EOS workshop

Monday 01 March 2021 - Thursday 04 March 2021

Book of Abstracts

Contents

OPS / 1

EOS at IHEP

Author: Haibo li[1]

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

Corresponding Author: lihaibo@ihep.ac.cn

In this presentation, we will report our current experiences and challenges with running EOS instances for used by IHEP CAS. Currently, IHEP has a total of 42PB storage, of which EOS accounts for 16PB, an increase of 10PB in 2020. At present, the LHAASO experiment mainly uses EOS as its mass storage system. In addition, the JUNO experiment has completed the construction of EOS testbed, and EOS is also considered after the evaluation. We will discuss our recent upgrade, including operating experiences, the progress of EOS CTA and EOS test on ARM. Finally, we will discuss our plans in the future.

EOS / 2

Powered by XRootD

Author: Michal Kamil Simon[1]

[1] *CERN*

Corresponding Author: michal.simon@cern.ch

XRootD is a distributed scalable system for low-latency file access. It is primary data access framework for the high-energy physics community, and the foundation of EOS project. In this contribution we give an overview of the release 5. In particular, we discuss the TLS based, secure version of the xroot/root protocol and the several enhance tailor made for EOS, like the so-called redirect collapse, I/O error recovery at MGM and kernel buffer support.

OPS / 3

EOS at the Joint Research Centre

Authors: Marco Scavazzon[1]; Armin Burger[2]; Franck Eyraud[3]; Pier Valerio Tognoli[4]

[1] *European Commission*

[2] *European Commission - Joint Research Centre*

[3] *Contractor of European Commission*

[4] *European Commission Joint Research Centre*

Corresponding Authors: pier-valerio.tognoli@ext.ec.europa.eu, franck+cern@yrnm.net, marco.scavazzon@ec.europa.eu, armin.burger@ec.europa.eu

The Joint Research Centre (JRC) of the European Commission has set up the Big Data Analytics Platform to enable the JRC projects to process and analyse big data, extracting knowledge and insights in support of EU policy making.

Since 2016, EOS is the main storage component of the platform. In 2020, the total gross capacity of this instance has reached 19 PiB.

The Big Data Analytics Platform is actively used by more than 40 JRC projects, covering a wide range of data analysis activities. To support the growing needs for data storage and processing capacity, the platform has been extended over the last year. Eight new FSTs have been added, for a space increase of 3.5 PiB. In addition, to increase the security of the platform, the team started to migrate the EOS nodes in a segregated VLAN.

The presentation will give an overview about the Big Data Analytics Platform and its EOS storage back-end, presenting the current status, experiences made and the issues identified.

EOS / 4

The virtue of composability

Author: Michal Kamil Simon[1]

[1] *CERN*

Corresponding Author: michal.simon@cern.ch

In this contribution we report on the new XRootD client declarative API that is in line with the modern C++ programming practices (ranges v3 inspired, support for lambdas and std::futures), offers much improved code readability and genuine composability.

EOS / 5

High throughput erasure coding with XRootD client

Author: Michal Kamil Simon[1]

[1] *CERN*

Corresponding Author: michal.simon@cern.ch

In this contribution we give the design details of the new Intel ISAL based XRootD erasure coding library and discuss the preliminary results obtained on the Alice O2 cluster.

CLOUD / 6

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 2020 challenges, how we tried to address them and talk about the roadmap for the service for 2021.

CLOUD / 7

Backing up CERNBox: Lessons learned.

Authors: Roberto Valverde Cameselle[1]; Joao Calado Vicente[1]

[1] *CERN*

Corresponding Authors: joao.calado.vicente@cern.ch, roberto.valverde.cameselle@cern.ch

CERNBox is the cloud sync and share service implementation at CERN which is used by physicists and collaborators across the globe. Data stored in CERNBox is becoming more and more critical and having a backup system is crucial for its preservation.

