GRID COMPUTING IN MEDICAL APPLICATIONS

P. Cerello, INFN, Sezione di Torino, Torino, Italy.

Abstract

Medical Applications can exploit GRID Services in many ways: some of them are computing intensive and require the use of big farms, some other are based on the availability of a transparent access to distributed data, so as they can be analysed, if possible interactively.

Classes of applications are reviewed in terms of the requirements they set for GRID Services and the advantages of a GRID approach. The corresponding use cases are described.

The structure and basic functionality of interfaces to GRID services are also discussed. An assessment of the present situation in terms of available services, development of application interfaces and deployment of prototypes is attempted.

Prospects of a widespread use of GRID services in medical applications are discussed in the last section.

INTRODUCTION

The added value which is expected from GRID Computing can be summarised as the capability, for a distributed community of authorised users, to access distributed resources and analyse distributed data in a transparent, quick and simple way.

Different medical applications can benefit of different kinds of functionality offered by baseline GRID services. With respect to High Energy Physics applications, there are very important differences that lead to specific requirements for the development of the interface to and the deployment of GRID services.

Medical applications typically deal with patient-related sensible data: in general, the global community of users which constitutes a Virtual Organisation (VO) shall not be granted equal access rights to all the VO-owned distributed datasets. Therefore, data and metadata management services must provide adequate features.

Also, it is quite common that a given VO would not share computing and storage resources and data related services with other VOs. Therefore, a vision based on “the” GRID concept would rather be replaced with a configuration in which many single-VO GRIDs live in parallel on the infrastructure provided by the low-level network services.

In the medical world, GRID sites are typically located in hospital facilities, with strict rules on data protection and ownership. Users, on the other hand, can be classified in two different groups: *end-users* (i.e. medical doctors), who would like to access a local or distributed system in the same way from a local *application-specific* Graphic User Interface, and *developers*, who would take advantage of a federated distributed database to improve the performance of their algorithms.

The access to GRID services should take place from an application specific wrapper rather than the other way around, particularly for *end-users*. It is acceptable, on the other hand, that for developers and for some computing intensive applications, the execution of an application task might be triggered via a GRID interface to the job submission services.

USE CASES AND REQUIREMENTS

The set of requirements for each application is strictly related to the set of use cases that must be implemented. It is therefore natural to define some classes of applications with similar features: for example, the simulation of a treatment planning is mostly computing intensive, the management of the electronic patient record is essentially a data and metadata management problem, while medical imaging analysis requires both large computing resources and high level data management services. The present section reviews some examples, describing the use cases that several applications implement and the related requirements to GRID services.

All the examples, however, share at least a common set of requirements related to authentication, authorisation and accounting for the community of users belonging to a VO.

Simulation-based Applications

Treatment planning and/or feedback are the typical simulation based applications: they rely on a Monte Carlo simulation engine, with the capability to describe the body geometry at the required resolution (for example, by importing the output of a raw CT scan [1] – Fig. 1).

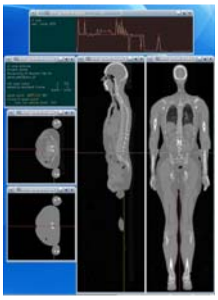

Fig. 1: A visualisation of the description of the body geometry in FLUKA obtained from an imported CT scan.

The simulation of the physics processes taking place when the body is exposed to a therapeutic beam allows the reconstruction of a 3D map of the released dose. Also, the Monte Carlo code can accurately reproduce a 3D map generated by a post-irradiation PET-scan (Fig. 2).

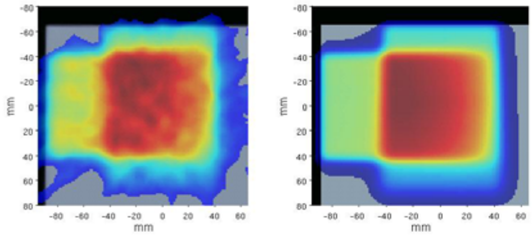

Fig.2: Comparison between a measured 3D map obtained from a post-irradiation PET scan and the FLUKA simulation.

