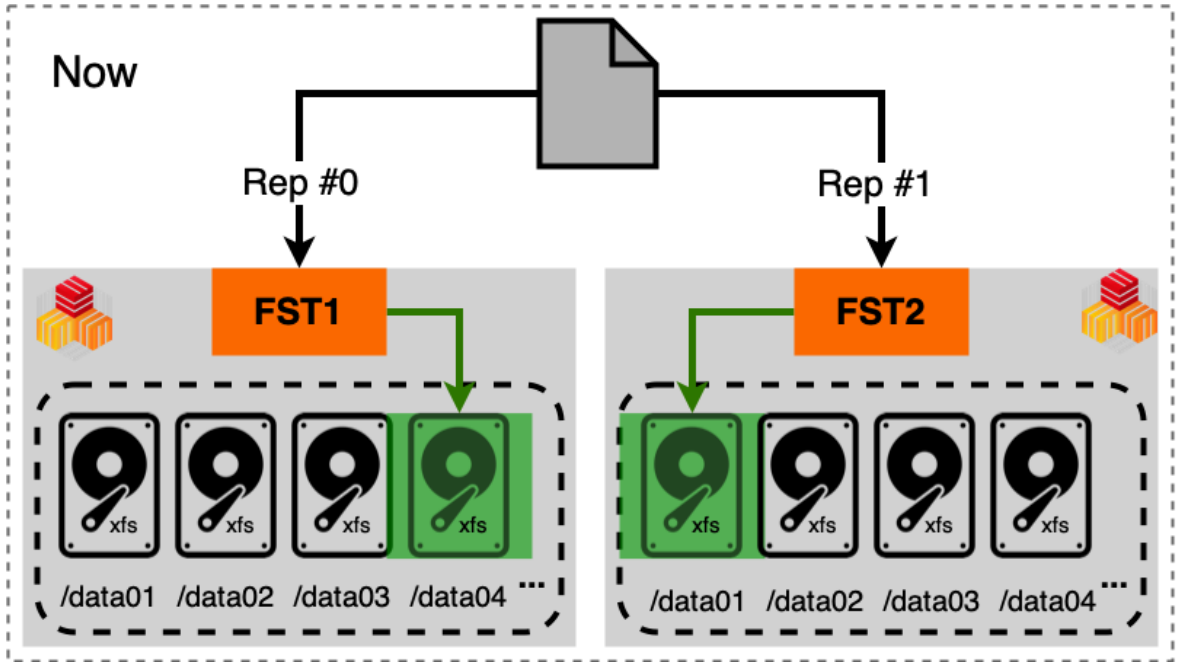
PoC Evaluation Criteria

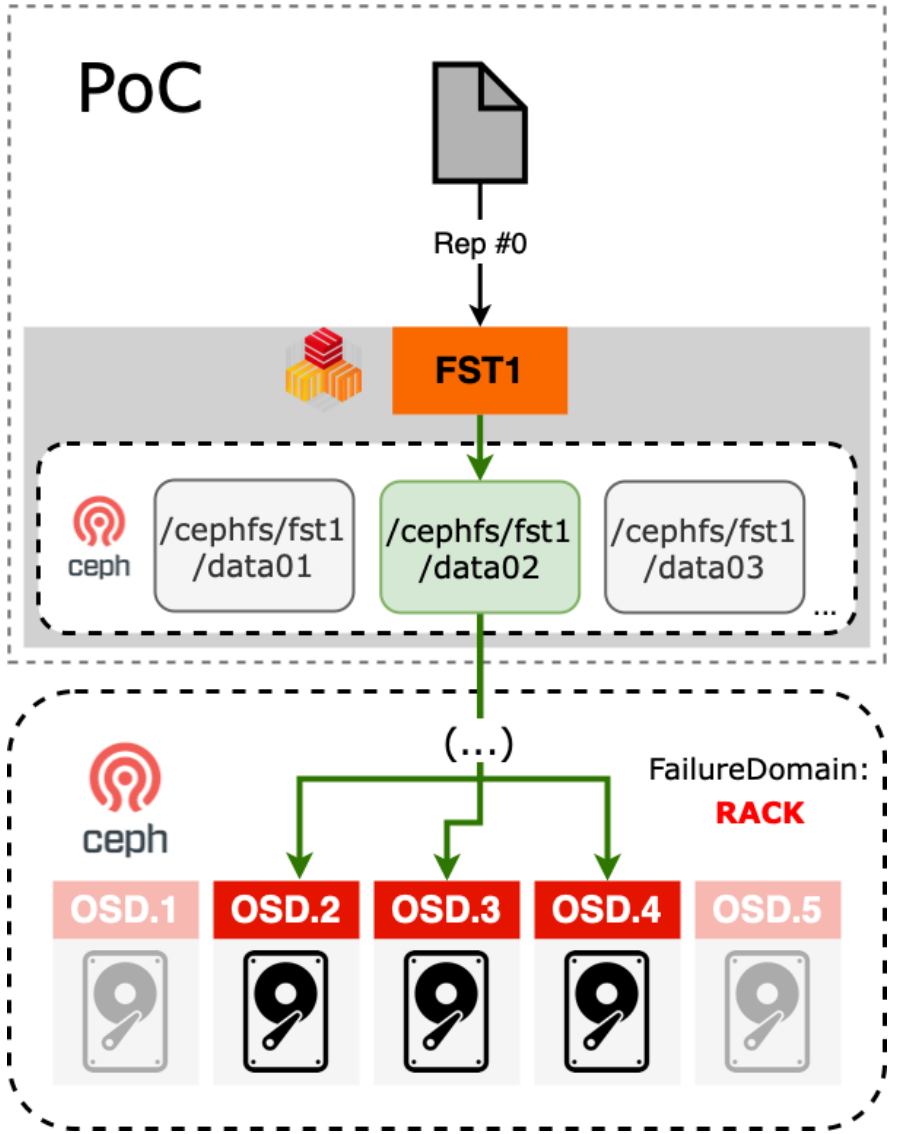
- **Reliability / Durability**
  - EOS consistency check (`fsck`) should confirm that data is **safely stored** on CephFS
- **Performance**
  - CephFS backend should not negatively impact performance (IOPS, throughput, latency)
- **Availability**
  - Frontend host failure should have **minimal impact**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

PoC Testing

- EOSHOMECANARY testing instance:
  - **default** space: disk-based storage servers
  - **cephfs** space: virtual CephFS storage servers
- We ran a microtest suite against the PoC over a 3 month period.
- Three configs: EOS dual replica, EOS single replica, CephFS

PoC Testing - Replica Layout





PoC Results: Reliability / Durability

- **fsck** confirmed that adding a CephFS backend did not introduce any data durability issues

