EOS workshop

Monday 7 March 2022 - Thursday 10 March 2022

CERN

Book of Abstracts

Contents

EOS 1 / 1

Introduction

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

CERNBOX / 2

EOS log aggregation with Grafana Loki.

Author: Sami Mohamed Chebbi[1]

[1] *CERN*

Corresponding Author: sami.mohamed.chebbi@cern.ch

EOS provides a very detailed log system which provides useful information of all the user and system operations that are performed at any time. Each EOS daemon has its own log file and tracing operations that involve different components can be a time consuming task (MGM -> FST1 -> FST2). With Grafana Loki and Promtail, we setup a logging aggregation system that allows tracing operations across different EOS components from a central and unified interface, considerably reducing the time needed for debugging sessions or usage patterns investigation.

EOS 3 / 3

Prometheus EOS exporter

Author: Aritz Brosa Iartza[1]

Co-author: Roberto Valverde Cameselle [1]

[1] *CERN*

Corresponding Authors: roberto.valverde.cameselle@cern.ch, aritz.brosa.iartza@cern.ch

Prometheus is a modern, simple and scalable monitoring system with an easy to use query language based in labels. EOS Operators team has developed a fully-functional EOS Prometheus exporter in Golang to monitor all EOS metrics. This includes space, group, node, filesystem, I/O and namespace stats collectors. In this talk, the tool will be showcased and made available to the EOS Community.

CERNBOX / 4

EOS for CERNBox Report

Author: Roberto Valverde Cameselle[1]

[1] *CERN*

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

EOS provides the backend to CERNBox, the cloud sync and share service implementation used at CERN. EOS for CERNBox is storing 12PB of user and project space data across 9 different instances

running in multi-fst configuration. This presentation will give an overview of 2021 challenges, how we tried to address them and talk about the roadmap for the service for 2022.

CERNBOX / 5

CERNBox backup evolution

Authors: Gianmaria Del Monte[1]; Roberto Valverde Cameselle[1]

[1] *CERN*

Corresponding Authors: gianmaria.del.monte@cern.ch, roberto.valverde.cameselle@cern.ch

More than 300 million CERNBox files are processed daily using **cback** backup tool, which ensures that files are safely stored in a different geographical area and using a different storage backend. The backup tool has not stop evolving and was extended to support CephFS mount backup along with EOS mounts under the same infrastructure. This talk will present the current status of the project and the roadmap for this year, highlighting the plans for a self-service restore from CERNBox interface and mount access via the **cback-portal** backup gateway.

CERNBOX / 6

Converging Storage Layers with Virtual CephFS Drives for EOS/CERNBox

Authors: Andreas Joachim Peters[1]; Dan van der Ster[1]; Roberto Valverde Cameselle[1]

[1] *CERN*

Corresponding Authors: daniel.vanderster@cern.ch, roberto.valverde.cameselle@cern.ch, andreas.joachim.peters@cern.ch

The CERNBox service is currently backed by 13PB of EOS storage distributed across more than 3,000 drives. EOS has proven to be a reliable and highly performing backend throughout. On the other hand, the CERN Storage Group also operates CephFS, which has been previously evaluated in combination with EOS as a potential solution for large scale physics data taking [1]. This work seeks to further explore the operational benefits of a combined EOS/CephFS solution as a CERNbox backend. First, we present the functional validation work done using a canary instance and existing micro benchmarks. Next, we show how the solution was gradually introduced to production, observing the relative impacts of metadata and backend storage on user perceived small op performance. Finally, the qualitative impact of the solution is discussed: potential for enhanced QoS (e.g. policy driven low latency vs low-cost areas), simplication of hardware operations across the entire lifecycle, and how the work may enable future cloud-based deployments.

[1] https://doi.org/10.1007/s41781-021-00071-1

CTA 2 / 7

dCache integration with CTA

Author: Tigran Mkrtchyan[1]

[1] *DESY*

Corresponding Author: tigran.mkrtchyan@desy.de

The ever increasing amount of data that is produced by modern scientific facilities like EuXFEL or LHC puts a high pressure on the data management infrastructure at the laboratories. This includes poorly shareable resources of archival storage, typically, tape libraries. To achieve maximal efficiency of the available tape resources a deep integration between hardware and software components are required.

