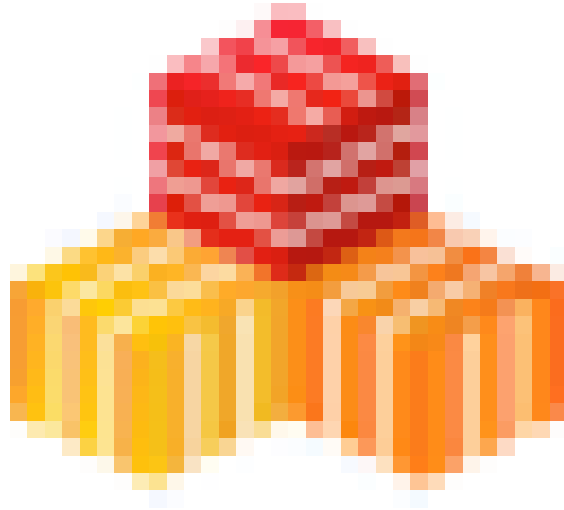


# EOS 2024 Workshop

Thursday, 14 March 2024 - Friday, 15 March 2024

CERN



## Book of Abstracts



# Contents

CERNBox/Reva deployment at CERN . . . . .	1
EOS Operation Status at KISTI Tier-1 for ALICE experiment . . . . .	1
Site Report Vienna Tier-2 . . . . .	1
EOS at the Fermilab LHC Physics Center . . . . .	1
Comparison of High-Performance Distributed File Systems on two Platforms: Linux and Windows . . . . .	2
eosxd-csi: Mounting EOS volumes in Kubernetes . . . . .	3
EOS 5.2 Status . . . . .	3
EOS III . . . . .	3
EOS ALMA9 Migration . . . . .	3
EOS on ARM . . . . .	4
EOSXd Evolution and CFSd . . . . .	4
EOS on SMR Status . . . . .	4
EOS Namespace Locking Evolution . . . . .	4
EOS FlatScheduler & Freespace Engine . . . . .	4
Group Balancer Free Space Engine . . . . .	4
HTTP Improvements and SciTags . . . . .	5
EOS and fixes & low level changes for openssl, xrootd & eosxd . . . . .	5
A new REST API Gateway for EOS . . . . .	5
EOS4Physics Operations . . . . .	5
Long Term Monitoring with Prometheus + Thanos . . . . .	5
EOS Benchmarks at CERN . . . . .	6
EOS Hardware Procurement . . . . .	6
CERNBox Update . . . . .	6

CIFS Update . . . . .	6
EOS on lxplus/lxbatch . . . . .	6
CTA Update . . . . .	6
Configuring EOS . . . . .	6
Implementing FSCK for Erasure Coded Files in EOS . . . . .	7
A native S3 interface EOS/XRootD . . . . .	7
Roadmap 2024 . . . . .	7
XRootD Status and Plans . . . . .	7
CERN Storage Day: Welcome to the IT Department . . . . .	7
CERN IT Storage and Data Management Group Overview . . . . .	7
Large Scale Disk Storage: EOS . . . . .	7
Large Scale Tape Storage: CTA . . . . .	8
User and Collaboration Storage: CERNBox . . . . .	8
Infrastructure Storage: Ceph . . . . .	8
More Filesystems: AFS,FILER,DFS,SAMBA,,... . . . .	8
Content Delivery Storage: CVMFS . . . . .	8
File Transfer Service (FTS) . . . . .	8
Scientific Data Management: Rucio . . . . .	9
EOS site report of the Joint Research Centre . . . . .	9
NDMSPC - EOS and N-Dimensional Analysis with ROOT, Enhanced by Web Interface and VR Visualization . . . . .	9
Purdue EOS status report . . . . .	9
Meet the team: EOS & XRootD . . . . .	10
Meet the team: Ceph FS . . . . .	10
Meet the team: CERN Tape Archive (CTA) Operations . . . . .	10
EOS Status at IHEP . . . . .	10
Roadmap Discussion . . . . .	10
Shared EOS instance at JINR . . . . .	11
ALICE File Consistency Check System: A new solution based on EOS FSCK . . . . .	11
Meet the team: FTS/CVMFS/DFS . . . . .	11

Hands-on CERNBox/Reva and Cephfs . . . . .	12
Hands-on CERNBox/Reva and EOS . . . . .	12
Meet the team: Ceph General . . . . .	12
Meet the team: Ceph S3 . . . . .	12
Meet the team: Ceph RBD . . . . .	12
Meet the team: CERN Tape Archive (CTA) Newcomers . . . . .	12
CVMFS: Openlab room . . . . .	13
Group Photo and Logistics . . . . .	13



