

CERN Tape Archive Workshop : CTA 2024

Monday, 18 March 2024 - Tuesday, 19 March 2024

CERN

Book of Abstracts

Contents

CTA Operator Tools News	1
The CTA Tape Alerting System	1
Archive metadata	1
CTA operations best practices	2
CTA status at IHEP	2
Container-based CTA installation for storing midsize archive data	2
PIC's journey to CTA	3
CTA site report of AARNet	3
Experiences migrating CTA to Alma Linux 9	3
CTA evaluation at JINR	4
From stagerrm to evict : disk buffer management in CTA	4
Improving CTA performance by splitting user and repack scheduler backends	4
CTA status at RAL	5
Welcome and Introduction	5
CTA Outlook and Roadmap	5
Discussion and Close-out	6
A new simple and flexible configuration of tape drives and the CTA tape daemon	6
Understanding and debugging common CTA problems	6
One Year of Running CTA at DESY	7
Keynote: The Future of Tape	7
CTA site report for Fermi National Accelerator Lab (FNAL)	7
BoF : Backing up data between WLCG sites	8
Restoring deleted tape files in EOSCTA	8

CTA Operations and Site Reports / 1**CTA Operator Tools News****Author:** Richard Bachmann¹¹ CERN**Corresponding Author:** richard.bachmann@cern.ch

At the 2nd CTA Workshop (2023), we announced a repository of Free and Open Source operator tools for CTA, including the Repack automation system (ATRESYS) tool. Since then, the repository has seen a number of new tool additions, as well as changes to the underlying libraries.

In this talk, we will introduce these new features, including tools for: automating the supply of tapes to tape pools; automated validation of data on tape; EOSCTA disk file metadata operations; and higher-level CTA operator tasks. Together, these make up the backbone of the tape operations ecosystem at CERN. We present these contributions in the hope that they will be as useful to the CTA community at large.

CTA Operations and Site Reports / 2**The CTA Tape Alerting System****Author:** Richard Bachmann¹¹ CERN**Corresponding Author:** richard.bachmann@cern.ch

The Tape Alerting System (TAS) acts automatically on issues detected in the tape infrastructure by disabling the affected elements, notifying operators of the situation requiring their attention, and thereby protecting the system from further disruption or damage to the infrastructure.

This is done by performing a scan of a configurable time frame of past tape sessions and then executing a number of detection jobs corresponding to the set of configurable error conditions.

This presentation will give an introduction to the TAS, walk you through the potential error conditions this operator tool reacts to, and give instructions on how to install it.

CTA Development and Roadmap / 3**Archive metadata****Author:** Julien Leduc¹¹ CERN**Corresponding Author:** julien.leduc@cern.ch

The CTA software and service have been designed to match the Run-3 write performance requirements at WLCG Tier-0. The queuing system, coupled with a small SSD-based cache, has demonstrated its performance predictability over the first two years of the run.

This performance was achieved with FIFO scheduling, resulting in a pure temporal collocation on tapes, with only the legacy “storage class” concept to separate data into different tape pools.

As tape data is becoming “warmer” in experiment data workflows, the performance of data retrieval from tape is becoming more important. Optimising retrieval requires additional metadata to deliver staging efficiency gains, which can be exploited in the evolution of the CTA tape scheduler.

This presentation will focus on various scheduling issues observed during Run-3, and how archive metadata will enable improvements to the CTA tape site write and read efficiencies.

CTA Operations and Site Reports / 4

CTA operations best practices

Author: Julien Leduc¹

¹ CERN

Corresponding Authors: pablo.oliver.cortes@cern.ch, julien.leduc@cern.ch

During three years in production at CERN as the WLCG Tier-0 tape storage service, the CTA software has evolved and gained various features driven by the service requirements. Discussions between developers and the operations have led to design choices that shaped the service and helped refine CTA operations best practices.

This presentations aims at clarifying the current operations and user data transfer best practices when dealing with a CTA tape system.