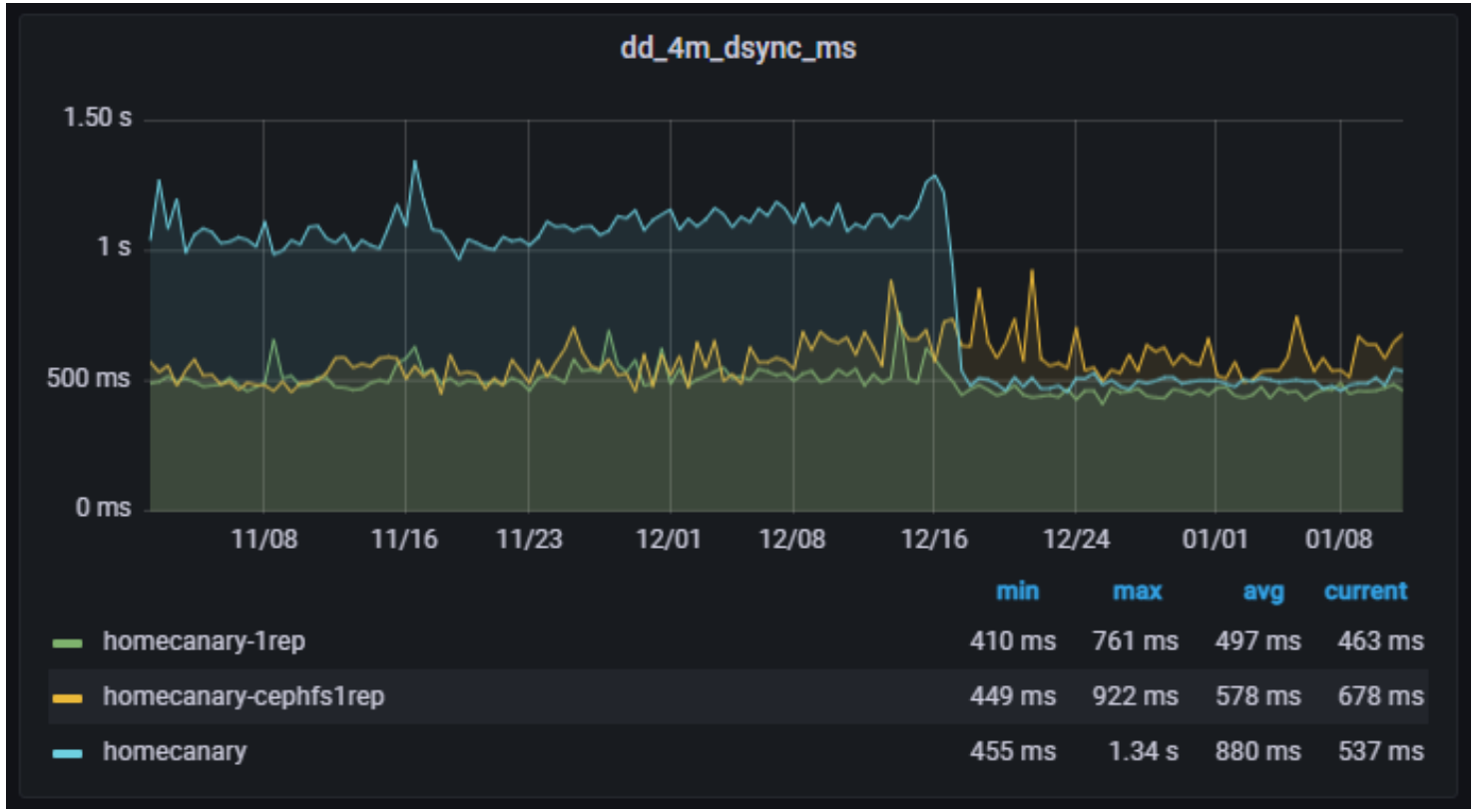
- We found an unrelated replication issue [EOS-5045]