**Meet the CERNBox Team / 1**

## **CERNBox/Reva deployment at CERN**

**EOS Operations / 3**

### **EOS Operation Status at KISTI Tier-1 for ALICE experiment**

**Authors:** Jeong Heon Kim<sup>1</sup>; Sang Un Ahn<sup>1</sup>

<sup>1</sup> *Korea Institute of Science & Technology Information (KR)*

**Corresponding Authors:** jh.kim@kisti.re.kr, sang.un.ahn@cern.ch

We have been running disk and archive (replacing tape) storage for ALICE experiment with EOS for several years. In 2023, we upgraded our EOS instances to v5 from v4. More recently we deployed EOS based on Podman (instead of Docker), which is adapted as a native container runtime in EL9 distributions. In this work, we present the current status of EOS operation at KISTI Tier-1 centre based on Podman container with systemd integration.

**EOS Operations / 4**

### **Site Report Vienna Tier-2**

**Author:** Erich Birngruber<sup>1</sup>

**Co-author:** Umit Seren<sup>1</sup>

<sup>1</sup> *Austrian Academy of Sciences (AT)*

**Corresponding Authors:** erich.birngruber@gmi.oeaw.ac.at, uemit.seren@gmi.oeaw.ac.at

We share our operational experience of running a converged EOS instance for 3 Experiments (CMS, Alice, Belle).

- In 2022 we've extended the capacity of the EOS cluster from 9 to 15 FSTs
- In 2023 we've successfully updated to EOS 5.2.2
- The system deployment and all operational tasks are fully automated with Ansible. Most recently a gap was closed with the automation of host certificate rotation.

**EOS Operations / 5**

### **EOS at the Fermilab LHC Physics Center**

**Author:** Dan Szkola<sup>1</sup>

<sup>1</sup> *Fermi National Accelerator Lab. (US)*

**Corresponding Author:** dszkola@fnal.gov

Fermilab has been running an EOS instance since testing began in June 2012. By May 2013, before becoming production storage, there was 600TB allocated for EOS. Today, there is approximately 13PB of storage available in the EOS instance.

The LPC cluster is a 4500-core user analysis cluster with 13 PB of EOS storage. The LPC cluster supports several hundred active CMS users at any given time.

An update of our current experiences and challenges running an EOS instance for use by the Fermilab LHC Physics Center (LPC) computing cluster. Planning and implementation of our upgrade to EOS 5 and moving to Almalinux before EOL for SL7.

**EOS Operations / 6**

## **Comparison of High-Performance Distributed File Systems on two Platforms: Linux and Windows**

**Author:** Gregor Molan<sup>1</sup>

<sup>1</sup> *Comtrade 360's AI Lab*

**Corresponding Author:** gregor.molan@cern.ch

Subtitle: Superiority of EOS-based Comtrade Distributed File System (CDFS) for Earth Observation Data Storage.

### **Introduction**

The rising quantity of data collected necessitates transitioning to the next generation of reliable, high-performance data storage solutions. Despite the clear need for high-performance storage across many industries, there has yet to be a consensus on the optimal high-performance storage technology that would suit all user needs.

The project's goal was speed-testing to provide a possibility to combine them with the currently available state-of-the-art technology results, technological capabilities and existing expertise.

### **Methodological approach**

We choose representative high-performance storage solutions compared on two different platforms for storage clients.

- Linux: - Ceph - EOS - IBM Spectrum Scale - Hadoop
- Windows: - Ceph - EOS-drive - EOS through Samba - EOS-wnc - Hadoop

An example of specifications of the testing environment is the following.

- *Management node:* - 32 threads - 2x Intel® Xeon® Silver 4208 Processor - 384 GB of RAM - 2x SSDs of 2 TB.
- *Storage nodes:* - 32 threads - 2x Intel® Xeon® Silver 4208 Processor - 64 GB of RAM - 1x SSD of 2 TB for the operation system - 6x HDDs of 2 TB for data.
- *Client nodes:* - 12 threads - 1x Intel® Core™ i5-12400 Processor - 16 GB of RAM - 1x SSD with 1 TB.



## Results

The results are from the testing performed separately in an isolated environment for each high-performance solution but on the same hardware. For each platform, there were three categories of tests related to file sizes: (1) small, (2) medium, and (3) large files.

According to the results, some high-performance file systems have evident advantages, as shown in our presentation. These results should be the starting point for an even more exact comparison between these file systems. They are good starting points in choosing the right high-performance file system.

