



The AMGA Metadata Service

Thursday 2 March 2006 14:40 (20 minutes)

We present the ARDA Metadata Grid Application (AMGA) which is part of the gLite middleware. AMGA provides a lightweight service to manage, store and retrieve simple relational data on the grid, termed metadata.

In this presentation we will first give an overview of AMGA's design, functionality, implementation and security features. AMGA was designed in close collaborations with the different EGEE user communities and combines high performance, which was very important to the high energy physics community, with fine-grained access restrictions required in particular by the BioMedical community. These access restrictions also make full use of the EGEE VOMS services and are based on grid certificates. To show to what extent the users' requirements have been met, we will present performance measurements as well as show uses-cases for the security features.

Several applications are currently using AMGA to store their metadata. Among them are the MDM (Medical Data Manager) application implemented by the BioMedical community, the GANGA physics analysis tool from the Atlas and LHCb experiments and a Digital Library from the generic applications.

The MDM application uses AMGA to store relational information on medical images stored on the grid plus information on patients and doctors in several tables. User applications can retrieve images based on their metadata for further processing. Access restrictions are of the highest importance to the MDM application because the stored data is highly confidential. MDM therefore makes use of the fine-grained access restrictions of AMGA.

The GANGA application uses AMGA to store the status information of jobs running on the grid which can be controlled by GANGA. AMGA's simple relational database features are mainly used to ensure consistency when several GANGA clients of the same user are accessing the stored information remotely.

Finally, the Digital Library project makes similar use of AMGA as the MDM application but provides many different schemas to store not only images but information on texts, movies or music. Another difference is that there is only a central librarian updating the library while for MDM updates are triggered by the many image acquisition systems themselves.

This presentation will also discuss future developments of AMGA, in particular its features to replicate or federate metadata. They will mainly allow users to make use of a better scaling behaviour but could also allow better security by using federation to physically separate

metadata. The replication features will be compared to current proprietary solutions.

AMGA provides a very lightweight metadata service as well as basic database access functionality on the Grid. After a brief overview of AMGA's design, functionality, implementation and security features we will show performance comparisons of AMGA with direct database access as well as other Grid catalogue services. Finally the replication features of AMGA are presented and a comparison done with proprietary database replication solutions.

Author: Dr KOBLITZ, Birger (CERN-IT)

Co-author: Mr SANTOS, Nuno (CERN-IT)

Presenter: Dr KOBLITZ, Birger (CERN-IT)

Session Classification: 2b: Data access on the grid

Track Classification: Data access on the grid