From the GRID point of view, the simulation of many particles can be treated as a parallel problem and split into many processes, running in parallel on CPUs provided by one or more computer centres that are typically remote with respect to the user that issues the request.

The input size is generally small and there is typically no remote use of protected data and no requirement to access external databases.

Therefore, it is probably the simplest use case to be deployed on a GRID system: indeed, several applications of this kind have already been interfaced to different flavours of GRID middleware.

Image Analysis on demand

Medical Imaging Analysis is another class of applications that can exploit the availability of GRID services. If many hospitals belong to the same VO, they can share a common analysis algorithm release, which is made available by a *GRID Service Provider* acting as a server. From a local Graphic User Interface (GUI), an authorised user can upload a new image (exam) to the local Storage Element, register it to the Data Catalogue and then, depending on the features of the image, ask for the execution of the analysis algorithm (Fig. 3).

Depending on license issues, the analysis can take place locally (after retrieving the binary release from the server) or on the GRID Provider site (after the image transfer). The first option is more suitable in terms of bandwidth requirements, but it might not be allowed if the code is licensed. In any case, the result of the analysis is shown on the client GUI. The analysis algorithm can be executed as a batch job or as an interactive process, depending on the system configuration and the implementation choices.

In terms of GRID Services, the image analysis on demand is definitely a complex use case: it requires storage and computing resources on the client side, as well as a high-level GUI; on the provider side, computing resources and data and metadata management services are mandatory; file transfer functionality between the client and the server is also required. A prototype implementation, based on Computer Assisted Detection algorithms for the analysis of mammograms was developed by the GPCALMA project [2].

The most important advantage of image analysis on demand is the possibility for sites to make use of proprietary algorithms that are too expensive for a small hospital on a pay per image basis: the license-related costs are sustained by the GRID Service Provider.

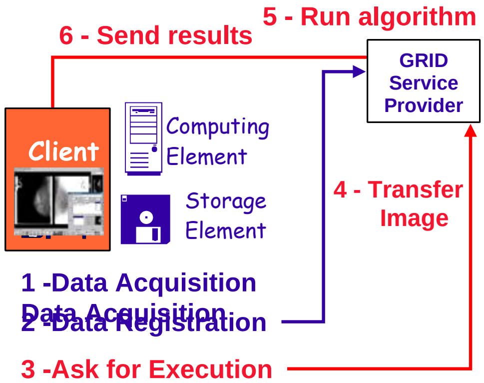

Fig. 3: Computer Assisted Detection algorithms can be triggered on demand from a remote repository/installation. After the acquisition, from a scan of analog images of the output of a digital device, data are registered in a VO dedicated instantiation of a Data Catalogue. A request of execution, depending on license-related legal issues, either retrieves the latest available algorithm version from the Provider or transfers the image(s) to the Provider. The results are then displayed on the user's GUI.

Tele-radiology and Epidemiology

It is sometimes useful to be able to share with some expert colleague the responsibility of a diagnosis or to be able to show, during a training session for young radiologists, some selected cases that are particularly significant.

Assuming a GRID VO is deployed in *N* sites, each of them owns a fraction of a distributed database of images/exams, which are all registered, with the associated *metadata*, in the VO Data Catalogue instantiation, hosted by a GRID Service Provider.

Any authorised user, from any site in the system, can query the Data & Metadata Catalogue and retrieve a list of images that meet a selected set of requirements (Fig. 4). In case of a training session or a co-diagnosis, it is required to retrieve and locally visualise the selected images.

Similarly, it is possible to study the distribution of any metadata for the selected set: that feature is all that is needed for any epidemiology study, when statistical distributions of metadata values are compared for different input data sets.

The required functionality can be implemented with a set of services similar to the image on demand use case.

Electronic Patient Record (EPR)

A database containing all the patient-related clinical history is, in terms of GRID Services, a Data and Metadata Catalogue, which is filled with new entries from any of the VO sites. It may look like a pretty simple implementation, but, with respect to the ones previously discussed, a strict and safe management of private patient data is mandatory.