PoC Results: Performance

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

PoC Results: Performance

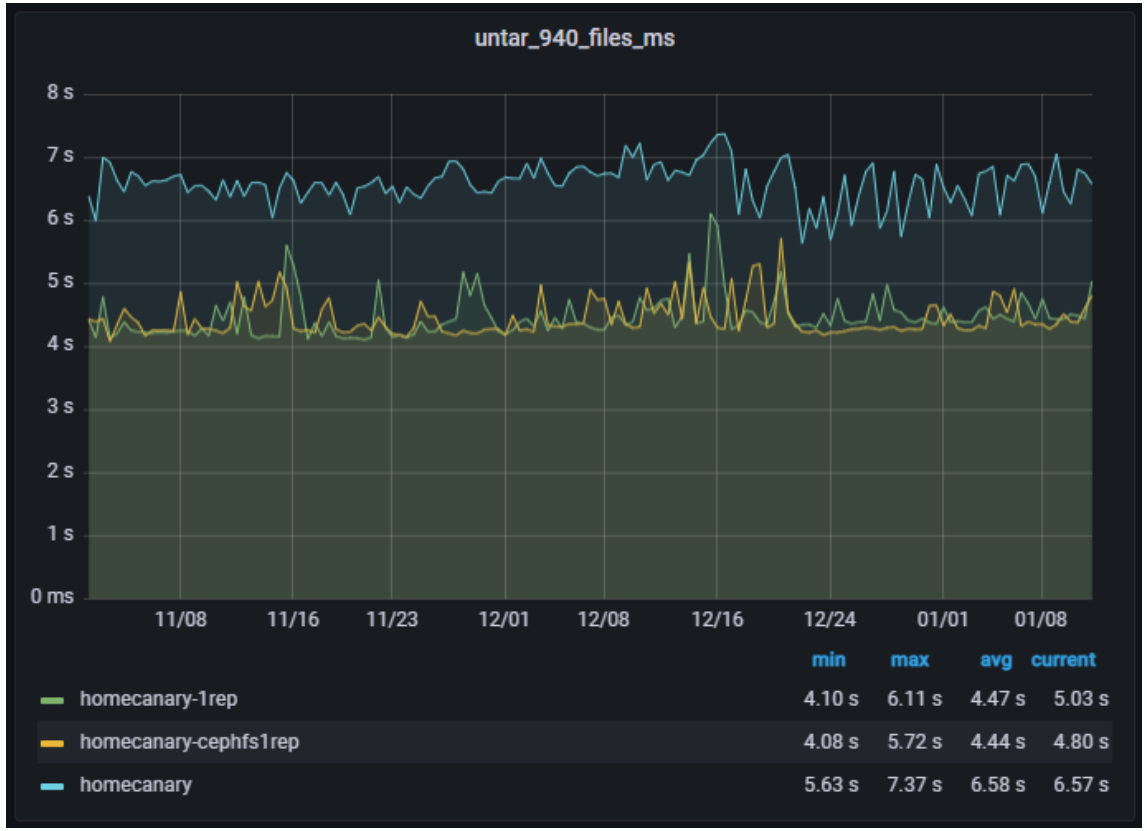
- Example microtest: Time to write 4MB O_DSYNC:



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

PoC Results: Performance

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



Single replica performance is similar.

