CERN Tape Archive Workshop : CTA 2025

Thursday 27 March 2025 - Friday 28 March 2025

CERN

Book of Abstracts

Contents

CTA Operations and Site Reports / 1

CTA at PIC. Site report.

Author: Esther Acción[1]

Co-authors: Elena Planas [1]; Jordi Casals [1]

[1] *PIC*

Corresponding Authors: eplanas@pic.es, eaccion@pic.es, jcasals@pic.es

This year, we have upgraded CTA to Alma 9 and worked on automating the platform installation using Puppet. Additionally, we have tested the CTA Operations modules along with other features, such as policy mount rules. In February-March, we plan to conduct performance tests by allocating more resources to our test environment.
Unfortunately, we are still facing issues with humidity in the room, and our plans to build a dedicated space for a new library have been delayed due to administrative constraints.

CTA Operations and Site Reports / 2

CTA Status at DESY

Author: Mwai Karimi[None]

Corresponding Author: mwai.karimi@desy.de

CTA has been in operation at DESY for nearly two years, during which time additional experiments have been integrated, and over 80PB of data has been written to tape. This presentation will provide an overview of recent developments, along with insights and experiences from running dCache+CTA so far

CTA Development and Roadmap / 3

Progress Report on CTA Scheduler DB

Author: Jaroslav Guenther[1]

[1] *CERN*

Corresponding Author: jaroslav.guenther@cern.ch

The CERN Tape Archive (CTA) scheduling system manages the workflow of archive, retrieve, and repack requests, relying on a Scheduler database (Scheduler DB) for transient metadata storage. We present the development of a new relational database (PostgreSQL) backend for the Scheduler DB. The aim is to improve the limitations of the current (object-store based) implementation. This talk will provide a status update on the implementation of archive and retrieve workflows, discuss key advantages of the relational approach, and outline the roadmap.

CTA Development and Roadmap / 4

Migrating CTA Frontend from XRootD/SSI to gRPC

Author: Konstantina Skovola[None]

Corresponding Author: konstantina.skovola@cern.ch

The CTA Frontend serves the physics workflow event requests made by the disk system buffer, for example EOS or dCache, and cta-admin commands submitted by operators and automated scripts. Communication with the frontend is currently based on XRootD/SSI. However, not all disk front-ends support the SSI extensions to XRootD, which creates a constraint for sites using CTA as a tape backend. As a result, we have decided to migrate CTA to a more widely-supported protocol.
This talk will present the ongoing work on migrating from XRootD/SSI to the industry-standard gRPC framework. We will present the implementation of the rpc methods, discuss how authentication works with gRPC vs XRootD/SSI and finally discuss the roadmap for completely switching to gRPC.

CTA Operations and Site Reports / 5

Fermilab site report and Enstore compatibility

Authors: Dmitry Litvintsev[1]; Eric Vaandering[1]; Jeff Derbyshire[None]; Lucas Peter Trestka[1]; Rafael Arturo Rocha Vidaurri[None]; Scarlet Norberg[None]; Terrance Jones[2]; Tim Messer[1]; Victor Leopoldo Munoz Flores[1]

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

[2] *FL/CD/SCD/SDS/SSA*

Corresponding Authors: timothy.alan.messer@cern.ch, norberg@ou.edu, jeffderb@fnal.gov, ltrestka@fnal.gov, rrochavi@fnal.gov, jonest@fnal.gov, ewv@fnal.gov, litvinse@fnal.gov, victormf@fnal.gov

Fermilab will move to CTA this spring with dCache as the frontend file system. The modifications made to CTA to be able to read Enstore (Fermilab's legacy tape management software) files will be discussed, as will our solution to read the existing Enstore Small File Aggregation (SFA) files.

Operational issues arising during our push to production will be highlighted. Details on our deployment, infrastructure and monitoring will be included. Finally, our remaining steps in fully migrating the Enstore installations to CTA will be covered.

CTA Operations and Site Reports / 6

CTA status at IHEP

Authors: Yujiang BI[1]; qiuling yao[2]

Co-authors: LI Haibo lihaibo ; Yaodong Cheng [3]

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

[2] *IHEP*

[3] *IHEP, CAS*

Corresponding Authors: lihaibo@ihep.ac.cn, biyujiang@ihep.ac.cn, yaoql@ihep.ac.cn, yaodong.cheng@cern.ch

We will share our experiences and challenges with using CTA last year. We optimized the CTA configuration and upgraded the EOS&CTA. Additionally, we expanded the scale of two experimental applications, Tier 1 of LHCB and HEPS, and enhanced the monitoring of the CTA system.

