



Contribution ID: 18

Type: Poster

Distributed databases services in the EGEE Grid and beyond

Tuesday 23 September 2008 16:03 (1 minute)

Describe the activity, tool or service using or enhancing the EGEE infrastructure or results. A high-level description is needed here (Neither a detailed specialist report nor a list of references is required).

A wide-range of mission-critical applications depend on the production infrastructure that has been established through the EGEE projects. Many of the services that are part of this infrastructure rely on databases - not only middleware and storage-related services, but also key infrastructure and operations services.

Report on the impact of the activity, tool or service. This should include a description of how grid technology enabled or enhanced the result, or how you have enabled or enhanced the infrastructure for other users.

Building on a small portfolio of techniques, highly reliable and flexible distributed database services, that form the basis for key grid services and infrastructures, have been established. These have allowed widespread production usage of heterogeneous grid infrastructures to support massive scale data and computing challenges. The techniques are independent of application area and are applicable to all types of grid - ranging from research to production. Most importantly, they address two of the key challenges that still face grids today, namely the provision of grid services with low cost of ownership, as well as low cost of entry, both for new grid service providers as well as new application communities.

Describe the added value of the grid for your activity, or the value your tool or service adds for other grid users. This should include the scale of the activity and of the potential user community, and the relevance for other scientific or business applications.

In the specific case of the Large Hadron Collider (LHC) experiments, a large proportion of non-event data, including detector conditions, calibration, geometry and production book-keeping meta-data, are stored in (distributed) relational databases. These services - together with data and storage management services - are rated consistently by the LHC VOs as amongst the most critical for their production work.

This contribution describes the distributed database service architecture, based on database clusters and data replication and caching techniques, that is now implemented at CERN and ten WLCG Tier-1 sites. It includes the technology chosen, the service set-up, monitoring and production experience gained with this infrastructure at CERN and Tier1 sites throughout several experiment data challenges and the Common Computing Readiness Challenges.

Author: GIRONE, Maria (CERN)

Presenter: GIRONE, Maria (CERN)

Session Classification: Demos and Posters

Track Classification: Poster