The CERN Tape Archive (CTA) is an open-source storage management system developed by CERN to manage LHC experiment data on tape. Although today CTA's primary target is CERN Tier-0, the data management group at DESY considers the CTA as an main alternative to commercial HSM systems.

dCache has an flexible tape interface which allows connectivity to any tape system. There are two ways that a le can be migrated to tape. Ether dCache calls a tape system specific copy command or through interaction via an in-dCache tape system specific driver. The latter has been shown (by NDGF, TRIUMF and KIT Tier-1s), to provide better resource utilization and efficiency. Together with the CERN Tape Archive team dCache developers working on seamless integration of CTA into dCache.

This presentation will show the design of dCache-CTA integration, current status and first test results at DESY.

CTA 1 / 8

CTA at AARNet

Author: Denis Lujanski Not Supplied[None]

Corresponding Author: denis.lujanski@aarnet.edu.au

In this presentation, we will report on how we at AARNet deployed CTA along with restic backup client as a backup/ archive solution for our production EOS clusters. The solution has been in production since late 2021. This presentation will aim to cover why we chose CTA, how CTA is deployed, and how it is integrated into our backup workflow.

CTA 2 / 9

Evaluation of CTA for use at Fermilab

Authors: Bo Jayatilaka[1]; Brian Yanny[2]; David Alexander Mason[1]; Eric Vaandering[1]; Ren Bauer[1]; Robert Illingworth[1]

[1] *Fermi National Accelerator Lab. (US)*

[2] *Fermilab*

Corresponding Authors: illingwo@fnal.gov, yanny@fnal.gov, renbauer@fnal.gov, dmason@fnal.gov, bo.jayatilaka@cern.ch, ewv@fnal.gov

Fermilab is the primary research lab dedicated to particle physics in the United States and also is home to the largest archival HEP data store outside of CERN. Fermilab currently employs a HSM based on Enstore, a Fermilab product, and dCache, for tape and disk, respectively. This Enstore+dCache HSM manages nearly 300 PB of active data on tape. Because of the necessary development work to ensure Enstore will work at expected HL-LHC data scales, Fermilab is exploring the use of CTA to replace it. We will report on the progress of this evaluation, including the deployment of CTA using containerized systems as well as the ability to read tapes formatted with CPIO tape wrappers.

EOS 1 / 10

EOS site report of the Joint Research Centre

Authors: Armin Burger[1]; Franck Eyraud[1]; Marco Scavazzon[1]

[1] *JRC*

Corresponding Authors: marco.scavazzon@ec.europa.eu, franck+cern@yrnm.net, armin.burger@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 process and analyze a wide range of data, providing knowledge and insights in support of EU policy making.

EOS is the main storage system of the BDAP for scientific data. It is in use at JRC since 2016. The gross capacity of 20 PB is currently in the phase of being increased by 7 PB, with an additional increase foreseen throughout 2022. The Big Data Analytics Platform is actively used by more than 50 JRC projects, covering a wide range of data analysis activities.

The presentation will give an overview about EOS as storage back-end of the Big Data Analytics Platform. It covers the general set-up, current status, experiences made, and an outlook of planned activities and changes in 2022.

EOS 1 / 12

EOS deployment at GRIF

Author: Emmanouil Vamvakopoulos[1]

[1] *Université Paris-Saclay (FR)*

Corresponding Author: emmanouil.vamvakopoulos@ijclab.in2p3.fr

In this communication, we are going to present the deployment project of the EOS storage software solution at the GRIF site. GRIF is a distributed site made of four (4) different subsites, in different locations of the Paris region. The worst network latency between the subsites is within 2-4 msec with 3 of them connected with a 100G connection. The objective is to consolidate the four (4) currently independent DPM instances into (a new) one EOS instance. The EOS capabilities such as the service structure (simple roles based on xrootd), the failover mechanism, the redundancy of metadata nodes, and the backend key/store database (quarkdb)) ensure with a straightforward installation and configuration the high availability requirements for the geographically distributed environment of GRIF. We are going to discuss the future EOS service structure and capacities at GRIF and present a summary of the organization of the EOS filesystems over legacy raid6 storage devices on heterogeneous hardware.

