Contribution ID: **95** Type: **oral presentation**

GRelC Data Access Service: Extreme Performance Managing Grid Databases

Thursday 10 May 2007 14:40 (20 minutes)

Describe the scientific/technical community and the scientific/technical activity using (planning to use) the EGEE infrastructure. A high-level description is needed (neither a detailed specialist report nor a list of references).

The activity aims at integrating within the EGEE infrastructure the GRelC DAS, a GSI enabled web service allowing high performance access in grid environments both to relational and non-relational DBs.This service currently runs on the GILDA t-Infrastructure and fully interacts with several gLite services.The target community includes scientists belonging to different domains within the EGEE collaboration that need for their applications to transparently and securely access huge distributed DBs.

Report on the experience (or the proposed activity). It would be very important to mention key services which are essential for the success of your activity on the EGEE infrastructure.

We present the GRelC DAS, a data grid access service developed by the GRelC Team (SPACI Consortium and Univ. of Salento).Deployment and experimental results related to the GILDA test-bed as well as porting issues on gLite middleware will also be presented.The GRelC DAS is a GSI enabled web service addressing extreme performance, interoperability and security.It provides an uniform access interface to relational and non-relational data sources. The GRelC DAS supports both basic and advanced functionalities. The aim of this service is to efficiently, securely and transparently manage databases on the grid across VOs, with regard to emerging and consolidating grid standards and specifications and interoperation with gLite. We will detail the key components of the GRelC DAS architecture, presenting queries (with advanced functionalities such as chunking and compression),user management, technological choices, heterogeneous DB access and security issues (GSI support, VOMS integration, etc).

With a forward look to future evolution, discuss the issues you have encountered (or that you expect) in using the EGEE infrastructure. Wherever possible, point out the experience limitations (both in terms of existing services or missing functionality)

Several EGEE-VOs need access to relational DBs for their experiments. Even though in the last few years many efforts have been concentrated in this direction, performances have not been completely addressed. Within the EGEE middleware there is a lack of services concerning this topic and current solutions do not fully address user's/VO's requirements. The GRelC DAS tries to bridge this gap providing a well suited solution strongly integrated with the existing gLite components.

Describe the added value of the Grid for the scientific/technical activity you (plan to) do on the Grid. This should include the scale of the activity and of

the potential user community and the relevance for other scientific or business applications

The GRelC DAS: (i) could attract new communities that usually work intensively with databases (e.g. financial applications exploiting datawarehouse and data mining); (ii) help existing ones (bioinformatics, astrophysics, etc.), in performing new tests and experiments; and (iii) improve e-Science research results. Due to the transversal role of this service, many experiments and Virtual Organizations within the EGEE project would benefit from it. Within such a multifaceted environment new kinds of queries could also be invented to perform distributed computation based on a work sharing approach.
Moreover, replication (cross-DBMS) facilities that we developed within our service could help grid users and site admins to ease (i) deployment of data sources, (ii) data migration and (iii) database backup.

Primary authors: Prof. ALOISIO, Giovanni (SPACI Consortium and University of Salento, Italy); Prof. BARBERA, Roberto (INFN Sez. di Catania, Italy); Dr FIORE, Sandro (SPACI Consortium and University of Salento, Italy)

Co-authors: Mr NEGRO, Alessandro (SPACI Consortium and University of Salento, Italy); Dr GIORGIO, Emidio (INFN Sez. di Catania, Italy); Dr MIRTO, Maria (SPACI Consortium and University of Salento, Italy); Dr CAFARO, Massimo (SPACI Consortium and University of Salento, Italy); Mr VADACCA, Salvatore (SPACI Consortium and University of Salento, Italy)

Presenter: Dr FIORE, Sandro (SPACI Consortium and University of Salento, Italy)

Session Classification: Data Management

Track Classification: Data Management