

Advances in Service and Operations for ATLAS Data Management

Monday, 5 September 2011 14:50 (25 minutes)

ATLAS has recorded almost 5PB of RAW data since the LHC started running at the end of 2009. Many more derived data products and complimentary simulation data have also been produced by the collaboration and, in total, 55PB is currently stored in the Worldwide LHC Computing Grid by ATLAS. All of this data is managed by the ATLAS Distributed Data Management system, called Don Quixote 2 (DQ2).

DQ2 has evolved rapidly to help ATLAS Computing operations to manage these large quantities of data across the many grid sites at which ATLAS runs and to help ATLAS physicists get access to this data. In this paper we describe new and improved DQ2 services:

- Popularity service, which measures usage of data across ATLAS.
- Space monitoring and accounting at sites.
- Automated blacklisting service.
- Cleaning agents, which trigger deletion of unused data at sites.
- Deletion agents, to reliably delete unwanted data from sites.

We describe the experience of data management operation in ATLAS computing, showing how these services enable management of petabyte scale computing operations.

We illustrate the coupling of data management services to other parts of the ATLAS computing infrastructure, in particular showing how feedback from the distributed analysis system in ATLAS has enabled dynamic placement of the most popular data, helping users and groups to analyse the increasing data volumes on the grid.

Primary authors: Dr STEWART, Graeme Andrew (CERN); Dr GARONNE, Vincent (CERN)

Presenter: Dr STEWART, Graeme Andrew (CERN)

Session Classification: Monday 05th - Computing Technology for Physics Research

Track Classification: Track 1: Computing Technology for Physics Research