



Contribution ID: 251

Type: Poster

Consistency between Grid Storage Elements and File Catalogs for the LHCb experiment's data

Tuesday, 22 May 2012 13:30 (4h 45m)

In the distributed computing model of WLCG Grid Storage Elements (SE) are by construction completely decoupled from the File Catalogs (FC) where the experiment's files are registered. On the basis of the experience of managing large volumes of data in such environment, inconsistencies have often happened either causing a waste of disk space, in case the data were deleted from the FC, but still physically on the SE, or serious operational problems in the opposite case, when some data registered in the FC was not found on the SE. Therefore, the LHCbDirac data management system has been equipped with a new dedicated system to ensure the consistency of the data stored on the SEs with the information reported in the FCs implementing systematic checks. Objective of the checks is to spot any inconsistency above a certain threshold, that cannot only be due to the expected latency between data upload and registration, and in such case try and identify the problematic data. The system relies on information provided by the sites who should make available to the experiment a full dump of their SEs on weekly or monthly basis.

In this talk we shall present the definition of a common format and procedure to produce the storage dumps that has been coordinated with the other LHC experiments in order to provide a solution as generic as possible that can suit all LHC experiments and will reduce the effort for the sites who are asked to provide such data. We will also present the LHCb specific implementation for checking the consistency between SEs and FC and discuss the results.

Primary author: LANCIOTTI, Elisa (CERN)

Co-authors: ZHELEZOV, Alexey (Ruprecht-Karls-Universitaet Heidelberg (DE)); Dr TSAREGORODTSEV, Andrei (Universite d'Aix - Marseille II (FR)); SERFON, Cedric (Ludwig-Maximilians-Univ. Muenchen (DE)); REMENSKA, Daniela (NIKHEF (NL)); Dr BOUVET, David (Universite Claude Bernard-Lyon I (FR)); STAGNI, Federico (CERN); CLOSIER, Joel (CERN); CATTANEO, Marco (CERN); UBEDA GARCIA, Mario (CERN); RATNIKOVA, Natalia (KIT - Karlsruhe Institute of Technology (DE)); MAGINI, Nicolo (CERN); CLARKE, Peter (University of Edinburgh (GB)); CHARPENTIER, Philippe (CERN); NANDAKUMAR, Raja (STFC - Science & Technology Facilities Council (GB)); GRACIANI DIAZ, Ricardo (University of Barcelona (ES)); Dr SANTINELLI, Roberto (CERN); Dr ROISER, Stefan (CERN); MENDEZ MUNOZ, Victor (Universitat de Barcelona); BERNARDOFF, Vincent Roger Yvan (Univ. P. et Marie Curie (Paris VI) (FR)); ROMANOVSKIY, Vladimir (Institute for High Energy Physics (RU))

Presenter: LANCIOTTI, Elisa (CERN)

Session Classification: Poster Session

Track Classification: Distributed Processing and Analysis on Grids and Clouds (track 3)