Two years ago we started a prototype of a backup orchestrator based on the open source tool restic. In 2020 the project reached its maturity and was enabled as the main production system for backup and restore. At the time being, more than 3PB of backup data is stored in S3 and more than 36K backup jobs are scheduled every day over the eosxd mounts.

In this presentation, we will give an overview of the project focusing on challenges, what we have learnt and talk about future plans.

OPS / 8

Experience of using EOS LRU and File Converter engines for write buffering in the Data Lake prototype

Authors: Andrey Kirianov[1]; Andrey Zarochentsev[2]

[1] *NRC Kurchatov Institute PNPI (RU)*

[2] *St Petersburg State University (RU)*

Corresponding Authors: andrey.kiryanov@cern.ch, andrey.zarochencev@cern.ch

In this talk we will share our experience in implementing write buffering with background stage-out of files from a site accessing the Data Lake prototype using EOS built-in LRU and File Converter engines. This study was aimed at improving resource usage for CPU-only sites by reducing the data stage-out overhead.

CTA / 9

A brief overview of the CTA mount scheduling logic

Author: Cedric Caffy[1]

Co-authors: Vladimir Bahyl [1]; Eric Cano [1]; Michael Davis [1]; David Fernandez Alvarez [1]; Aurelien Gounon [1]; Oliver Keeble [1]; Julien Leduc [1]; Steven Murray [1]; Volodymyr Yurchenko [2]

[1] *CERN*

[2] *National Academy of Sciences of Ukraine (UA)*

Corresponding Authors: volodymyr.yurchenko@cern.ch, michael.davis@cern.ch, eric.cano@cern.ch, david.fernandez.alvarez@cern.ch, vladimir.bahyl@cern.ch, oliver.keeble@cern.ch, aurelien.gounon@cern.ch, cedric.caffy@cern.ch, steven.murray@cern.ch, julien.leduc@cern.ch

Accessing data in a tape archival system can be costly in terms of time. The time taken to mount a tape into a drive, to position the tape head to a file and to unmount the tape when this file has been read can take more than 2 minutes.

A tape drive cannot be used to archive or retrieve data during the mounting and unmounting of a tape. We therefore need a solution to avoid mounting a tape when it is not worth it. Indeed, imagine a user who retrieves a single file from a tape and then 5 minutes later wants another file from the same tape. Without the CTA scheduling logic, the drive would lose twice the amount of mount, unmount and positioning time! A CTA tape server contains the scheduling logic that decides when to mount a tape in order to optimise drive usage for reading and writing data.

The aim of this presentation is to explain the different elements taken into account by the scheduler of each CTA tape server to decide whether or not a tape is worth mounting.

CTA / 10

ALICE and the CTA Garbage Collectors

Author: Steven Murray[1]

Co-authors: Vladimir Bahyl [1]; Eric Cano [1]; Michael Davis [1]; David Fernandez Alvarez [1]; Aurelien Gounon [1]; Oliver Keeble [1]; Julien Leduc [1]; Volodymyr Yurchenko [2]; Cedric Caffy [1]

[1] *CERN*

[2] *National Academy of Sciences of Ukraine (UA)*

Corresponding Authors: michael.davis@cern.ch, volodymyr.yurchenko@cern.ch, david.fernandez.alvarez@cern.ch, canoc3@cern.ch, julien.leduc@cern.ch, aurelien.gounon@cern.ch, steven.murray@cern.ch, oliver.keeble@cern.ch, cedric.caffy@cern.ch, vladimir.bahyl@cern.ch

In the standard layout of an EOSCTA deployment there are two SSD buffers in front of the tape drives. One is called the “default” space and is used for writing files to tape and the other is called the “retrieve” space and is used for reading them back. These buffers prevent direct file transfers between HDDs and tape drives. Such direct transfers would suffer from the unacceptable performance penalties incurred by mixing the preferred access patterns of disk and tape. A HDD usually has thousands of concurrently open files with data bandwidth being shared across them. A tape drive on the other hand simply reads or writes one file at a time at high speed. The mechanical thrashing of a HDD that is associated with thousands of open files may be acceptable to end users but it is unacceptable to a tape drive requiring high bandwidth for a single file.