PoC Results: 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

PoC Results: Availability

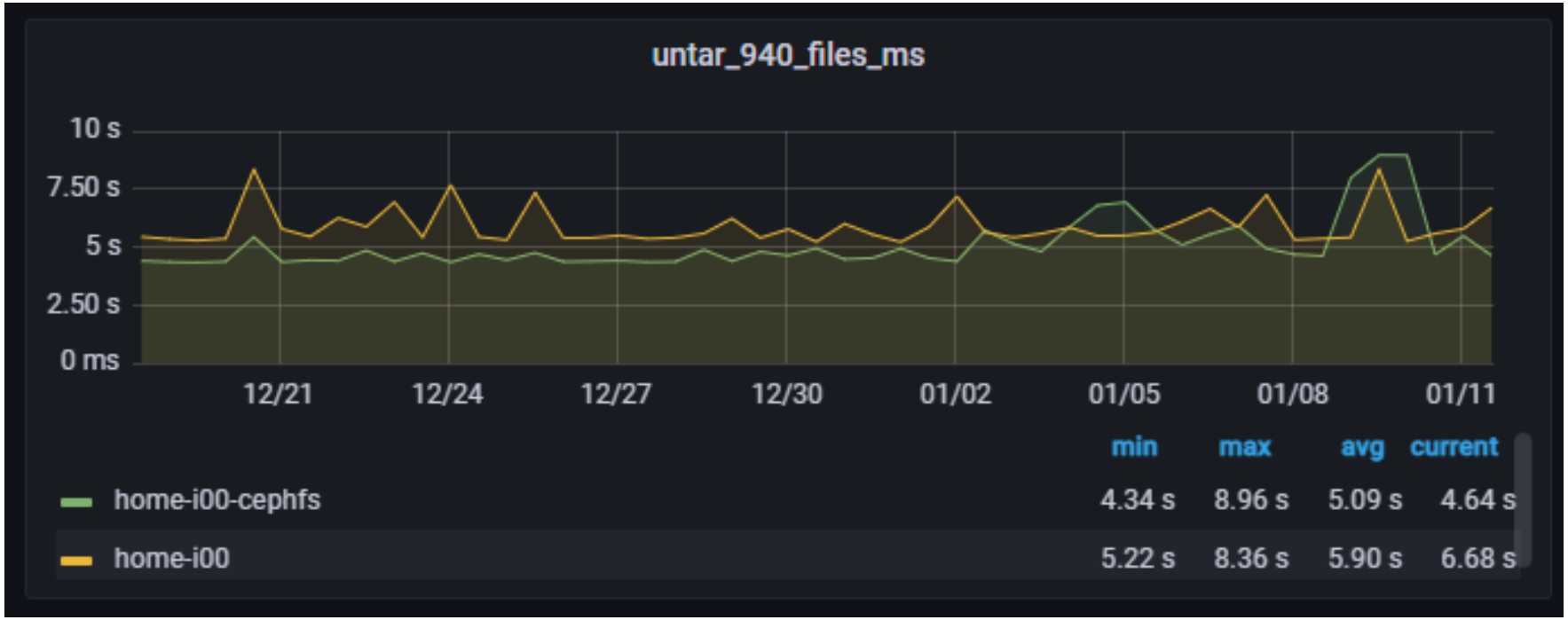
- When a frontend fails, we need to **redistribute** its virtual disks to the other remaining frontends.
- Operationally it is best if we can use as many other frontends in parallel
  - **Ex 1: with 1 virtual FST** -- that single FST is taken over by one other box, whose load now doubles.
  - **Ex 2: 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.

Production 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.

Production Testing Results

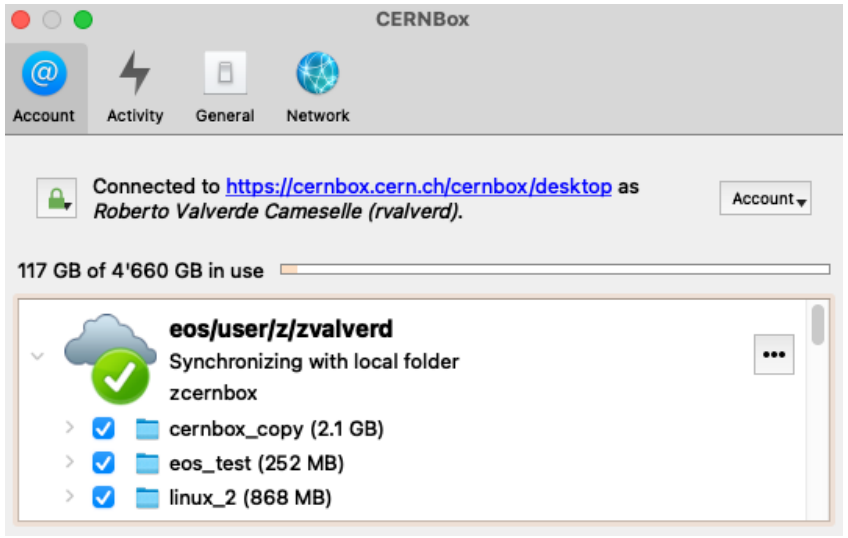
- The results roughly match what we observed on the PoC.

