

LHC Data Storage: Preparing for the Challenges of Run-3

Wednesday, 19 May 2021 11:16 (13 minutes)

The CERN IT Storage Group ensures the symbiotic development and operations of storage and data transfer services for all CERN physics data, in particular the data generated by the four LHC experiments (ALICE, ATLAS, CMS and LHCb).

In order to accomplish the objectives of the next run of the LHC (Run-3), the Storage Group has undertaken a thorough analysis of the experiments' requirements, matching them to the appropriate storage and data transfer solutions, and undergoing a rigorous programme of testing to identify and solve any issues before the start of Run-3.

In this paper, we present the main challenges presented by each of the four LHC experiments. We describe their workflows, in particular how they communicate with and use the key components provided by the Storage Group: the EOS disk storage system; its archival back-end, the CERN Tape Archive (CTA); and the File Transfer Service (FTS). We also describe the validation and commissioning tests that have been undertaken and challenges overcome: the ATLAS stress tests to push their DAQ system to its limits; the CMS migration from PhEDEx to Rucio, followed by large-scale tests between EOS and CTA with the new FTS "archive monitoring" feature; the LHCb Tier-0 to Tier-1 staging tests and XRootD Third Party Copy (TPC) validation; and the erasure coding performance in ALICE.

Primary authors: Dr ARSUAGA RIOS, Maria (CERN); BAHYL, Vladimir (CERN); BATALHA, Manuel; CAFFY, Cedric (CERN); CANO, Eric (CERN); CAPITONI, Niccolo (CERN); CONTESCU, Cristian (CERN); DAVIS, Michael (CERN); FERNANDEZ ALVAREZ, David (CERN); GUENTHER, Jaroslav (CERN); KARAVAKIS, Edward (CERN); KEEBLE, Oliver (CERN); LEDUC, Julien (CERN); LUCHETTI, Fabio (CERN); MASCETTI, Luca (CERN); MURRAY, Steven (CERN); PATRASCOIU, Mihai (CERN); PETERS, Andreas Joachim (CERN); SIMON, Michal Kamil (CERN); SINDRILARU, Elvin Alin (CERN); TOEBBICKE, Rainer (CERN)

Presenter: Dr ARSUAGA RIOS, Maria (CERN)

Session Classification: Storage

Track Classification: Distributed Computing, Data Management and Facilities