Plenary Session / 7

Welcome and Introduction

Author: Michael Davis[1]

[1] *CERN*

Corresponding Author: michael.davis@cern.ch

Welcome to the fourth annual CERN Tape Archive Workshop (CTA 2025).

The LHC experiments have completed another successful year of data taking, with new records in terms of throughput and total data archived. This presentation introduces the CTA Project, Team and Community, as well as an overview of the challenges and achievements during the third year of LHC Run-3.

CTA Development and Roadmap / **8**

CTA Roadmap

Author: Michael Davis[1]

[1] *CERN*

Corresponding Author: michael.davis@cern.ch

CTA software development is primarily driven by the needs of the CERN experimental programme. Looking beyond Run-3, data rates are set to continue to rise exponentially into Run-4 and beyond. The CTA team are planning how to scale the software and service to meet these new challenges.

CTA is also driven by the needs of the community outside CERN. The landscape of tape archival for scientific data is consolidating, and CTA is constantly adapting to a wider range of use cases.

This talk will present the short-term and medium-term roadmap for CTA development and new features.

CTA Development and Roadmap / **9**

Discussion and close-out

Author: Michael Davis[1]

[1] *CERN*

Corresponding Author: michael.davis@cern.ch

Final comments, questions and discussion.

Plenary Session / **10**

Cold storage using S3+GLACIER+CTA

Author: Michael Davis[1]

[1] *CERN*

Corresponding Author: michael.davis@cern.ch

S3 is a popular protocol for object storage, in use at CERN since 2018. CERN's Ceph S3 service provides a disk storage back-end for various use cases including backup. In principle it is possible to archive S3 objects to tape using the S3 GLACIER storage class extensions. However, this is not yet supported in an open source solution. This talk gives a brief overview of the landscape of support for S3+tape and describes CERN's planned proof-of-concept for an open source S3+tape implementation using CTA.

Plenary Session / **11**

CERN Perspective on Tape Technology Evolution

Author: Vladimir Bahyl[1]

[1] *CERN*

Corresponding Author: vladimir.bahyl@cern.ch

With LHC now in the middle of Run 3, we will present our current tape hardware setup and present our experience with the different components of the technology.
We will start with a reflection on the evolution of our capacity planning vs. increasing storage requirements of the experiments.
We will then report on performance characteristics of both LTO9 and TS1170 tape drives: RAO, environmental aspects and how the technology evolution is impacting our operations.
Lastly, we will share our thoughts about rack size scale-out tape libraries and consideration to replace FC with SAS.

CTA Operations and Site Reports / **12**

Offsite Tape Backup Collaboration

Author: Vladimir Bahyl[1]

[1] *CERN*

Corresponding Author: vladimir.bahyl@cern.ch

This talk is a follow-up from the 2024 BoF session on Offsite Tape Backup between sites. We will present the proof-of-concept architecture that we plan to develop in 2025. We propose to test it with one collaborating Tier-1 site (yet to be identified).

CTA Operations and Site Reports / **13**

Insights from Operating CTA at CERN

Author: Pablo Oliver Cortes[1]

[1] *CERN*

Corresponding Author: pablo.oliver.cortes@cern.ch

Operating the CERN Tape Archive all year-round does not come without surprises and challenges: massive recall campaigns, peak system throughput for archival during the data tacking period and (not so) transparent upgrades to critical services we depend on push the system to the limits, popping some nuts and bolts from time to time.

In this presentation, we will share insights gained from challenges such as temporary network disconnects and hardware problems. We will also discuss the procedures we have developed to mitigate these issues while working towards long term solutions.

CTA Development and Roadmap / **14**

CTA Tape Daemon Evolution

Author: Pablo Oliver Cortes[1]

[1] *CERN*

Corresponding Author: pablo.oliver.cortes@cern.ch

In 2024, the CTA Tape Daemon was updated to address issues in deployments with multiple drives per tape server. This was a first step towards a major refactoring of the daemon, as in its current state, its multi-process architecture presents problems such as logging information unrelated to the current process and inter-process communication bugs. It also causes confusion in internal discussions and in external bug reports.

In this session, we will provide an update on the current state of the CTA Tape Daemon, explore the upcoming changes designed to reduce the daemon's complexity by removing its parent-child process architecture, and promoting the maintenance process to an independent daemon, as it doesn't perform any drive related tasks. We will conclude with a discussion of possible future developments.