CTA 1 / 13

CTA Status and Roadmap

Author: Michael Davis[1]

[1] *CERN*

Corresponding Author: michael.davis@cern.ch

CTA entered into production at CERN in 2020 and physics data taking into CTA started in July 2021. 2022 will see the start of LHC Run-3, with combined experiment data rates up to 40 GB/s. This

presentation will give an overview of CTA's preparation and readiness for the upcoming Run, as well as a look forward to software features in the development pipeline.

EOS 1 / **14**

EOS migration tools

Author: Jaroslav Guenther[1]

[1] *CERN*

Corresponding Author: jaroslav.guenther@cern.ch

Migrating the AMS experiment data from EOSPUBLIC to EOSAMS02 stimulated development of tools which might be useful in general for similar exercises in the future. We will show the work in progress.

EOS 3 / **15**

EOS monitoring of finished transfers

Author: Jaroslav Guenther[1]

[1] *CERN*

Corresponding Author: jaroslav.guenther@cern.ch

Improving EOS monitoring of finished transfers. Hands-on eos io stat output.

EOS 3 / **16**

XRootD5 landscape

Author: Michal Kamil Simon[1]

[1] *CERN*

Corresponding Author: michal.simon@cern.ch

General update from XRootD project.

EOS 3 / **17**

Record and Replay

Author: Michal Kamil Simon[1]

[1] *CERN*

Corresponding Author: michal.simon@cern.ch

Presentation on the new recording plug-in that allows I/O sampling and the replay tool.

EOS 1 / 18

Native XRootD EC @ SLAC

Author: Michal Kamil Simon[1]

[1] *CERN*

Corresponding Author: michal.simon@cern.ch

Report on the latest tests done at SLAC with the native XRootD EC library.

EOS 3 / 19

xrdcp primer

Author: Michal Kamil Simon[1]

[1] *CERN*

Corresponding Author: michal.simon@cern.ch

A primer on xrdcp new (and old) features like zip append, metalling support, retries and many more.

CTA 1 / 20

Maintaining consistency in an EOSCTA system

Author: Richard Bachmann[1]

[1] *CERN*

Corresponding Author: richard.bachmann@cern.ch

This presentation summarizes the current effort to detect, and therebye subsequenly remedy, inconsistencies in the file metadata stored on EOS and CTA.
We show how we combine and validate EOSCTA namespaces in order to produce a summary of healthy files for experiments and a troubleshooting tool for operators.

CTA 1 / 21

The CTA project, team and community

Author: Oliver Keeble[1]

[1] *CERN*

Corresponding Author: oliver.keeble@cern.ch

Introduction to the CTA session.

CTA 1 / 22

Tape Drive Status Lifecycle

Author: Jorge Camarero Vera[1]

[1] *CERN*

Corresponding Author: jorge.camarero.vera@cern.ch

Explanation of the CTA Tape Drive status during a data transfer session.

CTA 1 / 23

Configuring user access control in CTA

Author: Volodymyr Yurchenko[1]

[1] *CERN*

Corresponding Author: volodymyr.yurchenko@cern.ch

CTA uses access mechanism provided by EOS and adds tape-specific layer. If one of these elements is misconfigured, a user won't be able to read a file, or, on the contrary, unauthorized access can be granted.

This talk explains how the combination of the ACL, Unix permissions and mount rules works in CTA. We show which tools we use for the permissions management and what are capabilities and limitations of our system.

CTA 1 / 24

EOSCTA file restoring

Author: Miguel Barros[1]

[1] *Universidade de Lisboa (PT)*

Corresponding Author: miguel.veloso.barros@cern.ch

This talk sumarizes the new file restoring feature of CTA, how it works, how to configure it, when it should be used and it's current limitations.

