

# Security in a Replicated Metadata Catalogue

Thursday 10 May 2007 09:40 (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 gLite-AMGA metadata has been developed by NA4 to provide simple relational metadata access for the EGEE user community. As advanced features, which will be the focus of this presentation, AMGA provides very fine-grained security also in connection with the built-in support for replication and federation of metadata. AMGA is extensively used by the biomedical community to store medical images metadata, digital libraries, in HEP for logging and bookkeeping data and in the climate community.

**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.**

The biomedical and digital libraries have been deployed using a centralized structure already for some time. They now intend to decentralize their activity to increase reliability and scalability without cutting back on security. The deployed systems make use of the EGEE workload management system in case of the biomed use case to schedule jobs to run over the stored data and both the biomed use case and the digital library one use a file-catalogue and the storage elements.

**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)**

Security in a replicate environment is a very complex problem, because it requires the nodes to establish some sort of trusted relationship. We will show how these problems have been tackled, which may be of interest also for other services in a Grid environment.

**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 biomedical community intends to deploy a distributed metadata system for medical images consisting of various sites, which range from hospitals to computing centres. Only safe sharing of the highly sensitive metadata as provided in AMGA makes such a scenario possible. Another scenario are digital libraries, which federate copyright protected (meta-)data into a common catalogue.

**Primary authors:** Dr KOBLITZ, Birger (CERN-IT); Mr SANTOS, Nuno (EPFL Lausanne)

**Presenter:** Dr BIRGER, Koblitz

**Session Classification:** Data Management

**Track Classification:** Data Management