Contribution ID: **32** Type: **Demo contribution**

# Secured Medical Data Management on the EGEE grid

*Wednesday 1 March 2006 18:30 (20 minutes)*

** Clinical data management versus computerized medical analysis

The medical community is routinely using clinical images and associated medical data for diagnosis, intervention planning and therapy follow-up. Medical imagers are producing an increasing number of digital images for which computerized archiving, processing and analysis are needed.

DICOM (Digital Image and COmmunication in Medicine) is today the most widely adopted standard for managing medical data in clinics. DICOM is including both the image content and additional information on the patient and the acquisition. DICOM was exclusively designed to respond clinical requirements. The interface with computing infrastructures for instance is completely lacking.

Grids are promising infrastructures for managing and analyzing the huge medical databases. However, the existing grid middlewares are often only providing low level data management services for manipulating files, making difficult the gridification of medical applications. Medical data often have to be manually transferred and transformed from hospital sources to grid storage before being processed and analyzed. To ease applications development there is a need for a data manager that: (i) shares access to medical data sources for computing without interfering with the clinical practice; (ii) ensures transparency so that accessing medical data does not require any specific user intervention; and (iii) ensures a high data protection evel to respect patients privacy.

** MDM: a grid service for secured medical data management

To ease medical applications devlopment, We developed a Medical Data Manager (MDM) service with the support of the EGEE uropean IST project. This service was developped on top of the new generation middleware release, gLite.

The data management in the gLite middleware is based on a set of Storage Elements which are exposing a same standard Storage Resource Manager SRM) interface. The SRM is handling local data at a file level. Additional services such as GridFTP or gLiteIO are coexisting on storage elements to provide transfer capabilities. In addition to storage resources, the gLite data management system includes a File Catalog (Fireman) offering a unique entry point for files distributed on all grid storage elements. Each file is uniquely identified through a Global Unique IDentifier (GUID).

The Medical Data Management service architecture is diagrammed in figure 1. On the left, is represented a clinical site: various imagers in an hospital are pushing the images produced on a DICOM server. Inside the hospital, clinicians can access the DICOM server content through DICOM clients. In the center of figure 1, the MDM internal logic is represented. On the right side, the grid services interfacing with the MDM are shown. To remain compatible with the rest of the grid infrastructure, the MDM service is based on an SRM-DICOM interface software which translates SRM grid requests into DICOM transactions addressed to the medical servers. Thus, medical data servers can be transparently shared between clinicians (using the classical DICOM interface inside hospitals) and image analysis scientists (using the SRM-DICOM interface to access the same data bases) without interfering with the clinical practice. An internal scratch space is used to transform DICOM data into files that are accessible through data transfer services (GridFTP or gLiteIO). For enforcing data protection, a highly secured and fault tolerant encryption key catalog, called hydra, is used. In addition, all DICOM files exported to the grid are anonimized. A metadata manager is in charge of holding the metadata extracted from DICOM headers and to ease data search. The AMGA ervice is used for ensuring secured storage of these very sensitive data. The AMGA server holds a relation between each DICOM slice and the image metadata.

The security model of the MDM relies on several components: (i) file access control, (ii) files anonymization, (iii) files encryption, and (iv) secured access to metadata. The user is coherently identified through a single X509 certificate for all services involved in security. The file access control is enforced by the gLiteIO service which accepts Access Control Lists (ACLs). The hydra key store and the AMGA metadata service both accept ACLs. To read an image content, a user needs to be authorized both to access the file and to the encryption key. The access rights to the sensitive metadata associated to the files are administrated independently. Thus, it is possible to grant access to an encrypted file only (e.g. for replicating a file without accessing to the content), to the file content (e.g. for processing the data without revealing the patient identity), or to the full file metadata (e.g. for medical usage). Through ACLs, it is possible to implement complex use cases, granting access rights to patients, physicians, healthcare practitioners, or researchers independently.

** Medical image analysis applications

On the client side, three levels of interfaces are available to access and manipulate the data hold by the MDM: (1) the standard SRM interface, can be used to access encrypted images provided that their GUID is known; (2) the encryption middleware layer can both fetch and decrypt files; (3) the fully MDM aware client provides access to the metadata associated to files in addition.

The Medical Data Manager has been deployed on several sites for testing purposes. Three sites are actually holding data in three DICOM servers installed at I3S (Sophia Antipolis, France), LAL (Orsay, France) and CREATIS (Lyon, France). An AMGA catalog has also been set up in CREATIS (Lyon) for holding all sites' metadata, and an hydra key store is deployed at CERN (Geneva, Switzerland).

The testbed deployed has been used to demonstrate the viability of the service by registering and retrieving DICOM files across sites. Registered files could be retrieved and used for computations from EGEE grid nodes transparently. The next important milestone will be to experiment the system in connection with hospitals by registering real clinical data freshly acquired and registered on the fly from the hospital imagers.

The Medical Data Manager is an important service for enabling medical image processing applications on the EGEE grid infrastructure. Several existing applications could potentially use the MDM such as the GATE, CDSS, gPTM3D, pharmokinetics, and Bronze Standard applications currently deployed on the EGEE infrastructure.

## Summary

This abstract describes the effort to deploy a Medical Data Management service on top of the EGEE grid infrastructure. The most widely accepted medical image standard, DICOM, was developed for fulfilling clinical practice. It is implemented in most recent medical image acquisition and analysis devices. The EGEE middleware is using the SRM standard for handling grid files. Our prototype is exposing an SRM compliant interface to the grid middleware, transforming on the fly SRM requests into DICOM transactions. The prototype ensures user identification, strict file access control and data protection through the use of relevant grid services. This Medical Data Manager is easing the access to medical databases needed for many medical data analysis applications deployed today. It offers a high level data management service, compatible with clinical practices, which encourages the migration of medical applications towards grid infrastructures. A limited scale testbed has been deployed as a proof of concept of this new service.

**Authors:** FROHNER, Akos (CERN); KOBLITZ, Birger (CERN); PERA, Christophe (CNRS); JOUVENOT, Daniel (CNRS); Dr MONTAGNAT, Johan (CNRS); SANTOS, Nuno (CERN); KUNSZT, Peter (CERN)

**Presenter:** Dr MONTAGNAT, Johan (CNRS)

**Session Classification:** Poster and Demo session + cocktail

**Track Classification:** Demo session