EOS 3 / 25

Encryption and Obfuscation Support in EOS

Author: Andreas Joachim Peters[1]

[1] *CERN*

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

With XRootD5 the on the wire protocol provides confidentiality of data inside the transport layer. However data files are human readable on storage nodes and can be accessed and downloaded by any EOS administrator and any person with read access. Filesystem level encryption on storage nodes does not solve this confidentiality problem.
To provide better data privacy the most recent versions of EOS support client and server side high-performance obfuscation and (with certain limitations) data encryption. The presentation will explain opportunities, challenges and limitations of the implementation.

CERNBOX / **26**

Share ACLs and EGroup-Ownership in EOS

Author: Andreas Joachim Peters[1]

[1] *CERN*

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

To consolidate the concept of sharing implemented inside EOS for any access protocol we are currently adding a new type of ACL which defines a 'share'. One of the new characteristics of a share ACL is that they are not influenced by POSIX or classic ACLs. We support additional ACL capabilities as 'can share'.
A second important new concept is the concept of ownership by an EGROUP. Ownership by individuals is problematic when people depart from CERN. This requires often a manual change of ownership of a departed person if files reside in shared spaces.
An EGROUP ownership is beneficial in particular in shared areas like project spaces, which are currently soved by the creation of one service account per project.

EOS 3 / **27**

Direct IO, IO priority and Bandwidth Policies in EOS

Author: Andreas Joachim Peters[1]

[1] *CERN*

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

In preparation for Run-3 we have faced the following problem: we have to balance the usage of IO resources between individual activities, which has led to the implementation of IO priorities and bandwidth regulation policies. While commissioning the ALICEO2 EOS instance we have observed, that write performance using the buffer cache is a bottleneck on storage nodes. Direct IO helps to improve write performance on storage nodes and additionally to reduce tails in data upload times by experiment DAQ systems. The presentation will explain and implementation and how to use these features in production.

CTA 2 / 28

CTA at RAL

Authors: Alastair Dewhurst[1]; Alison Packer[2]; George Patargias[1]; Tom Byrne[1]

[1] *STFC*

[2] *Science and Technology Facilities Council STFC (GB)*

Corresponding Authors: george.patargias@stfc.ac.uk, alastair.dewhurst@stfc.ac.uk, tom.byrne@stfc.ac.uk, alison.packer@stfc.ac.uk

This talk will present details of the deployment of Antares, the EOS-CTA service at RAL Tier-1, which replaces Castor.

EOS 3 / 29

Benchmarking TBits/s

Author: Andreas Joachim Peters[1]

[1] *CERN*

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

With 100GE technology and erasure coding we discovered new bottlenecks and challenges. This presentation will recap the state of the art of the ALICEO2 EOS instance and show benchmarks including a real and and replayed physics analysis use case.

EOS 3 / 30

Taming Batch Access to EOS at CERN

Author: Andreas Joachim Peters[1]

[1] *CERN*

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

Physics and CERNBOX instances at CERN are exposed to O(4) mount clients simultaneously. Overloads from batch access is not a new thing - since years the AFS filesystem suffers more or less frequently volume overloads. During overload episodes meta-data access at the MGM slows down significantly because thousands of batch nodes compete against few interactive clients and sync & share access. To give handles to the storage operation team EOS provides a set of access limitation features, which will be introduced in this talk.

EOS 1 / 31

EOS 5 during Run-3 Roadmap

Author: Andreas Joachim Peters[1]

[1] *CERN*

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

This presentation will introduce the roadmap for EOS5 during the Run-3 period.

EOS 1 / 32

EOS GroupBalancer improvements

Author: Abhishek Lekshmanan[None]

Corresponding Author: abhishek.lekshmanan@cern.ch

This is a talk introducing the GroupBalancer and what it does. We also cover about the current in place GroupBalancer improvements introduced from 4.8.78 release, the ways to configure this for deployments, some figures from existing deployements and what the roadmap for the future holds with these functionalities.