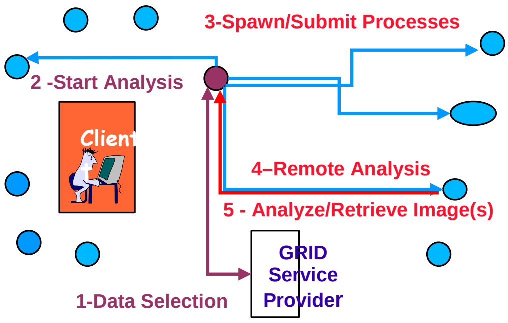

Fig. 4: Tele-radiology or epidemiology studies are based on a query to a VO Data Catalogue, which retrieves the list of entries meeting the set of requirements defined by the user. The analysis of the selected images can be executed in parallel on the sub-samples residing on each remote site. The partial results are sent to the Client site and then merged or post-analysed.

In addition, in order to provide meaningful and complete information about the clinical history of a patient, it is extremely important that a large fraction (if not all) the medical facilities that are part of a regional/ national infrastructure adopt the same EPR system.

Screening Programs

Perhaps, the most relevant impact of the use of GRID Services can be on screening programs, which are organised with the goal of an early diagnosis of cancer-related diseases on an asymptomatic population.

It has been proven that, for example in the case of breast cancer, screening can provide an early diagnosis which turns into much larger survival rates for the target population. It is still under discussion whether the same effectiveness can be obtained for other organs; however, lung cancer screening programs are being started in several countries, and plans are being drawn for colon cancer.

A national/continental screening program generates an amount of data which is in the same order of magnitude of a LHC experiment: for breast cancer, on women in the 50-69 age range, and for lung cancer (assuming that only the high-risk sub-sample of the population is monitored), about 1 PB/year would be collected on the European scale. In addition, such a database would be intrinsically distributed, with a number of involved sites much larger than for HEP experiments (and therefore a much smaller average size, since the global amount of data is similar).

A possible implementation of a screening use case is shown in Fig. 5: the basic idea is that the image analysis algorithm can be used as a selector of the sub-sample of exams that must be quickly analysed.

Each data collection site would daily add new exams to its local database. The Computer Assisted Detection (CAD) algorithm would run everyday, on each site, on the new data, select the exams with highest probability of a malignant structure and transfer only that fraction of data to the diagnosis centre. From their local application GUI, radiologists would daily monitor the data they receive and make decisions. All the other data would be analysed asynchronously.

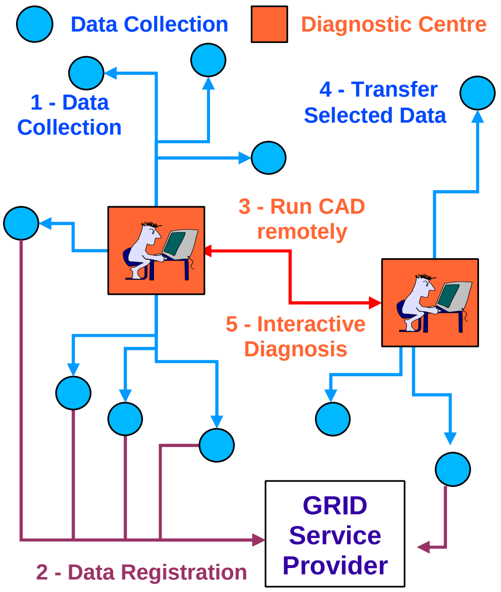

Fig. 5: The large scale screening use case: data collection takes place in parallel, asynchronously, in many sites. Data are registered in a VO Data Catalogue, which can be either a single central instance or a distributed catalogue with a central part that manages only the data mapping to the physical site location. A synchronous process, with a pre-defined periodicity (e.g., daily) would analyse the recent data and identify the subset of exams that requires a quick feedback by radiologists. Only that subset is transferred to a diagnostic centre and it is reviewed as quickly as possible.

