

Contribution ID: 667

Type: Poster

## Data Processing and Preservation System for the Cherenkov Telescope Array Observatory

The Cherenkov Telescope Array Observatory (CTAO) will be the world's leading facility for very-high-energy gamma-ray astrophysics, producing unprecedented volumes of raw and processed data. Efficient processing, storage, and long-term preservation of these data are critical for ensuring scientific reproducibility and accessibility. The Data Processing and Preservation System (DPPS) is designed to meet these challenges through a distributed, automated, and scalable infrastructure.

DPPS processes observational and simulated data through automated workflows executed across multiple CTAO data centers. It employs a modular pipeline framework integrated with a workload management system to handle complex processing chains, ensuring efficient transformation of raw telescope data into high-level science products. The system is built on modern, open-source technologies, including DIRAC middleware for workload management with workflow automation based on the Common Workflow Language (CWL), Rucio for data management, and CERN's File Transfer Service (FTS) for efficient data movement. Additionally, containerized software environments and Continuous Integration/Continuous Deployment (CI/CD) pipelines ensure reproducibility and maintainability.

In this contribution, we present the architecture, status, and future development plans of DPPS. We discuss the technical choices behind workflow automation, distributed computing, and data preservation, as well as the challenges associated with processing and managing petabyte-scale data efficiently.

## Collaboration(s)

CTAO

Author: Dr DALCHENKO, Mykhailo (University of Geneva)
Presenter: Dr DALCHENKO, Mykhailo (University of Geneva)
Session Classification: PO-2

Track Classification: Gamma-Ray Astrophysics