CTA Operations and Site Reports / 5

CTA status at IHEP

Authors: LI Haibo lihaibo^{None}; Yaodong CHENG¹; Yaosong Cheng²; Yujiang BI³; qiuling yao⁴

¹ IHEP, Beijing

² Institute of High Energy Physics Chinese Academy of Sciences, IHEP

³ Institute of High Energy Physics, Chinese Academy of Sciences

⁴ IHEP

Corresponding Authors: chengys@ihep.ac.cn, lihaibo@ihep.ac.cn, biyujiang@ihep.ac.cn, ycheng@cern.ch, yaoql@ihep.ac.cn

The presentation will provide an overview of the status and progress of CTA at IHEP. Last year, we completed the data migration from CASTOR to CTA, added new CTA instances for JUNO, HEPS, and LHCb Tier1. We compiled CTA 5.8.10 on alma9 and have commenced performance testing. Additionally, CTA has been utilized on the new tape library and LTO9 tapes.

CTA Operations and Site Reports / 6

Container-based CTA installation for storing midsize archive data

Author: Viktor Kotliar¹

Co-authors: Denis Syrko ²; Dmitry Bolotin

¹ National Research Centre Kurchatov Institute (RU)

² *Pirogov Russian National Research Medical University*

Corresponding Authors: victor.kotlyar@cern.ch, bolotin.dmitriy@gmail.com, denissyenko@ya.ru

This paper presents one of the possible architectures for a container-based CTA installation designed for a midsize storage system. The primary objective of this setup is to store archive data efficiently and integrate it with an existing computing cluster.

CTA Operations and Site Reports / 7

PIC's journey to CTA

Authors: Elena Planas¹; Elisabet Carrasco Santos²; Esther Accion Garcia³; Jordi Casals Hernandez⁴; Vanessa Acin Not Supplied⁵

¹ *PIC*

² *PIC-IFAE*

³ *The Barcelona Institute of Science and Technology (BIST) (ES)*

⁴ *Port d'Informació Científica*

⁵ *Institut de Física d'Altes Energies*

Corresponding Authors: elisabet.carrasco.santos@cern.ch, vacin@pic.es, esthera@pic.es, jcasals@pic.es, eplanas@pic.es

Throughout 2023, we have continued conducting operational tests to familiarize ourselves with the functionality of CTA and understand the differences from Enstore. In addition to configuration tests of the application to observe how resource allocation works or to configure automatic tape provisioning, efforts have been made to deploy a centralized logging system to enhance the monitoring and operation of the application. Furthermore, preparations have begun for the migration tests from Enstore to CTA.

CTA Operations and Site Reports / 8

CTA site report of AARNet

Author: Denis Lujanski^{None}

Corresponding Author: dmaybir@hotmail.com

This presentation will aim to provide a yearly site report. This year at AARNet, we implemented tape limiting and offsite tape vaulting processes and workflows. CTA was also an integral part of decommissioning efforts of some of our storage platforms.

CTA Development and Roadmap / 9

Experiences migrating CTA to Alma Linux 9

Author: Jorge Camarero Vera¹

¹ *CERN*

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

Due to the upcoming end-of-life of CERN CentOS 7, CTA will be migrated to Alma Linux 9, following the recommendations of the CERN Linux team. Migrating the CTA codebase from CC7 to Alma 9 presented a range of compatibility hurdles. This talk will delve into the challenges encountered and the strategies used to overcome them, including: managing version changes in vital dependencies like Protobuf and Oracle client libraries; resolving differences between CC7 and Alma9 package offerings; and ensuring C++ code compatibility. Additionally, we will address the transition from Docker to Podman, the adaptation of the versionlock file for Alma 9, modernisation of scripts from Python 2 to Python 3, handling potential Bash version incompatibilities, and strategies for integrating the CTA docker image into a local minikube environment.

