

## On the Use of Grids for Distributed Mammogram Analysis

S. R. AMENDOLIA<sup>1</sup> - F. ESTRELLA<sup>2</sup> - C. DEL FRATE<sup>3</sup> - T. HAUER<sup>1</sup> - R. McCLATCHEY<sup>2</sup> - T. SOLOMONIDES<sup>2</sup>

<sup>1</sup>ETT Division, CERN, 1211 Geneva 23, Switzerland

<sup>2</sup>CCCS Research Centre, Univ. of West of England, Frenchay, Bristol BS16 1QY, UK

<sup>3</sup>Istituto di Radiologia, Università di Udine

### Abstract

Medical conditions such as breast cancer, and mammograms as images, are extremely complex with many degrees of variability across the population. An effective solution for the management of disparate mammogram data sources that provides sufficient statistics for complex epidemiological study is a federation of autonomous multi-centre sites which transcends national boundaries. Grid-based technologies are emerging as solutions for managing and collaborating distributed resources. The MammoGrid project, as one example of a HealthGrid, is developing a Grid-aware medical application which manages a European-wide database of mammograms and is investigating the potential of the Grid to support effective co-working among mammogram analysts throughout the EU.

### 1. Introduction

A distributed database of normalised images that reflects the spread of pathologies across the population would be an invaluable tool for the epidemiologist, while an understanding of the variation in medical image acquisition protocols is essential to a radiologist or radiographer (radiologic technician) in a screening programme. By exploiting emerging Grid technology [1], the aim of the MammoGrid [2] project is to develop a Europe-wide database of mammograms that will be used to facilitate a set of important healthcare applications and to explore the potential of the Grid to support effective co-working between healthcare professionals.

In particular, the project aims to prove that Grids infrastructures can be practically used for collaborative medical image analysis. This leads to several technical issues, including the standardization of mammograms, the design of an appropriate clinical workstation and the distribution of data, images and clinician queries across a Grid-based database. The MammoGrid project [2] aims to prove the viability of the Grid by harnessing its power to enable radiologists from geographically dispersed hospitals to share standardized mammograms, to compare diagnoses (with and without computer-aided detection of tumours) and to perform sophisticated epidemiological studies across national boundaries. MammoGrid has defined an imaging workstation architecture, an information infrastructure to connect radiologists across a Grid, and a DICOM-compliant [3] object model residing in multiple, distributed data stores, currently in Italy and the UK. The project is investigating a number of relevant technologies that are being harnessed together via the Grid including Mirada Solutions' Standard Mammogram FormTM[4]; aspects of the CERN/CMS CRISTAL kernel [5]; and AliEn [6] a lightweight Grid developed for CERN/ALICE.

The current status of MammoGrid is that a single 'virtual organization' (VO) AliEn-based solution has been demonstrated and images have been accessed and transferred between hospitals in the UK and Italy. The next stage is to provide rich metadata structures and a distributed database in a multi-virtual organisation environment to enable epidemiological queries to be serviced and the implementation of a service-oriented (OGSA-compliant) architecture for the MammoGrid

*Correspondence:* S.R. Amendolia, T. Hauer - ETT Division, CERN, 1211 Geneva 23, Switzerland - E-mail: {Salvator.Amendolia, Dmitry.Rogulin}@cern.ch; F. Estrella, R. McClatchey, T. Solomonides, CCCS Research Centre, Univ. of West of England, Frenchay, Bristol BS16 1QY, UK - E-mail: Florida.Estrella@cern.ch {Richard.McClatchey, Tony.Solomonides}@uwe.ac.uk; C. del Frate - Istituto di Radiologia, Università di Udine - E-mail: Chiara.delfrate@med.uniud.it

Information Infrastructure.

## 2. Current Status and Future Plans.

The MammoGrid project has recently delivered its first proof-of-concept prototype (see figure 1) enabling clinicians to store (digital) mammograms along with appropriately anonymised patient meta-data [5] and to provide controlled access to mammograms both locally and remotely stored. A typical database comprising several hundred mammograms is being created for user tests of the query handler. The prototype comprises a high-quality clinician visualization workstation used for data acquisition and inspection, a DICOM-compliant interface to a set of medical services (annotation, security, image analysis, data storage and querying services) residing on a so-called 'Grid-box' and secure access to a network of other Grid-boxes connected through Grids middleware. The very nature of a project like MammoGrid implies that it is inconceivable to define an exhaustive list or even complete classification of all possible queries which the radiologists may need to run against the distributed database. Inevitably, when the user community starts using such a system, the requirements will undergo adjustments and extension. It is proposed that the design fol-

lowed in MammoGrid, with extensive use of meta-data, is both capable of handling such complex queries in an efficient way and flexible enough to adapt to changing requirements. A design which handles queries using a reflexive data model [5] has been presented as the proposed query model for the MammoGrid infrastructure [7].

