

# GSAF: GRID Storage Access Framework

Thursday 10 May 2007 14:00 (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 GSAF Project is carried out by INFN Catania and IR&T Engineering s.r.l.(a SME located in Catania, <http://www.irt-engineering.com>). INFN leads research activities to port Industrial Use Cases over the Grid and the IR&T aims to design a web applications that adopts Grid as a Content Management System. We analyzed the high level requirements of a classic web application in order to identify common points and shared problems and isolate peculiarities for porting it over the Grid infrastructure.

**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.**

GSAF hides the complexity and the fragmentation of the underlying APIs (one for each data service) and ensures both coherence and flexibility to the underlying application. GSAF, both as framework and application, interacts to the Storage Element, File Catalogue, and Metadata Catalogue allowing both applications (through the APIs) and users (by means of the web interface) to perform the following functionalities: 1) collection management (collection tree browsing, collection creation, collection deletion); 2) entry management (entry list, entry search, entry creation, entry deletion); 3) metadata schema management (attribute list, attribute creation, attribute deletion); 4) File management (upload to SE, download from SE, register file on File Catalogue, deregister file from File Catalogue); 5) Access Control (login based on VOMS proxy file); 6) ACLs for collections (owner permission list, change access mode, change owner, group list, group add, group deletion).

**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)**

1)AMGA does not allow separation between user X.509 credentials and the software connection object. 2)There were not Java GFAL APIs and we wrote them by ourselves wrapping the C APIs. 3)There is some unexpected incompatibility between GFAL and lcg-\* tools, uploading a file to a SE and registering its entry into LFC at the same time looks impossible. 4)Existing APIs seem to support only a single user logged while we need to bound the proxy file with the working session of the user.

**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**

Every service of DMS is independent from each other and works in a “stand alone” mode. Then, applications must refer to those services in a decoupled way. This fragmentariness obliges software engineers and web designers to consider a vertical architecture in order to access any single service and this always induces the same problem: application must take care of coherence and synchronization of data manipulation. From the end user side it is possible (using CLI tools), to manage objects across file catalogue rather than the metadata catalogue or storage element but all logical relationships must be kept clearly in the mind and can not be demanded to any framework.. We focused our effort on both of these two different aspects. We created an Object Oriented framework to help developers to write Java applications that use Grid as Data Access Layer. Subsequently, we tried to satisfy end user needs developing an easy-to-use web interface to access Grid DMS remotely.

**Primary author:** Dr SCIFO, Salvatore (INFN)

**Presenter:** Dr SCIFO, Salvatore (INFN)

**Session Classification:** Data Management

**Track Classification:** Data Management