EOS 3 / 33

C++ Atomics: An Overview

Author: Abhishek Lekshmanan[None]

Corresponding Author: abhishek.lekshmanan@cern.ch

std::atomic introduced since C++11 is used as a building block for lock free programming. However while the default flags provide the maximum consistency; the do come with a performance penalty and may not be what you want in all cases. We will look under the hood, at a top level on what the processor sees when an atomic is encountered, the acquire and release semantics, which are fundamentally what mutexes use; and thus understand what the various memory order flags mean and when it is safe (or unsafe) to use them.

CTA 2 / 34

An HTTP Rest API as SRM replacement for tape access

Author: Cedric Caffy[1]

[1] *CERN*

Corresponding Author: cedric.caffy@cern.ch

Imagine a world where SRM is no longer needed to dialog with tape storage systems. A world where only one standard protocol can be used across the entire WLCG to access tape storage systems.

This dream will soon become reality on EOS…

After several discussions about the specifications of the new WLCG tape REST API, a prototype of the final API has been developed in EOS.

In order to give a good idea of the functionalities the API offers, I will do a comparison between the current XRootD workflows at CERN and the new HTTP ones that will be used once the REST API will be deployed.

EOS 2 / 35

WLCG tokens integration and support in EOS

Author: Elvin Alin Sindrilaru[1]

[1] *CERN*

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

.

EOS 1 / 36

EOS 5 highlights and functionality consolidation

Author: Elvin Alin Sindrilaru[1]

[1] *CERN*

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

.

EOS 1 / 37

EOS and XCache data access performance for LHC analysis at CERN

Authors: Dirk Duellmann[1]; Bernd Panzer-Steindel[1]; Markus Schulz[1]; Andrea Sciabà[1]; David Smith[1]

[1] *CERN*

Corresponding Authors: dirk.duellmann@cern.ch, markus.schulz@cern.ch, bernd.panzer-steindel@cern.ch, andrea.sciaba@cern.ch, david.smith@cern.ch

Physics analysis is done at CERN in several different ways, using both interactive and batch resources and EOS for data storage. In order to understand if and how the CERN computer centre should change the way analysis is supported for Run3, we performed several performance studies on two fronts: measuring the performance and utilisation levels of EOS with respect to the current analysis workloads, and looking at the performance of different storage configurations, including SSD-based and HDD-based XCache instances, with respect to specific, I/O intensive analysis workloads from ATLAS and CMS. The collected results indicate that the current infrastructure is adequate and works well below saturation, and that specific needs can be fulfilled by dedicated high performance/throughput servers. We expect this type of studies to continue and the CERN infrastructure to adapt to the evolving needs of the LHC analysis community.

EOS 1 / 38

Operation status of Custodial Disk Storage for the ALICE experiment

Author: Sang Un Ahn[1]

[1] *Korea Institute of Science & Technology Information (KR)*

Corresponding Author: sang.un.ahn@cern.ch

This is going to be a brief presentation regarding the operation status of Custodial Disk Storage (CDS) system provided for the ALICE experiment as a Tape. The CDS system is basically using EOS with its erasure coding implementation (RAIN) for the data protection. The CDS joined the WLCG Tape Challenges in the previous year and about a PB of data has been transferred from the experiment. A short reminder of system architecture and EOS RAIN configuration will be presented followed by operational activities and future plans.

CTA 2 / 39

CTA tape format support : BoF discussion

Author: Michael Davis[1]

[1] *CERN*

Corresponding Author: michael.davis@cern.ch

CTA uses the same tape format as CASTOR. There is interest from the community in adding support to read (but not write) tapes in alternate formats, such as OSM and Enstore. The main use case is to allow sites to migrate from their existing tape storage system to CTA without needing to physically repack all of their tapes.

This BoF session will be a round-table for stakeholders with an interest in reading tapes which are not in CASTOR/CTA format. The goal is to ensure that all the use cases are understood, and to converge on a technical solution/roadmap to add this functionality to CTA.