Within the next year a rigorous evaluation of the prototype will indicate the usefulness of the Grid as a platform for distributed mammogram analysis and in particular for resolving clinicians' queries. The system will be tuned for performance and for security prior to the release of a second prototype at the end of the project in mid 2005. It is intended that the MammoGrid medical services for this second prototype will adhere to emerging Grids standards. The Grid platform provides an effective distributed computing model for harnessing the use of and access to massive amounts of medical image data across Europe. Moreover, it will enable a standardized, distributed digital mammography resource for improving diagnostic confidence. The use of web services or Grid services allows for containment and modularity, minimizing the requirements for shared understanding. This platform creates a dynamic loosely coupled client-server environment effective for distributed image analysis.

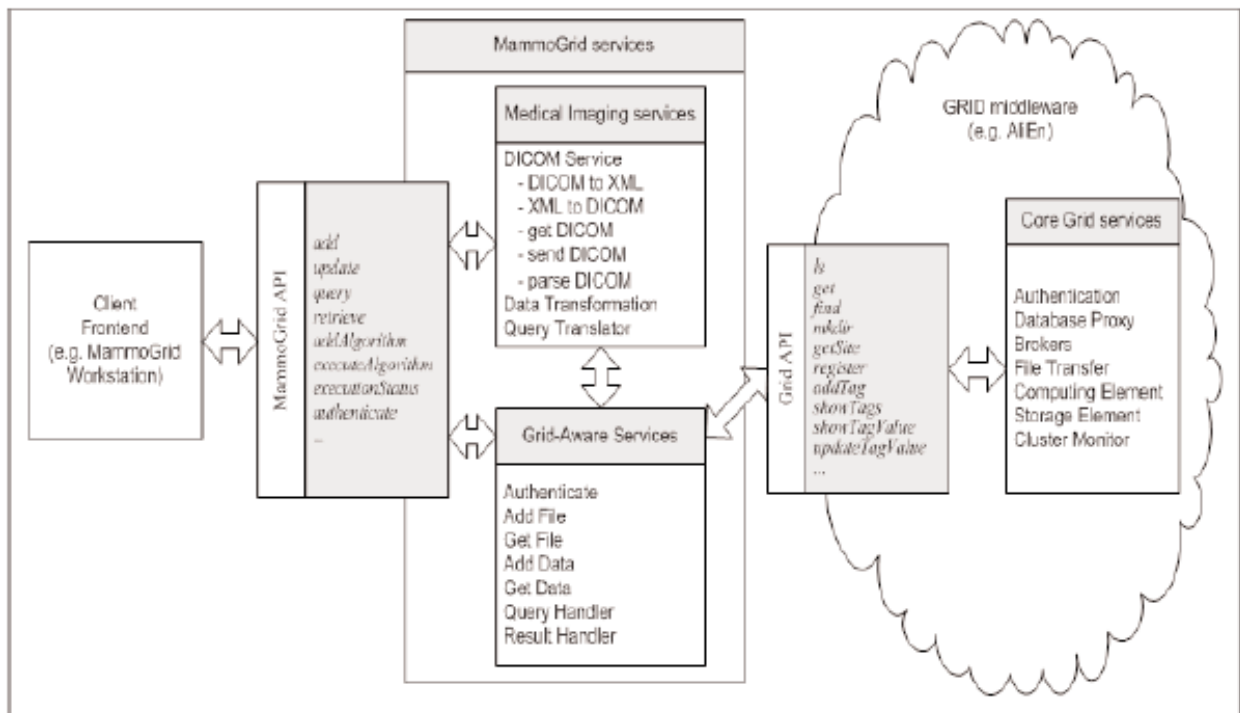


Figure 1: High Level MammoGrid prototype architecture

## References

- [1] I Foster, C Kesselman & S Tueke., "The Anatomy of the Grid - Enabling Scalable Virtual Organisations", Int. Journal of Supercomputer Applications, 15(3), 2001.
- [2] "MammoGrid ". EU Contract IST 2001-37614. See: <http://MammoGrid.vitamib.com>
- [3] DICOM Digital Imaging and Communications in Medicine. <http://medical.nema.org>
- [4] SMF: The Standard Mammogram Form. See: <http://www.mirada-solutions.com/smf.htm>
- [5] F. Estrella et al, "Meta-Data Objects as the Basis for System Evolution". LNCS Vol 2118, pp 390-399 ISBN 3-540-42298-6 Springer-Verlag, 2001.
- [6] P. Saiz et al., "AliEn" see: <http://alien.cern.ch>
- [7] S. R Amendolia et al., "Grid Databases for Shared Image Analysis in the MammoGrid Project". Proc of the 8th IDEAS conference, IEE Press, 2004.