



Contribution ID: 272

Type: **Poster presentation**

ATLAS DDM Workload Emulation

Monday, 14 October 2013 15:00 (45 minutes)

Rucio is the successor of the current Don Quijote 2 (DQ2) system for the distributed data management (DDM) system of the ATLAS experiment. The reasons for replacing DQ2 are manifold, but besides high maintenance costs and architectural limitations, scalability concerns are on top of the list.

The data collected so far by the experiment adds up to about 115 Peta bytes spread over 270 million distinct files. Current expectations are that the amount of data will be three to four times as it is today by the end of 2014. Further is the expansion of the WLCG computing resources pushing additional pressure on the DDM system by adding more powerful computing resources, which subsequently increases the demands on data provisioning. Although DQ2 is capable of handling the current workload, it is already at its limits. To ensure that Rucio will be up to the expected quality of service, a way to emulate the expected workload is needed. To do so, first the current workload observed in DQ2 must be understood in order to scale it up to future expectations.

This paper presents an overview of the theory behind workload emulation and discusses how selected core concepts are applied to the workload of the experiment. Further a detailed discussion is provided on how knowledge about the current workload is derived from central file catalogue logs, PanDA dashboard, etc. The discussion also addresses how this knowledge is utilized in the context of the emulation framework. Finally a description of the implemented emulation framework used for stress-testing Rucio is given.

Summary

Primary author: VIGNE, Ralph (University of Vienna (AT))

Co-authors: NAIRZ, Armin (CERN); SERFON, Cedric (CERN); SCHIKUTA, Erich (University of Vienna); STEWART, Graeme Andrew (CERN); Dr GOOSSENS, Luc (CERN); LASSNIG, Mario (CERN); BARISITS, Martin (CERN); BEERMANN, Thomas (Bergische Universitaet Wuppertal (DE)); GARONNE, Vincent (CERN)

Presenter: LASSNIG, Mario (CERN)

Session Classification: Poster presentations

Track Classification: Data Stores, Data Bases, and Storage Systems