## EOS Operations / 7

### **eosxd-csi: Mounting EOS volumes in Kubernetes**

**Author:** Robert Vasek<sup>1</sup>

<sup>1</sup> CERN

**Corresponding Author:** robert.vasek@cern.ch

Kubernetes is seeing a huge adoption across the cloud, including the one at CERN. Already used in production, the EOSxd-CSI driver exposes EOS volumes as regular PersistentVolumeClaims that can be mounted by containerized workloads. This talk will touch on storage stack in Kubernetes, how EOS falls into the mix, and finally a demo showing the CSI driver in action.

## EOS Development / 8

### **EOS 5.2 Status**

**Corresponding Author:** elvin.alin.sindrilaru@cern.ch

## EOS Development / 9

### **EOS III**

**Corresponding Author:** andreas.joachim.peters@cern.ch

III = Instance Inventory Implementation

We have added recently new tools to gather statics about storage hardware, resource usage, hardware lifecycle. These tools compute the virtual cost and value of user data and the hardware of an EOS instances.

The presentation will introduce these new tools.

## EOS Operations / 10

### **EOS ALMA9 Migration**

**Corresponding Author:** ioanna.vrachnaki@cern.ch

**EOS Development / 11**

## **EOS on ARM**

**Corresponding Author:** abhishek.lekshmanan@cern.ch

**EOS Development / 12**

## **EOSXd Evolution and CFSd**

**Corresponding Author:** andreas.joachim.peters@cern.ch

EOSXd is the filesystem client implemented as a FUSE filesystem to provide POSIX-like access to EOS. This is a crucial component for general usage of EOS. The presentation will highlight the problems we faced and advancements since the last workshop.

EOS CFSd is a FUSE pass-through filesystem implementation allowing to add missing features to any general POSIX filesystems. The presentation will highlight some possible use-cases and performance measurements e.g. an example how CFSd adds kerberos based mapping to a CephFS filesystem without a significant performance impact.

**EOS Development / 13**

## **EOS on SMR Status**

**Corresponding Author:** andreas.joachim.peters@cern.ch

In close collaboration with the IT procurement team we have conducted R&D with EOS and SMR disks. The findings, current status and outlook on integration of SMR disks will be covered in this presentation.

**EOS Development / 14**

## **EOS Namespace Locking Evolution**

**Corresponding Author:** cedric.caffy@cern.ch

**EOS Development / 15**

## **EOS FlatScheduler & Freespace Engine**

**Corresponding Author:** abhishek.lekshmanan@cern.ch

**EOS Development / 16**

## Group Balancer Free Space Engine

**Corresponding Author:** abhishek.lekshmanan@cern.ch

EOS Development / 17

## HTTP Improvements and SciTags

**Corresponding Author:** cedric.caffy@cern.ch

EOS Development / 18

## EOS and fixes & low level changes for openssl, xrootd & eosxd

**Corresponding Author:** david.smith@cern.ch

EOS Development / 20

## A new REST API Gateway for EOS

**Co-author:** Elvin Alin Sindrilaru <sup>1</sup>

<sup>1</sup> CERN

**Corresponding Authors:** andreea.prigoreanu@cern.ch, elvin.alin.sindrilaru@cern.ch

EOS Operations / 22

## EOS4PHysics Operations

**Corresponding Author:** maria.arsuaga.rios@cern.ch

EOS Operations / 23

## Long Term Monitoring with Prometheus + Thanos

**Corresponding Author:** roberto.valverde.cameselle@cern.ch

The Storage and Data Management Group at CERN manages 20 EOS instances corresponding to almost 1000 servers and 100,000 disks. Having a good monitoring and alerting system is crucial not only for day-to-day activities but also as a tool to record the evolution of our services throughout the time. In this talk an overview of the monitoring tools that are used will be presented specially in regards of long-term metric preservation.

**EOS Development / 24**

## **EOS Benchmarks at CERN**

**Corresponding Author:** andreas.joachim.peters@cern.ch

Before and after Run 3 have conducted several performance benchmarks as preparation of DC'24 and Run-3 in 2024 on various EOS instances at CERN. The presentation will report the findings of these benchmarks.