The lifetime of the files within the two SSD buffers is relatively short. Files being written to tape are deleted from the default space as soon as they have been safely stored on tape. Files being retrieved from tape are deleted from the retrieve space as soon as they have been copied to their destination system.

The layout of the EOSCTA deployment for ALICE experiment is different from the standard layout because it has an additional HDD disk cache called the “spinners” space which sits between the retrieve SSD buffer and the ALICE end users. The spinners space is a true disk cache because the lifetime of files within it are relatively long. These files are automatically deleted by one of two garbage collectors when space needs to be freed up in order to make room for newly retrieved files. This workshop presentation describes the ALICE HDD disk cache and the automatic garbage collectors that free up space within it.

EOS / 11

EOS Version 5 Timeline and Roadmap

Authors: Andreas Joachim Peters[1]; Elvin Alin Sindrilaru[1]

[1] *CERN*

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

We will present an overview of the upcoming EOS Version 5 release (Diopside) and a development roadmap.

OPS / **12**

EOS on CephFS

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

[1] *CERN*

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

This presentation will highlight how to deploy EOS effectively using CephFS as a storage backend, the basic operational aspects for EOS and CephFS and performance expectations.

EOS / **13**

XRootD I/O Server Benchmarking on flash and disk

Author: Andreas Joachim Peters[1]

[1] *CERN*

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

This presentation will summarize few bandwidth and IOPS measurements using root:// and http:// protocol using XRootD Version 5 in front of disk, NVMe, SSDs and CephFS and an outlook with posisble future improvements.

CTA / **14**

EOS+CTA WorkFlows: Tape Archival and Retrieval

Authors: Michael Davis[1]; Vladimir Bahyl[None]; Cedric Caffy[None]; Eric Cano[None]; David Fernandez Alvarez[None]; Aurelien Gounon[None]; Oliver Keeble[None]; Julien Leduc[None]; Steven Murray[None]; Volodymyr Yurchenko[None]

[1] *CERN*

Corresponding Author: michael.davis@cern.ch

The CERN Tape Archive (CTA) is the tape back-end to EOS. EOS provides an event-driven interface, the WorkFlow Engine (WFE), which is used to trigger the processes of archival and retrieval. When EOS is configured with its tape back-end enabled, the CREATE and CLOSEW (CLOSE Write) events are used to trigger the archival of a file to tape, while the PREPARE event triggers the retrieval of a file from tape and the creation of a disk replica.

This talk will present the details of these tape-related workflows, including the state machine for the processes of archival and retrieval, and the metadata which is communicated between EOS and CTA.

CTA / 15

CTA best practices for data taking workflows

Author: Volodymyr Yurchenko[1]

Co-authors: Aurelien Gounon [2]; Cedric Caffy [2]; David Fernandez Alvarez [2]; Eric Cano [2]; Julien Leduc [2]; Michael Davis [2]; Oliver Keeble [2]; Steven Murray [2]; Vladimir Bahyl [2]

[1] *National Academy of Sciences of Ukraine (UA)*

[2] *CERN*

Corresponding Authors: julien.leduc@cern.ch, cedric.caffy@cern.ch, eric.cano@cern.ch, volodymyr.yurchenko@cern.ch, david.fernandez.alvarez@cern.ch, oliver.keeble@cern.ch, michael.davis@cern.ch, vladimir.bahyl@cern.ch, steven.murray@cern.ch, aurelien.gounon@cern.ch

There is significant diversity in the Data Acquisition (DAQ) systems of the non-LHC experiments supported at CERN. Each system can potentially have its own data taking software and helper scripts, and each can use their preferred data transfer commands and apply different checks and retry policies. The task of the CERN Tape Archive (CTA) team is to provide support for all of these different use cases and to define the best practices for integrating with an EOSCTA instance.