CTA Operations and Site Reports / 10

CTA evaluation at JINR

Author: Nikita Balashov¹

¹ *Joint Institute for Nuclear Research (RU)*

Corresponding Author: nikita.balashov@cern.ch

This presentation summarizes the current state of tape storage systems at the Joint Institute for Nuclear Research (JINR) and gives our current vision of its future in response to the discontinuation of Enstore support. We provide a brief description of two of our currently operating tape systems: a 90 PB instance built on dCache/Enstore and an 11 PB instance based by EOS/CTA. Additionally, we share initial insights gained from operating the EOS/CTA system.

CTA Development and Roadmap / 11

From stagerrm to evict : disk buffer management in CTA

Author: Joao Afonso¹

¹ *CERN*

Corresponding Author: joao.afonso@cern.ch

The original design of EOSCTA inherited its cache eviction mechanism —stagerrm— from CASTOR. Although similar, CTA's use-cases were not the same as CASTOR's, leading to several operational issues. This motivated the creation of a new command —evict— better adapted to modern disk buffer management in CTA.

In this talk, we will discuss the issues that we faced with stagerrm and how the new evict command helps to fix them. We will also cover the new features brought by evict and how they simplify EOSCTA operations.

CTA Development and Roadmap / 12

Improving CTA performance by splitting user and repack scheduler backends

Author: Joao Afonso¹

¹ CERN

Corresponding Author: joao.afonso@cern.ch

Recent improvements to CTA repack added several new features and tools for operators. However, we were still faced with severe performance issues when repacking tapes on a very large scale. An investigation showed that this was mostly due to limitations on the CTA SchedulerDB backend, which did not scale well to performing repack on the latest generation of very high capacity tapes, which can store 50 TB of data and millions of files. As a mitigation, while a new SchedulerDB is still in development, we decided to split the “user” and “repack” scheduler backends. This will effectively prevent repack jobs from interfering with user archival jobs. In this talk, we will discuss the limitations found during repack, the mitigations put in place, and why separating the scheduler into two backends was necessary to avoid performance issues during CTA operations.

CTA Operations and Site Reports / 13

CTA status at RAL

Author: George Patargias^{None}

Co-authors: Thomas Byrne ; Mahalakshmi Agilandamurthy ; Tim Folkes ; Alison Packer ; Alastair Dewhurst

¹

¹ Science and Technology Facilities Council STFC (GB)

Corresponding Authors: tom.byrne@stfc.ac.uk, george.patargias@stfc.ac.uk, alastair.dewhurst@cern.ch, tim.folkes@stfc.ac.uk, alison.packer@stfc.ac.uk, mahalakshmi.agilandamurthy@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 developments in the service since last year’s CTA workshop including the migration of CASTOR Facilities instance and discuss the service’s performance as well as the main operational issues since the beginning of LHC Run-3. Finally, we provide an overview of the future plans for the expansion of the service to cover additional use cases from physics (CLF) and astronomy (SKA).

Opening Plenary Session / 14

Welcome and Introduction

Author: Michael Davis¹

¹ CERN

Corresponding Author: michael.davis@cern.ch

Welcome to the third annual CERN Tape Archive Workshop (CTA 2024).

The LHC experiments have completed another successful year of data taking, with new records in terms of throughput and total data archived.

This presentation will give an introduction to the CTA Project, Team and Community, as well as an overview of the challenges and achievements during the second year of LHC Run-3.

CTA Development and Roadmap / 15**CTA Outlook and Roadmap**

Author: Michael Davis¹

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 / 16**Discussion and Close-out**

Author: Michael Davis¹

¹ CERN

Corresponding Author: michael.davis@cern.ch

Final comments, questions and discussion.