Depending on the input dataset size, which is in turn related to the CAD algorithm performance (i.e., the average number of false positive regions/exams that are identified), the file transfer rate might become a critical parameter. The deployment of such a use case requires a careful balancing of computing and storage resources at each site.

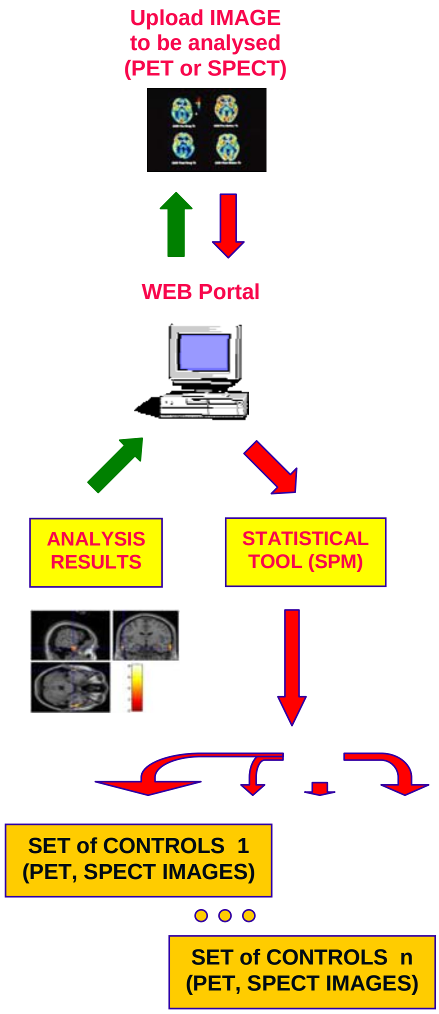

Fig. 6: Analysis of an image by comparison to a distributed set of reference images. The image to be analysed is uploaded to a Central Server, which forwards it to the sites where fractions of the reference database are stored. Analysis takes place in parallel. The partial results are sent back to the Central Server, which merges them and returns the result to the user.

Distributed Analysis by Image Comparison

An interesting class of medical imaging applications is based on the comparison of the exam being investigated with a set of similar exams that are part of a reference database, followed by a statistical analysis of the results.

An example is the use of the Statistical Parametric Mapping (SPM) algorithm [3] for the early diagnosis of the Alzheimer's disease (Fig. 6).

As for any statistical tool, the precision would increase with the number of reference cases, which, however, are typically difficult to obtain. Therefore, the use of a federated distributed reference database could provide added value to that class of applications.

For such a use case, client sites must be able to upload the candidate image/exam to a server site, which will forward it to all the clients containing a fraction of the reference database. Each of the sites would provide a partial statistical result to the server, which would merge them and send back the properly weighted result to the user that started the analysis (Fig. 6).

A prototype is being developed by the MAGIC-5 Collaboration and will be presented at the next HealthGrid Conference [4].

COMMON APPLICATION FEATURES

GRID prototypes should aim at the implementation and operation of services for Virtual Organisations that plan to implement any of the above-described use cases.

Prototypes should quickly evolve to systems that provide stable services on an ideally indefinite scale, without performance degradation when the number of users, sites and the data amount increase up to the highest values foreseen in the planning phase.

However, a GRID based system would be adopted on a large scale by commercial or research applications only if it proves to be economically efficient, in terms of money or time and effort.

Therefore, middleware must be easy to install and configure, up to the ideal situation in which one person on duty can deploy and monitor the behaviour of a full VO and at most one person per site can deal with standard interventions.

As already mentioned, stability and scalability are mandatory features.

Graphic User Interfaces to GRID Services

Another very important common feature is the possibility for applications (with the possible exception of simulation based use cases) to design and implement simple custom Graphic User Interfaces that act as wrapper to the calls to GRID services.

