

A new PostgreSQL backend for CERN Tape Archive scheduling for LHC Run 4

Friday, February 6, 2026 10:05 AM (30 minutes)

The CERN Tape Archive (CTA) stores over one exabyte of scientific data. To orchestrate storage operations (archival) and access operations (retrieval), the CTA Scheduler coordinates concurrent data movements across hundreds of tape servers, relying on a Scheduler Database (Scheduler DB) to manage the metadata of the in-flight requests. The existing objectstore-based design of the CTA Scheduler DB is a complex transactional management system. This talk presents the development of a new PostgreSQL-based backend for the CTA Scheduler as an off-the-shelf solution which simplifies implementation and is expected to significantly reduce future development and operational costs. We describe the implementation of all main CTA workflows and explain how PostgreSQL addresses the limitations of the objectstore-based system, providing the foundation for the tenfold increase in data throughput expected during LHC Run 4.

Author: Dr GUENTHER, Jaroslav (CERN)

Co-author: SKOVOLA, Konstantina

Presenters: Dr GUENTHER, Jaroslav (CERN); SKOVOLA, Konstantina

Track Classification: Day 1