**EOS Operations / 25**

## **EOS Hardware Procurement**

**Corresponding Author:** luca.mascetti@cern.ch

**EOS Operations / 26**

## **CERNBox Update**

**Corresponding Author:** emmanouil.bagakis@cern.ch

**EOS Operations / 27**

## **CIFS Update**

**Corresponding Author:** glopresti@gmail.com

**EOS Operations / 28**

## **EOS on lxplus/lxbatch**

**EOS Operations / 29**

## **CTA Update**

**Meet the Teams / 30**

## **Configuring EOS**

**EOS Development / 31**

## **Implementing FSCK for Erasure Coded Files in EOS**

**Corresponding Author:** mano.segransan@cern.ch

**EOS Development / 32**

## **A native S3 interface EOS/XRootD**

**Corresponding Author:** mano.segransan@cern.ch

**Final Session / 33**

## **Roadmap 2024**

**Corresponding Authors:** elvin.alin.sindrilaru@cern.ch, andreas.joachim.peters@cern.ch

The presentation will summarize plans of the EOS team for 2024 and beyond.

**EOS Development / 34**

## **XRootD Status and Plans**

**Corresponding Author:** amadio@cern.ch

**CERN Storage Technologies / 35**

## **CERN Storage Day: Welcome to the IT Department**

**Corresponding Authors:** jakub.moscicki@cern.ch, dirk.duellmann@cern.ch

**CERN Storage Technologies / 36**

## **CERN IT Storage and Data Management Group Overview**

**Corresponding Author:** oliver.keeble@cern.ch

**CERN Storage Technologies / 37**

## **Large Scale Disk Storage: EOS**

**Author:** Andreas Joachim Peters<sup>1</sup>

**Co-author:** Elvin Alin Sindrilaru<sup>1</sup>

<sup>1</sup> CERN

**Corresponding Authors:** elvin.alin.sindrilaru@cern.ch, andreas.joachim.peters@cern.ch

CERN Storage Technologies / 38

## Large Scale Tape Storage: CTA

**Corresponding Author:** joao.afonso@cern.ch

CERN Storage Technologies / 39

## User and Collaboration Storage: CERNBox

**Corresponding Author:** giuseppe.lopresti@cern.ch

CERN Storage Technologies / 40

## Infrastructure Storage: Ceph

**Corresponding Author:** enrico.bocchi@cern.ch

CERN Storage Technologies / 41

## More Filesystems: AFS,FILER,DFS,SAMBA,,...

**Corresponding Authors:** enrico.bocchi@cern.ch, sebastian.bukowiec@cern.ch

CERN Storage Technologies / 42

## Content Delivery Storage: CVMFS

**Corresponding Author:** valentin.volk@cern.ch

CERN Storage Technologies / 43

## File Transfer Service (FTS)

**Corresponding Author:** joao.pedro.batista.lopes@cern.ch

**CERN Storage Technologies / 44****Scientific Data Management: Rucio****Corresponding Author:** martin.barisits@cern.ch**EOS Operations / 45****EOS site report of the Joint Research Centre****Author:** Armin Burger<sup>1</sup>**Co-author:** Franck Eyraud<sup>1</sup><sup>1</sup> *EC Joint Research Centre***Corresponding Authors:** armin.burger@ec.europa.eu, franck.eyraud@ext.ec.europa.eu

The Joint Research Centre (JRC) of the European Commission is running the Big Data Analytics Platform (BDAP) to enable the JRC projects to store, process, and analyze a wide range of data and disseminate data products. The platform evolved as a core service for JRC scientists to produce knowledge and insights in support of EU policy making.

EOS is the main storage system of the BDAP for scientific data. It is in increasing use at JRC since 2016. The Big Data Analytics Platform is actively used by more than 90 JRC projects, covering a wide range of data analytics activities. The EOS instance at JRC has currently a gross capacity of 34 PB with an additional increase planned throughout 2024.

The presentation will give an overview about EOS as storage back-end of the Big Data Analytics Platform. It covers the general setup and current status, experiences made, issues discovered, and an outlook of planned activities and changes in 2024.

**EOS Development / 46****NDMSPC - EOS and N-Dimensional Analysis with ROOT, Enhanced by Web Interface and VR Visualization****Author:** Martin Vala<sup>1</sup><sup>1</sup> *Pavol Jozef Safarik University (SK)***Corresponding Author:** martin.vala@cern.ch

This talk delves into the synergy of NDMSPC (NDimensional SPaCe), EOS at CERN, and N-dimensional histograms via the ROOT framework. Discover how a web interface serves as a powerful analysis tool, enabling dynamic queries on projections in N-dimensional space. Interact and visualize the data seamlessly with JSROOT and VR (aframe), ushering in a new era of immersive exploration in high-energy physics.

**EOS Operations / 47****Purdue EOS status report**

**Author:** Stefan Piperov<sup>1</sup>

<sup>1</sup> *Purdue University (US)*

**Corresponding Author:** spiperov@purdue.edu

Purdue University switched to EOS storage a couple of years back, and we are still exploring/appreciating the benefits of this distributed file-system. In this talk we will give a brief status report and will outline our plans for using EOS at the CMS Tier-2 center at Purdue.

