Contribution ID: **142** Type: **oral presentation**

Don Quijote - Data Management for the ATLAS Automatic Production System

Monday 27 September 2004 14:00 (20 minutes)

As part of the ATLAS Data Challenges 2 (DC2), an automatic production system was introduced and with it a new data management component.

The data management tools used for previous Data Challenges were built as separate components from the existing Grid middleware. These tools relied on a database of its own which acted as a replica catalog.

With the extensive use of Grid technology expected for the most part of the DC2 production, no longer can a data management tool be independent of the Grid middleware. Each Grid relies on its own replica catalog and not on an ATLAS specific tool.

ATLAS DC will attempt to use uniformly the resources provided by three Grids: NorduGrid, US Grid3 and LCG-2. Lecagy system will be supported as well.

The proposed solution was to build a data management proxy system which consists of a common high-level interface, whose implementation depends on each Grid's replica and metadata catalog as well as the storage backend (mainly "classic" GridFTP servers and SRM).

Don Quijote provides management of replicas in a services oriented architecture, across the several "flavours" of Grid middleware used by ATLAS DC.

With a higher-level interface common across several Grids (and legacy systems) a user (such as the new automatic production system) can seamlessly manage replicas independently of their hosting environment. Given the services-based architecture, a lightweight command line tool is capable of interacting uniformly within each Grid and between Grids (e.g. moving files from LCG-2 to US Grid 3 while maintaining attributes such as the Global Unique Identifier).

Author: BRANCO, M. (CERN)

Presenter: BRANCO, M. (CERN)

Session Classification: Distributed Computing Services

Track Classification: Track 4 - Distributed Computing Services