For a standard (image) analysis use case, it is possible to define a few basic high-level functionalities. An example is the implementation of the prototype designed by the GPCALMA collaboration for the mammographic use cases:

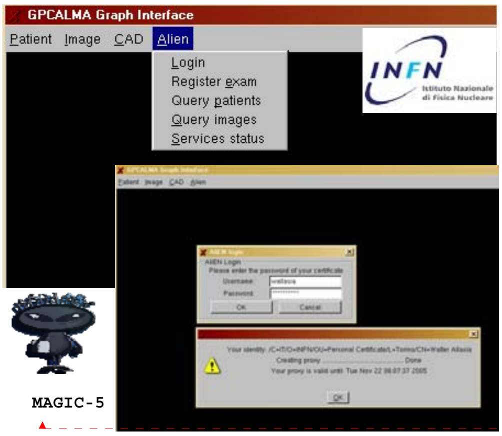

Fig. 7: An example of application custom GUI to the use of GRID Services: it provides access to the authentication service, to an interactive menu for the data registration, to dedicated menus for the query to the Data and Metadata Catalogue and to a script that monitors the present status of the VO sites.

- Access to the system: the interface should prompt the user with a login window that, upon insertion of the username and password, checks the user credentials and generates a proxy for the session (Fig. 7).
- Status of the system: a simple script would query the server that contains the VO configuration and then prompt the user for the list of (un)available sites.
- Query based on patient metadata (Fig. 8): it is particularly useful when trying to reconstruct the patient clinical history, including the retrieval of previous exams registered in the catalogue for a real time comparison with the most recent one. Access is granted on the basis that the user knows the patient code number.
- Query based on exam (image) metadata: it is possible for any user and does not require access rights, as it does not allow obtain any patient identification information. It is particularly useful for developers that need to select an input sample with some specific characteristics.

THE USER COMMUNITY

Several projects worldwide are developing Grid-based medical applications.

In Europe, a large community is involved in the EU-EGEE project [5] in the context of a *Biomedical Application Working Group*. The *GATE* project [6], starting from input MRI scans, computes radiotherapy planning for cancer patients to be treated. *Pharmacokinetics* focuses on the co-registration and analysis of a time sequence of MRI volumetric images, in order to generate diffusion maps of the contrast agent diffusion that characterises tumour tissues. *gPTM3D* [7] is an interactive tool for the 3D segmentation and volume reconstruction of complex organs (e.g. lung) or the entire body from modern CT-scans. All of them are computing intensive applications.

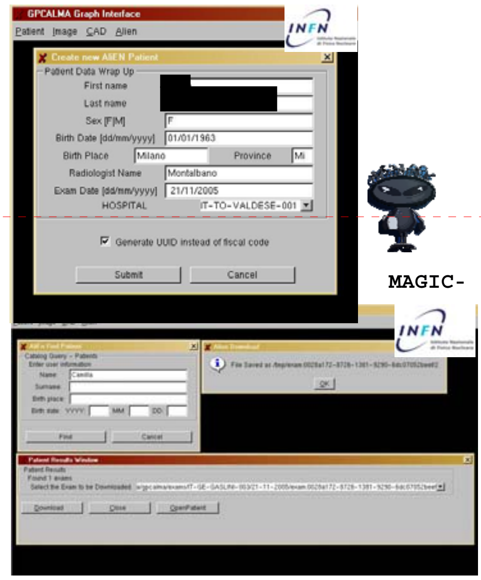

Fig. 8: An example of the menu for the Data Registration (upper window). Metadata are read from the input data set and prompted to the user, who can change their values if required and then start the registration process with the *Submit* button. The lower menu shows an example of the query based on patient metadata: all the exams meeting the query are listed. The selected one can be retrieved and interactively analysed in the local GUI.

Other projects, either national (e-Diamond [8], GPCALMA [2], MAGIC-5 [9]) or on a European scale (Mammogrid [10]), adopted a different approach and selected the *AliEn* [11] middleware for the implementation of GRID services. Here the focus is more on Data and Metadata management services, since their prototypes implement distributed analysis. In particular, GPCALMA (mammogram analysis) and MAGIC-5 (lung CT analysis) adopt *AliEn* Data Management services for data registration and search, while the analysis can take place either in batch mode, with *AliEn* jobs, or in interactive mode, with a *PROOF* session [12] configured and triggered on demand by the user via a GUI.

