A Grid Infrastructure for Image Registration and Segmentation

M. Burns, A.L. Rowland[†], D. Rueckert, J.V. Hajnal[†], D.L.G. Hill[‡]

Department Of Computing, Imperial College London
Imaging Sciences Department, Imperial College London[†]
Division of Imaging Sciences, King's College London‡

Abstract

Health care, medical research and drug discovery rely increasingly on the use of computationally intensive image processing algorithms to add value to the wealth of radiological images obtained through techniques such as Magnetic Resonance Imaging (MRI). Use of these algorithms on large cohorts remains problematic given the lack of a suitable IT infrastructure within many institutions and the problems associated with sharing imaging data in a secure and controlled way. Grid computing offers a potential solution to such problems through the promise of large scale distributed resource sharing and transparent access to distributed data sets. Unfortunately the adoption of grid technologies by the imaging community is not as widespread as it could be, mainly due to the absence of usable higher-level grid services. The Information Extraction from Images (IXI) project seeks to address this by developing a set of Open Grid Services Infrastructure (OGSI) compliant grid services (implemented with the Globus Toolkit 3) that allow the exposure of legacy image processing applications as services within a Grid envirnoment along with the ability to compose these services into complex image processing workflows, that execute across a distributed set of resources.

1. Introduction

The project seeks to 'grid enable' a repository of existing image-processing software such as that developed in [1, 2]. This software addresses two core image-processing problems namely that of image registration and image segmentation. Image registration enables quantitative comparisons to be made between images by determining the transformation required to match one image to another using varying numbers of degrees of freedom. Registration algorithms have a wide range of practical uses including but not limited to:

- Combining information from multiple imaging modalities, for example when relating functional information from nuclear medicine images to anatomy delineated in high resolution MR images.
- Monitoring changes in size, shape or image intensity over time intervals that might range from a few seconds in dynamic perfusion studies to several months or even years in the study of neuronal loss in dementia.
- Relating pre-operative images and surgical plans to the physical reality of the patient in the operating room during image guided surgery or the treatment suite during radiotherapy.
- Relating one individual's anatomy to a standardised atlas. [3]

The process of segmentation essentially involves delineating features of interest from within an image. Again this technology has a number of uses:

- Measuring volumes.
- Quantifying abnormalities.
- Visualisation of structures.
- Automatic identification of lesions.
- Measuring the shape of structures.

By grid enabling the above software we wish to:

- To show that the grid can scale medical image analysis to huge cohorts, using high performance compute resources located at various distributed locations.
- Achieve secure and efficient data sharing of multiple distributed image databases.
- Offer our image processing algorithms as services to the wider medical imaging community.

2. IXI Services

Since service orientation is the widely accepted model for building grid-based applications, our first goal was to expose existing imaging applications as grid services. These applications encompass a vast collection of software - a generic approach to service exposure was therefore desirable to avoid having to develop and deploy a new type of service each time a new image processing algorithm is implemented. We term this generic service the *IXI Core Service*.

2.1 IXI Core Service

The IXI Core Service represents the most basic type of service within the IXI grid infrastructure. This service is an OGSI compliant grid service wrapper (implemented with GT3) that allows us to expose our various image processing applications as services to the community. Generic IXI Core Services 'specialise' to offer a specific piece of image processing through the provision of a Medical Imaging Component Language (MICL) definition (see below). This definition describes *what* the service does, but not *how* or *where* it does it – these details are encapsulated behind the public interface of the service and are realised through the use of a component plug in model. Essentially the incoming SOAP message received by the service can be viewed as an abstract invocation of a particular algorithm (containing specific parameter values supplied by the client). Services can be configured with specific components that will take this message along with the MICL definition and translate this into a concrete invocation. The components configured with each service essentially depend on the resources available within our grid; at present we have tested components that will translate an incoming request into a simple command line and fork this on the host server and another that will take the incoming request map it to a Resource Specification Language (RSL) fragment and submit this to a Globus GRAM gatekeeper. Owing to the heterogeneity of grid resources one could also envisage numerous other components that may be beneficial, e.g. translating to Job Description Markup Language (JDML) and submitting to a high level grid scheduler (or resource broker) or interacting with a high performance parallel cluster. The benefits of this approach are obvious: clients wish to interact with the image analysis services themselves, not with the actual underlying resources and indeed not even with the existing (increasingly diverse and complex) grid middleware currently available to manage such resources. By providing one consistent interface to image analysis we effectively hide all these details from the client and through the use of the component plug-in model ensure that our services are not tied to any specific middleware package.