In this talk we will present an overview of typical DAQ workflows and discuss which protocols, commands and APIs we recommended to use with EOSCTA. We will provide examples of submitting archive and retrieve requests using FTS and XRootD tools. We will explain how to monitor the status of a file on tape. We will explain the best way to ensure a file is safely stored on tape. We will also give an overview of the CTA authentication policies.

CTA / 16

Running an EOS instance with tape on the back

Author: Julien Leduc[1]

Co-authors: Aurelien Gounon ; Cedric Caffy ; David Fernandez Alvarez ; Eric Cano ; Michael Davis [1]; Oliver Keeble ; Steven Murray [1]; Vladimir Bahyl [1]; Volodymyr Yurchenko

[1] *CERN*

Corresponding Authors: julien.leduc@cern.ch, vladimir.bahyl@cern.ch, michael.davis@cern.ch, steven.murray@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 disks if needed and it is optimized for the various tape workflows.

This talk will present the specific details of the EOS tape buffer tweaks and the Swiss horlogery gears in place to maximize tape hardware usage while meeting experiment workflow requirements.

OPS / 17

EOS service @CERN

Author: Maria Arsuaga Rios[1]

Co-author: Luca Mascetti [1]

[1] *CERN*

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

General description of the EOS service @CERN

EOS / **18**

Status and deployment of EOS XrdHttp with TPC support

Author: Elvin Alin Sindrilaru[1]

[1] *CERN*

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

Overview of the XrdHttp integration with EOS together with token support.

CLOUD / **19**

Multi-lock support for Office offline and online applications in CERNBox

Author: Giuseppe Lo Presti[1]

[1] *CERN*

Corresponding Author: giuseppe.lopresti@cern.ch

This short contribution will describe the offer of Office online and offline applications for our CERNBox users, and how we support their interplay to facilitate users' collaboration.

OPS / **20**

EOS for Physics at CERN in 2020

Author: Cristian Contescu[1]

Co-author: Maria Arsuaga Rios [1]

[1] *CERN*

Corresponding Author: cristian.contescu@cern.ch

Our team is in charge of providing storage and transfer services for the LHC and non-LHC experiments at CERN. In this presentation we are going to walk you through the activities of the EOS operations team at CERN in 2020. We are going to focus on the achievements, hurdles and lessons learned throughout the past year.

EOS / **21**

EOS developments overview 2020

Author: Elvin Alin Sindrilaru[1]

[1] *CERN*

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

Summary of the most important development done throughout 2020.

OPS / **22**

EOS for ALICE O2 - HW setup and OS challenges

Author: Cristian Contescu[1]

Co-authors: Michal Kamil Simon [1]; Andreas Joachim Peters [1]

[1] *CERN*

Corresponding Authors: michal.simon@cern.ch, andreas.joachim.peters@cern.ch, cristian.contescu@cern.ch

This presentation will briefly showcase the ALICE O2 HW setup for the pilot storage nodes and the OS challenges we have faced when trying to tweak it for maximum performance, in view of ALICE's Run3 data taking.

CLOUD / **23**

SAMBA: lessons learned

Authors: Aritz Brosa Iartza[1]; Giuseppe Lo Presti[1]

[1] *CERN*

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

Last year it was already presented the architecture of the SAMBA service within CERNBox, this year the topic will be the journey to improve the service, problems faced and the lessons learned for the future.

OPS / **24**

LHC Data Storage: RUN-3 preparation

Author: Maria Arsuaga Rios[1]

[1] *CERN*

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

LHC Data Storage: RUN-3 preparation

HANDS-ON / **25**

Practical use cases for eos-ns-inspect tools

Authors: Maria Arsuaga Rios[1]; Cristian Contescu[1]; Roberto Valverde Cameselle[1]

[1] *CERN*

Corresponding Authors: roberto.valverde.cameselle@cern.ch, cristian.contescu@cern.ch, maria.arsuaga.rios@cern.ch

Practical use cases for eos-ns-inspect tools

OPS / **26**

What's coming for EOS Data Durability

Author: Manuel Reis[1]

Co-author: Maria Arsuaga Rios [2]

[1] *Universidade de Lisboa (PT)*

[2] *CERN*

Corresponding Authors: maria.arsuaga.rios@cern.ch, manuel.b.reis@cern.ch

EOS Data Durability is a set of tools that automatically detects and repairs problematic files to ensure that data is not lost or compromised.

