

From DIRAC towards an Open Source Distributed Data Processing Solution

Description of content and intended audience- the outcome you expect to achieve.

The DIRAC software framework allows easy extension of its already existing functionality for the needs of new user communities. A number of user communities in different domains already adopted the DIRAC software as the basis for their distributed computing systems.

The outcome expected is to raise awareness on new communities that might be interested on a tool like DIRAC to simplify their access to distributed computing.

Relevant URL (if any)

<http://diracgrid.org>

Printable summary: this is the only section of the abstract that will be published in the Book of Abstracts.

The DIRAC Distributed Computing Project (the “DIRAC Project”) provides an open source software framework for building distributed computing systems. The DIRAC software framework is designed to create services for distributed computations in various environments such as grids, clouds and clusters. The DIRAC software was initially developed to meet the needs of the LHCb HighEnergy Physics experiment at CERN. It was first used for massive production of the LHCb modeling data and was later extended to support all the distributed computing operations, including Data Management and User Analysis. The DIRAC Project introduced several innovations focusing on the needs of large and distributed scientific collaborations.

Based on this framework, the DIRAC Project develops and maintains software systems supporting standard distributed computing tasks. The DIRAC Workload Management System provides all the necessary components to ensure efficient execution of user jobs using heterogeneous computing resources. It makes a special emphasis on fault tolerant operations in unstable distributed computing environments. Support is provided for large user communities with complex internal policies of computing resource usage for massive data processing. The DIRAC Data and Storage Management Systems provide tools for seamless access to various types of data storage, data cataloging and classification as well as for efficient and reliable massive data replication.

Primary authors: TSAREGORODTSEV, Andrei (CNRS); Dr GRACIANI DIAZ, Ricardo (University of Barcelona)

Co-authors: CASAJUS, Adria (CERN); Dr POSS, Stephane (AlpesLaser); MENDEZ, Victor (UAB)

Presenter: Dr GRACIANI DIAZ, Ricardo (University of Barcelona)