Contribution ID: **57**   Type: **Oral Presentation**

DIRAC Data Management Framework

Tuesday, 15 March 2016 15:00 (30 minutes)

DIRAC Project is developing software for building distributed computing systems for the needs of research communities. It provides a complete solution covering both Workload Management and Data Management tasks of accessing computing and storage resources. The Data Management subsystem of DIRAC includes all the necessary components to organize data in distributed storage systems. It has a versatile File Catalog (DFC) service to keep track of data file physical replicas. This service is a central component to build a logical File System of DIRAC presenting all the distributed storage elements as a single entity for the users with transparent access to the data. The DFC service provides also a Metadata Catalog functionality to classify data with user defined tags. This can be used for an efficient search of the data necessary for a particular analysis. The Data Management system provides also support for usual data management tasks of uploading/downloading, replication, removal files with a special emphasis on the bulk data operations involving large numbers of files. Automation of data operations driven by new data registrations is also possible. In this controbution we will make an overview of the DIRAC Data Management System and will give examples of its usage by several research communities.

Primary author: Dr TSAREGORODTSEV, Andrei (CPPM-IN2P3-CNRS)

Presenter: Dr TSAREGORODTSEV, Andrei (CPPM-IN2P3-CNRS)

Session Classification: Data Management Session I

Track Classification: Data Management