In North America, the Biomedical Informatics Research Network (BIRN) [13] "is building an infrastructure of networked high-performance computers,

Formatted: Font: 8 pt

data integration standards, and other emerging technologies" (including GRIDs).

In Asia, drug discovery and design applications are being developed in the framework of the DDGrid project [14].

A worldwide community, which has not been formed yet, would help focusing the goals and the best implementation choices to meet them in the shortest time with the highest efficiency.

PROSPECTS

Research working prototypes have already been (are being) implemented for many of the applications mentioned in this paper: the required functionality is available with more than one middleware flavour: *AliEn* [11], Globus [15], EGEE [5].

However, the path from prototypes to fully functional services running in medical facilities will not be easy (and nobody knows how quick it could be).

It would certainly be possible to start with simulation-based applications, where a node with User Interface functionality would be enough to exploit a GRID system.

Applications requiring data and private metadata management will require further testing (and possibly development) for services related to data protection and for the interface of Storage Element services to standard storage systems used in medical facilities. Similarly, distributed interactive analysis, which would be an important added value for many applications, requires further development.

Stability must be improved, but it is certainly true that the target for the LHC operation is more than satisfactory for medical applications.

In terms of the infrastructure, it is important to properly plan the use case implementation, by minimising the amount of data to be transferred. The deployment will be another extremely important issue: it must take into account the fact that sites are many more than in the case of HEP, with much smaller average size, and also the fact that many facilities do not plan to share their resources with other applications. Therefore, the "many Grids" model, with one Grid instantiation per VO is probably more suitable, at least in the first phase. If, at any time, services will be stable enough and sites will cross-validate each other's applications, it will be easy to merge different instantiations in a (partially) overlapping system.

During the whole process, efficiency should be taken as a gold standard to maximise the probability of success.

The High Energy Physics community is deeply involved in many medical applications, sometimes on the "gridification" side, sometimes directly in the development of application algorithms. I believe that the goal for a Research and Development community would be to provide high performance prototypes to be handed to some other entities we could call *Grid Service Providers*.

ACKNOWLEDGEMENTS

It is a pleasure to thank the organisers for the invitation to present the overview on the use of GRID services in medical applications. Some of the developers of all the applications that were mentioned supported my work: in particular, I would like to thank J. Montagnat, coordinator of the EU-EGEE Biomedical applications working group, L. Pinsky, who provided information about the US applications and the INFN/MAGIC-5 Collaboration, that includes the former GPCALMA.

REFERENCES

[1] A. Ferrari et al., "Update On the Status of the FLUKA Monte Carlo Transport Code", CHEP2006 proceedings
[2] P. Cerello et al., "GPCALMA: a grid-based tool for mammographic screening", METHODS OF INFORMATION IN MEDICINE 44 (2): 244-248 2005.
[3] http://www.fil.ion.ucl.ac.uk/spm
[4] http://valencia2006.healthgrid.org
[5] http://public.eu-egee.org
[6] http://www-lphe.epfl.ch/~PET/research/gate
]7[C. Germain-Renaud et al., "Interactive Volume Reconstruction and Measurement on the Grid", METHODS OF INFORMATION IN MEDICINE 44 (2) 2005.
[8] http://www.ediamond.ox.ac.uk
[9] R. Bellotti et al., "Distributed Medical Images Analysis on a GRID Infrasctructure ", accepted by FGCS
[10] http://mammogrid.vitamib.com
[11] http://alien.cern.ch
[12] M. Ballintijn et al., "The PROOF Distributed Parallel Analysis Framework based on ROOT", Proceedings of *Computing in High Energy and Nuclear Physics, 24-28 March 2003, La Jolla, California.*
[13] http://www.nbirn.net
[14]http://gridatasia.net/download/Shanghai/presentations /Drugdiscovery.pdf
[15] http://www.globus.org