CLOUD / **27**

EOS meets Helm: K8s-based instances for testing and external deployments

Authors: Fabio Luchetti[1]; Enrico Bocchi[1]; Samuel Alfageme Sainz[1]

[1] *CERN*

Corresponding Authors: samuel.alfageme.sainz@cern.ch, enrico.bocchi@cern.ch, fabio.luchetti@cern.ch

This contribution reports on the recent development of Helm charts for the deployment of EOS in kubernetes-orchestrated clusters. An excursus on the state of the art will lead to the underlying motivations and the description of several use cases where a container-based deployment of EOS comes in handy, from disposable clusters for internal testing to installations in commercial clouds for HEP analysis and education.

EOS / **28**

Windows drive for EOS-wnc

Author: Gregor Molan[1]

[1] *COMTRADE D.O.O (SI)*

Corresponding Author: gregor.molan@cern.ch

Context: Windows nature connection of EOS-wnc to Windows operating system.

Objectives: The connection of the EOS-wnc on Windows platform should be as it is for Windows local disks, external disk storages, it means as a Windows disk driver letter.

Method: A storage on Windows operating is presented as a "disk drive letter". Architecture of Windows storage drivers has following layers as follows:

1. IRP for Upper-filter driver
2. IRP for Storage-class driver
3. SRB for Lower-filter driver
4. SRB for Storage port driver
5. SRB for Bus-specific commands

where
IRP (I/O request packets): kernel mode structures that are used by Windows Driver Model (WDM)
SRB (SCSI Request Block): SCSI command descriptor blocks (CDBs)

For EOS-wnc is implemented "thin" Windows disk driver.

Result: Windows driver software is low-level software and bugs in low-level software can be extremely painful; bugs can cause a loose of all data on disk drive including operating system. Therefore, the EOS-wnc Windows drivers implementation is a "thin driver" to maximize stability and security of this part of software.

EOS / **29**

EOS-wnc console

Author: Gregor Molan[1]

[1] *COMTRADE D.O.O (SI)*

Corresponding Author: gregor.molan@cern.ch

Context: EOS-wnc console for EOS client on Windows operating system.

Objectives: The usage of the EOS-wnc on Windows platform with the functionalities of the EOS client on Linux platform should be on the same level as the usage of EOS Linux client.

Method: EOS client can be used as a set of command line interface (CLI) commands, where each EOS command is executed independently, or through a EOS client console, where EOS commands are executed in a console.

Result: EOS-wnc includes a Windows command eos.exe that is a console for all EOS client commands. Beside all features from EOS client console on Linux, there is the additional user-friendly feature: command-line completion (tab completion). It provides completion of commands, completion of command arguments, and completion of directory and file names. The last completion feature allows user to get "on the fly" directory content for any EOS command without previously using EOS ls command.

EOS / **30**

OSD-Model implementation on EOS-wnc

Author: Gregor Molan[1]

[1] *COMTRADE D.O.O (SI)*

Corresponding Author: gregor.molan@cern.ch

Context: Optimal Software Implementation Model (OSD-Model) is to supervise and control development of EOS-wnc, where EOS-wnc is an important extension of Linux based EOS system for Windows platform.

Objectives: OSD-Model is used to manage development process to assure appropriate performance of the EOS-wnc on Windows platform on the same level as the performance of EOS Linux client. EOS-wnc has the same functionalities as the EOS client on Linux platform, where EOS Linux fulfil the highest demands for CERN experiments.