Hands-on Session / **15**

Evolution of Continuous Integration for CTA

Author: Niels Alexander Buegel[None]

Corresponding Author: niels.alexander.bugel@cern.ch

CTA's Continuous Integration (CI) system has been around since the inception of the project. However, numerous limitations had piled up: the CI setup was monolithic, developments outside of CTA were difficult to test (EOS, XRootD and dCache), the pipelines were slow and large parts of the CI system were not nicely structured. Over the past year, the CTA team has made significant improvements to the CI in the areas of simplicity, flexibility, performance and robustness. The most impactful change was the migration from plain Kubernetes manifest files to Helm, allowing us to decouple the configuration of CTA from the EOS disk system configuration and opening up opportunities to test other disk buffer systems such as dCache. The new setup allows us to handle complex testing scenarios and perform regression testing on various components independently. We will discuss the challenges we encountered with our CI, the improvements we implemented to address them, and what we hope to do in the future.

Hands-on Session / **16**

Deploying a CTA Test Instance using Helm

Author: Niels Alexander Buegel[None]

Corresponding Author: niels.alexander.bugel@cern.ch

The current CTA test instance is deployed using Helm, a tool that streamlines the installation and management of Kubernetes applications. It allows us to clearly separate and template the different components that make up CTA. In this talk, we will walk through how the new containerized CTA setup works, covering how the components are organized, how Helm is used to manage configurations, and the deployment process.

Hands-on Session / **17**

Upgrading the CTA Catalogue

Author: Niels Alexander Buegel[None]

Co-author: Julien Leduc [1]

[1] *CERN*

Corresponding Authors: niels.alexander.bugel@cern.ch, julien.leduc@cern.ch

In this hands-on, participants will deploy their own containerized CTA + EOS setup and perform a catalogue upgrade from version 14.0 to 15.0. The goal is to walk through the full upgrade procedure so that participants understand the steps necessary to smoothly do a catalogue upgrade of CTA. Participants are expected to bring their own laptops and will be provided with an OpenStack VM.

Hands-on Session / **18**

Streamlining CTA version upgrades

Author: Joao Afonso[1]

[1] *CERN*

Corresponding Author: joao.afonso@cern.ch

The CTA software versioning numbering scheme distinguishes between standard code releases and “pivot”releases aimed at upgrading the CTA Catalogue schema. This separation provides us with a simple and replicable set of steps for upgrading CTA between any two versions.
This talk will present the strategy for versioning the CTA software, explain how CTA Catalogue schema upgrades are integrated into our release cycle, and how it all ties together with the planning of new RPM releases. It’s goal is to provide users with the knowledge necessary to navigate between different CTA versions, in particular during CTA Catalogue upgrades.

CTA Operations and Site Reports / **19**

Configuring ATRESYS to use Repack tapepools

Author: Idriss Larbi[None]

Corresponding Author: idriss.larbi@cern.ch

Until now, repack and archival of new files were carried out on the same tapepool. This could lead to the user's new archive files being mixed on the same tape as the old repack files, reducing performance during retrieve. With the latest version of CTA, we can create dedicated REPACK archive routes in order to repack to separate tapepools. We have modified ATRESYS to accommodate this new repacking method. This presentation gives a brief explanation of the new actions implemented in ATRESYS. A quick demonstration of a REPACK with ATRESYS will be presented to visualise the workflow.

CTA Operations and Site Reports / 20

Developments in the Antares service at RAL

Author: George Patargias[None]

Corresponding Author: george.patargias@stfc.ac.uk

Antares is the tape archive service at RAL that manages both Tier-1 and local Facilities data. In this talk, we present the main operational changes and developments in the service since last year's CTA workshop. Among others, these include the migration of the service from SL7, the CTA Frontend separation and the deployment of the new EOS nodes connected to the LHC-OPN network in our Tier-1 EOS cluster. In conjunction with the Facilities migration to TS1170 tape media, we present some work done on data reorganization with the aim to achieve higher recall efficiency. Finally, we discuss various ongoing operational problems and conclude with the development plans for the immediate future.

CTA Operations and Site Reports / 21

CTA's Retrieve Backpressure Mechanism: Status and Future

Author: Pablo Oliver Cortes[1]

[1] *CERN*

Corresponding Author: pablo.oliver.cortes@cern.ch

CTA was designed with two goals in mind: throughput to and from the tape system and minimising the stress on the tape infrastructure (minimising the number of tape mounts). These two constraints become particularly challenging in retrieval dataflows when elements external to the system start to misbehave.