CERNBOX / 40

Samba: service evolution and experience with bind mounts

Author: Aritz Brosa Iartza[1]

Co-author: Giuseppe Lo Presti [1]

[1] *CERN*

Corresponding Authors: giuseppe.lopresti@cern.ch, aritz.brosa.iartza@cern.ch

In this talk we present the evolution of the CERNBox Samba service that we operate in front of EOS. An important recent change is the adoption of a new layout based on bind mounts: this allows to operate a smaller number of EOS mounts and to enable federating multiple EOS instances in a single namespace. We will discuss further measures adopted to address the ever increasing load from the Windows clients, and present an outlook of future extensions of the service.

EOS 1 / 41

EOS service @CERN 2022

Author: Maria Arsuaga Rios[1]

[1] *CERN*

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

General description of the EOS service @CERN

EOS 2 / 42

LHC Data Storage: RUN 3 Data Taking Commissioning

Author: Maria Arsuaga Rios[1]

[1] *CERN*

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

LHC Data Storage: RUN 3 Data Taking Commissioning

EOS 2 / 43

EOS and Ceph integration with Kubernetes

Author: Federico Fornari[None]

Co-authors: Alessandro Costantini [1]; Alessandro Cavalli [1]; Daniele Cesini [1]; Doina Cristina Duma [1]; Antonio Falabella [1]; Enrico Fattibene [1]; Luca Mascetti [2]; Lucia Morganti [1]; Andreas-Joachim Peters [2]; Andrea Prosperini [1]; Vladimir Sapunenko [1]

[1] *INFN-CNAF*

[2] *CERN*

Corresponding Author: federico.fornari@cnaf.infn.it

Due to the increasing interest on data management services capable to cope with very large data resources, allowing the future e-infrastructures to address the needs of the next generation extreme scale scientific experiments, the national center of INFN (Italian Institute for Nuclear Physics) dedicated to Research and Development on Information and Communication Technologies (CNAF) and the Conseil Européen pour la Recherche Nucléaire (CERN) joined their experiences on storage systems to evaluate and test different technologies for next-generation storage challenges.
The activity focused on the integration, using Kubernetes as orchestrator, of different storage systems (EOS and Ceph) with the aim to combine the high level scalability and stability of EOS services with the reliability and redundancy features provided by Ceph.
In particular, EOS services have been deployed as containers and orchestrated by Kubernetes, the well-known open-source container-orchestration system for automating computer application deployment, scaling and management.
The activity leverages in the possibility to integrate the two storage solutions by deploying them as containers and orchestrated by Kubernetes. In this respect, Kubernetes has been adopted to test different cluster-deployment scenarios (both on cloud and bare-metal) and assess their performances, bringing important improvements in terms of system operations, management and scalability.
The results obtained by measuring the performances of the different combined technologies, comparing for instance block device and file system as backend options provided by a Ceph cluster deployed on physical machines, will be shown and discussed.

EOS 3 / 44

EOS Windows client productisation

Author: Gregor Molan[1]

[1] *Comtrade 360's AI Lab*

Corresponding Author: gregor.molan@cern.ch

Context: Productisation of Windows native connection of EOS to Windows operating system.

Objectives: The professional implementation of the EOS with the Windows platform should allow seamless usage of EOS as a Windows local disk with all the EOS benefits, as it is low latency, high throughput, and high reliability.

Method: Implementation of the EOS client for the Windows platform is based on providing four communication channels:
- Communication with data stored on Windows
* Low-level reading and writing of data stored on Windows disks
* Access to Windows Active Directory
- Communication with EOS cluster
* High speed and secure data access on the EOS cluster using cURL.
* Secure access to the EOS server with the use of gRPC.
- Communication with Windows OS
* This is a kind of "meta-communication".
* The same user experience as it is for using EOS client on Linux.
- Communication with Windows users
* Another "meta-communication".
* EOS data presented as Windows drive.
Additionally, implementation of EOS client on Windows is based on the "performance aware development"based on continuous performance testing with immediate feedback according to possible performance issues.