**Meet the EOS/XRootD Team / 48**

## **Meet the team: EOS & XRootD**

**Meet the CEPH Team / 49**

## **Meet the team: Ceph FS**

Discussion on the Ceph Filesystem.

**Meet the CTA Team / 50**

## **Meet the team: CERN Tape Archive (CTA) Operations**

This session is open to all, but is geared towards sites with specific questions about CTA and tape operations.

**EOS Operations / 51**

## **EOS Status at IHEP**

**Author:** LI Haibo lihaibo<sup>None</sup>

**Corresponding Author:** lihaibo@ihep.ac.cn

The Institute of High Energy Physics (IHEP) has been utilizing the EOS storage system since 2016, with the current capacity nearing 60PB. We use EOS to provide both disk and tape storage solutions for HEP experiments. This report will introduce some of work we carried out in 2023, which include:

1. The adoption of EOS as SE for the WLCG Tier2, replacing the previous DPM system, and the provision of SE and CTA services for the LHCb Tier1 via EOS.
2. The upgrade of our system from EOS V4 to V5, which addressed the issue of high concurrency demands posed by the LHAASO experiment instance.
3. The establishment of a new ARM EOS cluster, which has reached a capacity of 2PB.

Additionally, we have conducted extensive testing and validation of EOS on AlmaLinux. Our plan is to migrate our entire computing cluster to AlmaLinux 9.2 by the end of June this year.



**Final Session / 52**

## Roadmap Discussion

**EOS Operations / 53**

### Shared EOS instance at JINR

**Author:** Nikita Balashov<sup>1</sup><sup>1</sup> *Joint Institute for Nuclear Research (RU)***Corresponding Author:** nikita.balashov@cern.ch

The Joint Institute for Nuclear Research (JINR) utilizes a diverse storage ecosystem, encompassing a number of dCache, EOS and Ceph storage instances to address varied scientific needs. In this report we focus on one of the largest EOS instances at JINR, operational since 2019 as a shared storage system for experimental data. Presently, it has the capacity of 22 PB and hosts over 7 PB of data from numerous experiments and projects. We give an overview of the current setup, development plans and share our experience of operating it.

**EOS Development / 54**

### ALICE File Consistency Check System: A new solution based on EOS FSCK

**Corresponding Author:** andreea.prigoreanu@cern.ch

Undoubtedly, the processing of ALICE experiment data relies on the quality and integrity of data. Currently, ALICE uses a distributed file crawler that periodically evaluates samples of files from each storage element in order to gather statistics about the number of corrupted or inaccessible files. The main issue with this solution lies in its inability to provide a comprehensive overview of a storage element status, as the analysis results are based on examining a random selection of files. This presentation will describe a new solution for the ALICE File Consistency Check System. The new approach will overcome the limitations of the file crawler by using the powerful consistency checking tools provided by EOS. The idea behind this project is to collect all the existing errors on an EOS instance from the reports generated by the FSCK command with the goal of reconciling the contents of the local storage with the central catalogue and, where possible, recover the lost content from other replicas. The output of the FSCK report command will be accessed through the new HTTP interface available in the latest versions of EOS. Hence, this solution not only produces a more accurate integrity analysis but automates the recovery of data loss as well.

**Meet the FTS/CVMFS/DFS Team / 55**

### Meet the team: FTS/CVMFS/DFS

CVMFS: 513/R-070 - Openlab Space  
FTS: 28/S-029  
DFS: 31/2-28

**Meet the CERNBox Team / 56**

## **Hands-on CERNBox/Reva and Cephfs**

**Meet the CERNBox Team / 57**

## **Hands-on CERNBox/Reva and EOS**

**Meet the CEPH Team / 58**

## **Meet the team: Ceph General**

Ask us anything about Ceph.

**Meet the CEPH Team / 59**

## **Meet the team: Ceph S3**

Discussion on ceph-provisioned S3 object storage.

**Meet the CEPH Team / 60**

## **Meet the team: Ceph RBD**

Discussion on Ceph RADOS Block Devices.

**Meet the CTA Team / 61**

## **Meet the team: CERN Tape Archive (CTA) Newcomers**

This session is open to all, but is geared towards sites who are newcomers to CTA or are interested in evaluating CTA for their archival storage needs.

**Meet the FTS/CVMFS/DFS Team / 62**

## **CVMFS: Openlab room**

63

## **Group Photo and Logistics**

**Corresponding Author:** [luca.mascetti@cern.ch](mailto:luca.mascetti@cern.ch)