Method: Development process is managed with an OSD-Model in such a way, that graph vertices are requested functionalities and graph edges are test cases and f-influences between requested functionalities. Graph weights in the functionality graph are
(a) estimations for development costs for functionalities and functionality influences,
(b) estimations for test costs for functionality influences,
(c) functionality and f-influence significance,
(d) value for end user related to functionalities and f-influences.

Result: For each of required EOS-wnc command is defined their value and their significance. Defined are influences (f-influences) between required EOS-wnc commands with their values and their significance, similarly as values and significance of functionalities. According to available development resources, that could be changing during the development process, algorithms of the proposed OSD-Model determine the set of functionalities and f-influences to get the optimal EOS-wnc. In this case, the optimal EOS-wnc is the software that is at least on the same level of performance as the EOS Linux client.

OPS / **32**

AARNet FST Investigations

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

CTA / **33**

AARNet CTA Investigations

Corresponding Author: david.jericho@aarnet.edu.au

OPS / **34**

The first disk-based custodial 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

We present a disk-based custodial storage for the ALICE experiment at CERN to preserve its raw data alternative to tape with the EOS QRAIN. In this presentation, we describe the detailed system deployment of disk-based custodial storage, the integration to the ALICE experiment and the current status of system monitoring such as hardware error detection and power consumption measurement.

HANDS-ON / **35**

EOS in 5 Minutes

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

How to install EOS in 5 minutes and run it.

CLOUD / **36**

Making Reva talk to EOS: ultimate scalability and performance for CERNBox

Author: Fabrizio Furano[1]

[1] *CERN*

Corresponding Author: fabrizio.furano@cern.ch

The Reva component, at the heart of the CERNBox project at CERN will soon get new plugins that build on the experience accumulated with the current production deployment, where its data is stored centrally in EOS at CERN.

Making Reva natively interfaced to EOS through high performance gRPC and standard HTTPS interfaces will open a new scenario in terms of scalability and manageability of the CERNBox service, whose requirements in terms of data will continue to grow in the next decade. In this contribution we will technically introduce this near-future scenario.

HANDS-ON / **37**

EOS and SqashFS

Co-author: Andreas Joachim Peters [1]

[1] *CERN*

Corresponding Authors: andreas.joachim.peters@cern.ch, jaroslav.guenther@cern.ch

A quick tutorial, how to use squashfs images for software/small file distribution.

CLOUD / **38**

CERNBox: Horizon 2030

Author: Hugo Gonzalez Labrador[1]

[1] *CERN*

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

CERNBox is a sync and share collaborative cloud storage solution built at CERN on top of EOS. The service is used by more than 37K users and stores over 12PB of data. CERNBox has responded to the high demand in our diverse community to an easily and accessible cloud storage solution that provides integrations with other CERN services for big science: visualisation tools, interactive data analysis and real-time collaborative editing.

In this presentation we take a glimpse to the evolution of the service and the vision we have for it for the next decade.

EOS / **39**

SRE fundamentals in EOS

Author: Hugo Gonzalez Labrador[1]

[1] *CERN*

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

The EOS system is an advanced distributed storage system that deals with many extreme uses-cases (massive data injection from the LHC, latency-critical online home directories and massive throughput accesses from batch farms).

EOS implements many site reliability engineering best practices to support these uses cases at scale and also to support the work done by the operations team maintaining the production clusters.

In this presentation we explain some of the functionalities implemented in the core of EOS (logging, retry mechanism, QoS) that allows a smooth operation of the service while accommodating the diverse use-cases cited above.

EOS / **40**

EOS Basic Concepts and Design

Authors: Andreas Joachim Peters[1]; Elvin Alin Sindrilaru[1]

[1] *CERN*

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

This talk will give summary of the main concepts and features of EOS as a storage system.

- namespace design
- user concept
- access control
- access protocols

- high availability
  - meta data
  - data
- scheduling

EOS / **41**

Welcome & Technicalities & Introduction

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

HANDS-ON / **42**

EOS Operations: bits and pieces

Author: Roberto Valverde Cameselle[1]

[1] *CERN*

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

In this presentation we will be sharing some tips and recommendations about different operational procedures on EOS, from techniques to reduce the load on FST's system disk to how to use geoscheduler mechanism for draining and for adding capacity to the instances.