CTA Operations and Site Reports / 17**A new simple and flexible configuration of tape drives and the CTA tape daemon**

Authors: Pablo Oliver Cortes¹; Vladimir Bahyl¹

¹ CERN

Corresponding Authors: vladimir.bahyl@cern.ch, pablo.oliver.cortes@cern.ch

The configuration of the CTA tape daemon was inherited from CASTOR, which relied on a manual process to create and store the configuration for each tape drive. As the scale of CERN's tape operations have grown to hundreds of drives, this approach has not scaled well. The CTA team have therefore developed a new semi-automated process to determine the configuration of each drive. The configuration has been simplified and refactored, making it easier to operate multiple tape drives per tape server. In this two-part presentation we will describe: the procedure to automatically obtain the drive information from the library; and the new tape daemon configuration file structure and how to configure multiple drives per tape server.

Hands-on Sessions / 18

Understanding and debugging common CTA problems

Author: Pablo Oliver Cortes¹

¹ *CERN*

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

The goal of this hands-on session is to be able to debug common problems that could be faced by CTA administrators or operators. We will briefly give a broad overview of all the logging generated by the different CTA services, along with a peek into our monitoring dashboards. This gives a starting point to be able to understand and debug problems that may appear while operating CTA.

After that, a series of problems will be presented to be solved during the session.

CTA Operations and Site Reports / 19

One Year of Running CTA at DESY

Author: Mwai Karimi^{None}

Corresponding Author: mwai.karimi@desy.de

CTA has been in production at DESY for almost an year, with dCache as the frontend. Over the course of this time we have migrated all our data from OSM, moved all experiments to CTA and so far written close to 25PB of data. In this talk we give an overview of the past year and share our overall experience with CTA.

Opening Plenary Session / 20

Keynote: The Future of Tape

Author: Mark Lantz¹

¹ *IBM*

Corresponding Author: mla@zurich.ibm.com

Today, magnetic tape is the lowest cost technology for storing large volumes of data. Historically, areal density scaling has been the main driver of the exponential decrease in cost, i.e. cents/GB of capacity, of both tape and hard disk drives. Recently, HDD scaling has stagnated while tape continues to scale at historical rates. The INSIC 2019-2029 Tape Technology Roadmap projects that tape areal density will scale with a 40% compound annual growth rate enabling an approximate doubling of capacity every two years until at least 2029, at which point tape systems are expected to reach an areal density of 278 Gb/in². The feasibility of these projections have been validated by a recent research single channel tape areal density demonstration of 317 Gb/in² on SrFe tape. This demo as well as the latest tape technology roadmap will be discussed and contrasted with hard disk drives.

CTA Operations and Site Reports / 21

CTA site report for Fermi National Accelerator Lab (FNAL)

Corresponding Author: ewv@fnal.gov

In 2024, Fermilab will replace the legacy Enstore tape management system with CTA for CMS data. This will be followed by the migration of the second instance of Enstore containing all other scientific data on tape at Fermilab.

We will detail the results of on-going scale tests of CTA. These tests target the 10% scale in tape bandwidth compared to the existing CMS Enstore instance. Tests will be performed with both EOS and dCache as buffer file systems.

We will also cover code development efforts necessary to make all tapes written by Enstore readable in CTA.

22

BoF : Backing up data between WLCG sites

Corresponding Author: vladimir.bahyl@cern.ch

Recent years have witnessed an increase in the number and sophistication of cyberattacks on academic sites. One way to improve resilience is to store backups in a completely separate administrative domain to the data being protected. This Birds-of-a-Feather session is for site administrators who would like to discuss the possibility of backing up to each others sites.

Hands-on Sessions / 23

Restoring deleted tape files in EOSCTA

Corresponding Author: julien.leduc@cern.ch

Deletions are a rare event in tape archive storage operations. When they happen, they are often the result of a user mistake or a bug. As deletion is asynchronous on the tape side, nothing is physically deleted until the tape is reclaimed. CTA keeps a copy of all the deleted file metadata so that files deleted unintentionally can be restored. This hands-on session will allow participants to understand what happens to deleted tape files in the CTA tape lifecycle and how to use the CTA restore tools.