Result: Developed high-performance EOS client on Windows appropriate supported with adequate software support and adequate selling business model. The decision of potential customers is supported with professional comparison results between EOS end other concurrent distributed file systems.

CTA 1 / 45

EOS and CTA Status at IHEP

Authors: Haibo li[1]; Qiuling Yao[2]; Yaodong Cheng[None]; Yujiang Bi[3]; Lu Wang[4]; Yaosong Cheng[5]

[1] *Institute of High Energy Physics Chinese Academy of Science*

[2] *IHEP*

[3] *Institute of High Energy Physics, Chinese Academy of Sciences*

[4] *Computing Center,Institute of High Energy Physics, CAS*

[5] *IHEp*

Corresponding Authors: chengys@ihep.ac.cn, biyujiang@ihep.ac.cn, yaoql@ihep.ac.cn, wanglu@ihep.ac.cn, lihaibo@ihep.ac.cn, chyd@ihep.ac.cn

EOS is now the main Storage System for IHEP experiments like LHAASO and JUNO. And Castor has been used for backup experiment data for a long time at IHEP, and has difficulty to satisfiy data backup requirement of new experiments like LHAASO, JUNO. As EOSCTA became stable to replace Castor in production, we started EOSCTA evaluation and the castor migration. In this talk, we will give a brief introduction of current EOS status at IHEP, and mainly talk about our effort on CTA deployment and CTA migration.

EOS 2 / 46

Data flowing on the Stream

Authors: Andreas-Joachim Peters[1]; Cristian Contescu[1]

[1] *CERN*

Corresponding Author: cristian.contescu@cern.ch

In this talk we will highlight the operational challenges we faced while bringing up a high-throughput EOS instance for the Run 3 ALICE data acquisition. The journey started in 2020 and we are still perfecting the instance to this day.
During this time all storage nodes got migrated from CentOS 7 to CentOS 8 and, later on, CentOS Stream 8, and not without inherent challenges which we are going to detail in this talk.

CERNBOX / 47

Enabling lightweight and federated accounts access in CERNBox

Author: Ishank Arora[1]

[1] *CERN*

Corresponding Author: ishank.arora@cern.ch

Access to CERNBox via social account providers and external emails provides a highly scalable and traceable mechanism to allow sharing of data and knowledge with people external to CERN, and encourage collaboration across boundaries and institutes. In this talk, we'll talk about how we adapted our service to accommodate such accounts with restricted scopes and describe the developments that were needed to our EOS storage connector to facilitate sharing of resources with them.

CERNBOX / 48

Managing locks in CERNBox and EOS

Author: Giuseppe Lo Presti[1]

[1] *CERN*

Corresponding Author: giuseppe.lopresti@cern.ch

This contribution illustrates how we have evolved file locking in CERNBox and EOS. Initially introduced to support Office online applications, the functionality has been extended to be an integral part of Reva, the engine powering CERNBox. We will describe the implementation in the EOS storage system, and the foreseen extensions to cover Linux file locks (flocks) as supported for FUSE and Samba clients.

EOS 2 / 49

EOS deployment at Purdue

Author: Stefan Piperov[1]

[1] *Purdue University (US)*

Corresponding Author: stefan.piperov@cern.ch

As part of its storage migration plan, the CMS Tier-2 center at Purdue University is preparing an EOS deployment of ~10PB, which will serve as the main Storage Element of the site, as well as a basis for the future Analysis Facility that's in development at the moment. We adopted a fully containerized approach with Kubernetes, which allows us to better share available hardware resources between different services. Our aim is to achieve high degree of data protection with low hardware overhead by utilizing Erasure Coding algorithms with high stripe size.

CERNBOX / 50

CERNBox: today and tomorrow

Author: Hugo Gonzalez Labrador[1]

[1] *CERN*

Corresponding Author: hugo.gonzalez.labrador@cern.ch

CERNBox is key enabler service built on top of EOS for users at CERN and beyond. The service is used by more than 37K users and stores over 15PB of data, representing all the user communities at the laboratory.

