



Contribution ID: 451

Type: **Poster**

DIRAC File Replica and Metadata Catalog

Tuesday, May 22, 2012 1:30 PM (4h 45m)

File replica and metadata catalogs are essential parts of any distributed data management system, which are largely determining its functionality and performance. A new File Catalog (DFC) was developed in the framework of the DIRAC Project that combines both replica and metadata catalog functionality. The DFC design is based on the practical experience with the data management system of the LHCb Collaboration. It is optimized for the most common patterns of the catalog usage in order to achieve maximum performance from the user perspective. The DFC supports bulk operations for replica queries and allows quick analysis of the storage usage globally and for each Storage Element separately. It supports flexible ACL rules with plug-ins for various policies that can be adopted by a particular community. The DFC catalog allows to store various types of metadata associated with the files and directories and to perform efficient queries for the data based on complex metadata combinations. Definition of file ancestor-descendent chains is also possible. It is implemented in the DIRAC distributed computing framework following the standard grid security architecture. In this contribution we describe the design of the DFC and its implementation details. The performance measurements are compared with other grid file catalog implementations. The experience of the DFC Catalog usage in the ILC Collaboration is discussed.

Primary authors: Dr TSAREGORODTSEV, Andrei (Universite d'Aix - Marseille II (FR)); POSS, Stephane (Unknown)

Presenter: Dr TSAREGORODTSEV, Andrei (Universite d'Aix - Marseille II (FR))

Session Classification: Poster Session

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