OPS / **43**

A scheme to implement local server computation on EOS system based on Xrootd plug-in

Authors: minxing zhang[1]; Yaodong Cheng[2]; Haibo li[3]; Yujiang Bi[4]

[1] *The Institute of High Energy Physics of the Chinese Academy of Sciences*

[2] *IHEP, CAS*

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

[4] *Chinese Academy of Sciences (CN)*

Corresponding Authors: chyd@ihep.ac.cn, zhangmx@ihep.ac.cn, lihaibo@ihep.ac.cn, biyj@cern.ch

Particle physics computing model has a kind of high statistical calculation, such applications need to access a large amount of data for analysis, the data I/O capability is very high requirements.For example, the LHAASO experiment generates trillions of events each year, and the large raw data needs to be decode to encode and mark before it can be analyzed. In this process, very high I/O bandwidth is required, otherwise an I/O bottleneck will form.When using the EOS file system, the user cannot know the physical storage location of the file, and when the user needs to access the file, it needs to search the MGM, transfer the file from the FST to the client, and the client provides the target file to the user.In this process, if the user needs to perform such IO intensive operations as mentioned above, there are two limitations on I/O bandwidth, one is the storage node's hard disk read and write efficiency, the other is the network bandwidth between the FST and the client.In this case, if the data storage unit and the computing unit can be integrated into one, the data handling can be significantly reduced, and the parallelism and energy efficiency of computing can be greatly

improved.Currently, the potential of this kind of integrated memory and computing storage is attracting the attention of many companies and standards bodies.SNIA has formed a working group to establish standards for interoperability between computable storage devices, and the OpenFog Consortium is also working on standards for computable storage.
Therefore, we propose a scheme to implement local server computation on EOS system based on Xrootd plug-in.Flags can be added after a file is accessed when a user needs to use computable storage.After receiving the access request, the client will forward the request to the FST where the file is located and perform the default decode calculation in the background on the FST.After testing, we found that using this method to simultaneously decode 10 1G raw files stored on the same FST can save about 45.9% of the time compared to the traditional method.The next work plan is to sink the computable module onto the hard disk to reduce the CPU consumption of the FST, and to customize the acceleration module on the hardware to increase the speed of the computation.

HANDS-ON / **44**

OCIS meets EOS

Author: Samuel Alfageme Sainz[1]

[1] *CERN*

Corresponding Author: samuel.alfageme.sainz@cern.ch

In this hands-on we show how to connect an out-of-the-box OCIS (ownCloud Infinite Scale) and connect it to an existing EOS instance.

OPS / **45**

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 11PB 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 11 PB of EOS storage. This is an increase of about 80% over 2018. The LPC cluster supports several hundred active CMS users at any given time.

OPS / **46**

WebEOS for websites hosting

Corresponding Author: alexandre.lossent@cern.ch

This presentation will briefly describe the usage of EOS for website hosting at CERN.

OPS / **47**

EOS at the Austrian T2

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

The institutes at the Vienna Biocenter (GMI, IMBA, IMP) have run HPC services for their life sciences research for several years. With our new infrastructure “CLIP”, additional partners came on board in 2019, including the Austrian high-energy physics community.

Beginning in 2020 the Austrian grid T2 setup was modernized and based on the CLIP infrastructure.

We run a converged EOS instance for Alice, Belle2 and CMS and want to share our experience in getting our setup into production.

OPS / **48**

EOS - ALICE choice for Run3 + large O2 disk buffer

Corresponding Author: latchezar.betev@cern.ch

STORAGE ANALYTICS / **49**

Applying Data Analytics to Storage Systems and Data Access

Corresponding Authors: andrea.sciaba@cern.ch, olga.chuchuk@cern.ch, federico.gargiulo@cern.ch, asciaba@cern.ch