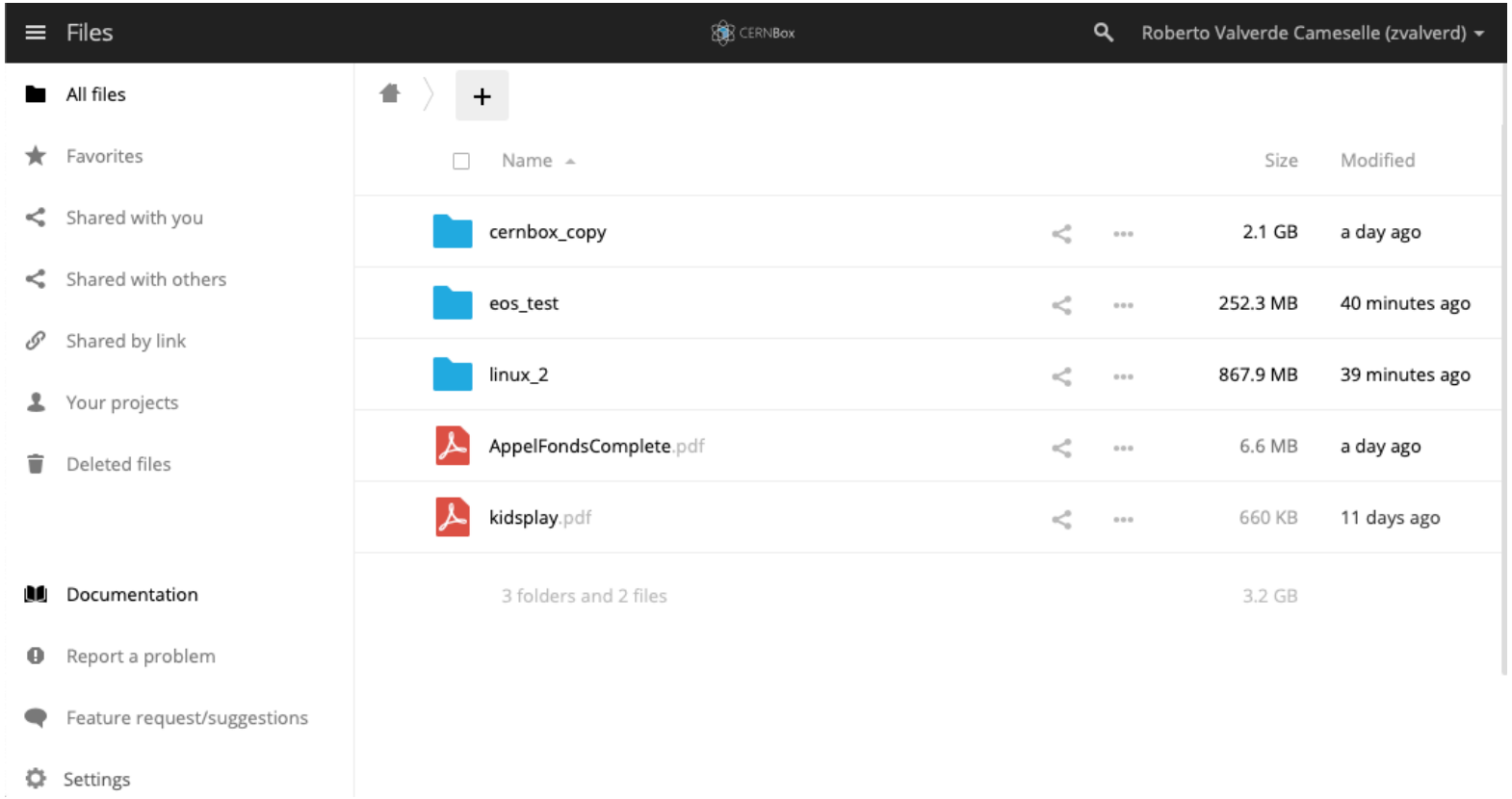


We enabled the same microtest suite in Dec 2021.

Production Testing Results

- 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 (II)

- 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 persistent volumes

Thank you!

home.cern