In this talk we will explain the current status of the service, the challenges we faced in 2021 and our vision for the future: CERNBox as the gateway for a federation of heterogeneous storage spaces.

EOS 1 / 51

High-capacity, high-throughput EOS storage for ALICE data taking

Author: Latchezar Betev[1]

[1] *CERN*

Corresponding Author: latchezar.betev@cern.ch

The ALICE detector and data acquisition system was substantially upgraded for Run3 and beyond. One of the main elements of the upgrade was the O2 processing cluster, which compresses the detector data in real time. The output of the compression is then written to EOS buffer for subsequent asynchronous data processing and archival. The requirements for the EOS storage are substantial: 120GB/sec write speed, 40GB/sec read speed and 100PB of total capacity. In addition, EOS must offer a sufficient data protection through erasure coding for the data until it is copied to archival storage. This presentation shows the ALICE experience with the EOS buffer deployment, testing and in production.

EOS 2 / 52

EOS at the Fermilab LHC Physics Center

Author: Dan Szkola[1]

[1] *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.

An update of our current experiences and challenges running an EOS instance for use by the Fermilab LHC Physics Center (LPC) computing cluster. The LPC cluster is a 4500-core user analysis cluster with 13 PB of EOS storage. This is an increase of about 71% over 2020. The LPC cluster supports several hundred active CMS users at any given time.

EOS 2 / **53**

EOS site report Vienna

Author: Erich Birngruber[1]

[1] *Austrian Academy of Sciences (AT)*

Corresponding Author: erich.birngruber@gmi.oeaw.ac.at

Update on the setup and operations at the Vienna Tier-2 site.

Community Feedback & Open Discussion / **54**

Community Feedback & Open Discussion

CTA 1 / **55**

How to enable EOS for tape

Corresponding Author: julien.leduc@cern.ch

An EOSCTA instance is an EOS instance commonly called a tape buffer configured with a CERN Tape Archive (CTA) back-end.
This EOS instance is entirely bandwidth oriented: it offers an SSD based tape interconnection, it can contain spinning disks if needed and it is optimized for the various tape workflows.
This talk will present how to enable EOS for tape using CTA and the Swiss horology gears in place to maximize tape hardware usage while meeting experiment workflow requirements.

EOS 3 / **56**

ScienceBox 2.0: From EOS Storage to Jupyter notebooks in Kubernetes

Authors: Enrico Bocchi[1]; Samuel Alfageme Sainz[1]; Aritz Brosa Iartza[1]; Abhishek Lekshmanan[None]

[1] *CERN*

Corresponding Authors: abhishek.lekshmanan@cern.ch, samuel.alfageme.sainz@cern.ch, enrico.bocchi@cern.ch, aritz.brosa.iartza@cern.ch

This contribution reports on the recent revamping of ScienceBox: The container-based stack for science with EOS, CERNBox, and SWAN services for Kubernetes-orchestrated clusters.
ScienceBox has been rebuilt from its foundations using modern cloud-native technologies for better service configuration and improved reliability, without compromising on deployment flexibility. Rethinking the whole package also allowed for better alignment of the production services at CERN with their container-based version.
Sciencebox has been tested and deployed on a variety of infrastructures, ranging from tiny deployments on developers' laptops to container orchestration platforms on commercial cloud providers with GPU accelerators and 100s of TBs of storage.

EOS 3 / 57

EOS Durability Summary

Author: Manuel Reis[1]

[1] *Universidade de Lisboa (PT)*

Corresponding Author: manuel.b.reis@cern.ch

EOS durability machinery is a set of (operator's) scripts, tools and EOS components to classify, monitor and repair unhealthy files. EOS filesystem check (fsck) was enabled in 2021, but one should keep track of the instances' state, and investigate root causes for the problems found.

CERNBOX / 58

Authentication Logic on /eos

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

Understanding the configuration and logic used by eosxd on /eos/ is not straight forward in particular in containerized environments. This short presentation tries to explain the basics.