21st International Conference on Computing in High Energy and Nuclear Physics **CHEP2015** Okinawa Japan: April 13 - 17, 2015

Contribution ID: **298** Type: **oral presentation**

Archiving tools for EOS

Tuesday, 14 April 2015 18:00 (15 minutes)

Archiving data to tape is a critical operation for any storage system, especially for the EOS system at CERN which holds production data from all major LHC experiments. Each collaboration has an allocated quota it can use at any given time therefore, a mechanism for archiving "stale" data is needed so that storage space is reclaimed for online analysis operations.

The archiving tool that we propose for EOS aims to provide a robust interface for moving data between EOS and the tape storage system while enforcing data integrity and verification.

The archiving infrastructure is written in Python and is fully based on the XRootD framework. All data transfers are done using the third-party copy mechanism which ensures point-to-point communication between the source and destination, thus providing maximum aggregate throughput. Using ZMQ message-passing paradigm and a process-based approach enabled us to archive optimal utilisation of the resources and a stateless architecture which can easily be tuned during operation.

In conclusion, we make a comparative analysis between archiving a data set in a "managed" way using the archiving tool and the old plain copy method, highlighting the speed-up gain, the data integrity checks performed and the behaviour of the system in different failure scenarios. We expect this tool to considerably improve the data movement work-flow at CERN in both directions between disk and tape.

Primary author: SINDRILARU, Elvin Alin (CERN)

Co-authors: Mr PETERS, Andreas Joachim (CERN); DUELLMANN, Dirk (CERN)

Presenter: Mr PETERS, Andreas Joachim (CERN)

Session Classification: Track 3 Session

Track Classification: Track3: Data store and access