In this presentation, we will explore the internal logic behind CTA's retrieval workflow, the mechanisms that come into action when the system starts to experience problems, and the various approaches we have taken to address these challenges.

CTA Operations and Site Reports / 22

Tape system at TRIUMF, recent updates, CTA evaluation for site

Author: Simon Liu[1]

Co-authors: Reda Tafirout [1]; Yun-Ha Shin

[1] *TRIUMF (CA)*

Corresponding Authors: simonliu@triumf.ca, yshin@triumf.ca, tafirout@triumf.ca

The Tapeguy is TRIUMF's home build tape system for ATLAS T1 data center, It was designed to be a stable system that can reliably store and retrieve LHC produced data as a tiered HSM system. we also open to other solutions, evaluated CERN CTA at site in 2024. the talk will present current tapeguy status, recent updates, and the evaluation done at site.

Plenary Session / **23**

Scaleway S3 Glacier: tape archiving for the masses

Authors: Florian FLORENSA[1]; Jules Hezard[1]; Saalik HATIA[1]

[1] *Scaleway*

Corresponding Authors: fflorensa@scaleway.com, shatia@scaleway.com, jhezard@scaleway.com

For six years Scaleway offered S3 Glacier-class storage fully operated on mostly powered-off SMR disks.
This enabled us to offer our customers fast restore time while remaining cost effective, using a mix of commodity and custom hardware.

As our S3 service grew, and since we cannot control nor predict public cloud workloads, the former glacier stack was unable to keep up while becoming more difficult to maintain and operate.
Storing on tape presents its own set of challenges, optimising a disk-oriented workflow to suit a tape-based backend, while being unable to predict the immediate needs of our clients.

This presentation will talk about lessons learned and what the Scaleway Glacier future looks like.

CTA Development and Roadmap / **24**

Optimising recall from tape using Archive Metadata

Author: Julien Leduc[1]

[1] *CERN*

Corresponding Author: julien.leduc@cern.ch

During Run-3, CTA has demonstrated very high write efficiency at nominal DAQ rates. For retrieval, CTA relies on time-based colocation of data on tape, but this has proved to be much less efficient than expected. Furthermore, the ratio of tape reads to writes is expected to significantly increase during Run-4, as some LHC experiments move towards the "tape carousel"model. Two years ago, we started to analyse the constraints from the experiments and tape sites, in order to devise a way to improve retrieve efficiency. The resulting Archive Metadata proposal is a practical abstraction layer between experiment data management and tape storage endpoints. It allows to express how group of files are likely be staged together.
This talk will present how this project has evolved over the past two years, its initial proposal implementation and finally an overview of the collected metadata during the 2024 Heavy Ion run.
Finally, we will also give an overview of the plan for upcoming research and development using Archive Metadata.

CTA Operations and Site Reports / 25

Hardware provisioning for CTA

Author: Julien Leduc[1]

[1] *CERN*

Corresponding Author: julien.leduc@cern.ch

The CERN Tape Archive (CTA) was designed to meet the demands of data archival from the LHC experiments, in terms of both data volume and throughput. In order to ingest data at the rates demanded by the LHC data acquisition (DAQ) systems, the system is built on EOS and CTA's scalable architecture principles. To optimise the performance of both disk and tape hardware and to achieve the desired I/O rates, the hardware must be provisioned appropriately. This talk will present a high-level view of the global CERN tape service configuration and dive into the details of how the required performance is achieved. We start from benchmarks of the SSDs used as the building blocks for the tape buffer and work up to the full service configuration, based on commodity hardware as used at CERN/WLCG Tier-0. This talk aims to outline the factors that a Tier-1 site who wants to deploy CTA should take into consideration.

CTA Development and Roadmap / 26

Tape reading efficiency gain by collocation

Author: Luc Goossens[1]

[1] *CERN*

Corresponding Author: luc.goossens@cern.ch

Tape reading efficiency, defined as the ratio between the effective average data reading rate and the maximal data reading rate, is reduced by two operations the tape drive inevitably needs to do and during which it can not read any data. The first is mounting the tape containing the file into the drive, after possibly having unmounted the tape that was in it before. The second is spooling the tape to the position of the file to be read. In this presentation we will explore which factors determine this reading efficiency loss and what could be done to mitigate it. We will in particular look at collocation, i.e. the strategy of writing files that are likely to be recalled together without interleaving them with unrelated files. To validate our hypotheses we analysed one month (Feb 2025) of ATLAS tape recall activity totalling about 6 PB of data contained in 1.6 million files spread over 1800 datasets.