Data Provenance

Data provenance, the process of determining how a particular output was produced is becoming increasingly important in medical imaging applications. Such provenance information can be used for a variety of purposes including result verification or error correction and when used in conjunction with a workflow service can also save significant computation time (by not re-computing results that already exist) – recording of provenance information may even be a legal requirement (e.g. in the case of drug discovery studies). The IXI Core Service facilitates the publication of a variety of provenance data by exposing this data as OGSI SDEs that are populated on completion of a successful execution. By default we publish a range of generic provenance data including the full parameter list used to produce a particular image, computation time and envirnoment details (operating system, library versions, etc.). An extensible mechanism exists whereby service providers can plug in provenance item generators to provide more application specific provenance data.

We are currently investigating how the IXI services may be integrated with the Chimera Virtual Data System to provide a complete service based provenance system.

Medical Imaging Component Language (MICL)

MICL is an XML Schema based language that allows us to augment each IXI Core Service with a rich semantic description. We define an initial (but extensible) number of common base types commonly found in most imaging applications (e.g. image formats, image types, image transformations) and a rich element declaration allowing each algorithm to be fully described in terms of the parameters supported, the types of its input/outputs, the relationship between various inputs/outputs and a range of various other application meta-data e.g. contributing institution, versioning information, mandatory parameters, default values, naming conventions of outputs).

To facilitate flexible service discovery we also define further metadata including the particular class of image processing to which the algorithm belongs (essentially a hierarchical classification) and other domain specific attributes of the actual algorithm that may be of interest to potential users. e.g. details such as similarity measure and interpolation mode for registration algorithms.

IXI Core Service Discovery

Each MICL definition is stored as an XML file and is associated with a specific IXI Core Factory Service (i.e. a service capable of producing transient IXI Core Service instances). This file is processed by the factory and exposed to potential clients as an OGSI Service Data Element (SDE). By querying this SDE, clients can effectively determine the type of image processing offered by the service, the precise details of the inputs that are needed and the outputs that will be produced. Of course before this querying can take place clients will still need to first locate or discover a factory service. To facilitate this we use the GT3 Index Service / ServiceGroupRegistration portType. After reading the appropriate MICL definition the factory service will initiate an aggregation between the MICL SDE and the Index Service – the Index Service can therefore be viewed as a cache of MICL SDEs representing the various image processing services within our grid. As this cache is also exposed to clients as service data, flexible querying is possible (e.g. using XPath) on any MICL element/attribute to locate potential interesting or useful analysis services.

2.2 IXI Workflow Service

Most real world imaging studies tend to utilise sequential pipelines of image processing algorithms that are defined once by the researcher and then applied to a specific data sets. The ability to compose IXI Core Service instances together is analogous to a grid-enabled version of such image processing pipelines and therefore constitutes an essential requirement. The IXI Workflow Service, an OGSI compliant grid service that exposes a public interface allowing the submission, modification and monitoring of workflows meets this requirement. Workflows are expressed in a custom XML Schema based language that allows the structure of the required image-processing pipeline to be defined. Typically a workflow will be composed of a number of distinct processing stages. Within each stage we define a number of tasks – these tasks essentially represent an iterator over the input data set (with each task being realized by an IXI Core Service instance). Tasks can reference data produced at any preceding stage/task thus allowing transparent (potentially distributed) data flow between workflow stages.

It is essential that this data flow occurring between workflow stages is compatible in terms of the type of data produced and consumed at each stage. By inspecting the appropriate MICL definitions the IXI Workflow Service can determine the compatibility between stages. Where a type mismatch occurs the IXI Workflow Service has two options: it can search for a conversion service (a specific type of IXI Core Service) that is capable of casting between the two types and transparently insert this service into the workflow or failing this the workflow has to be rejected and the user informed of the nature of the incompatibility.

Since each workflow has the potential to execute for hours or even days – the ability the dynamically modify a workflow, mid execution is highly desirable. For example users may wish to correct an erroneous input parameter before it propagates further down the pipeline or visually inspect outputs at certain stages before continuing with the workflow. The IXI Workflow Service therefore permits the user to interact or ‘steer’ submitted workflows. Users may:

- Add breakpoints, where upon reaching such a breakpoint the IXI Workflow Service will wait for a (user) confirmation to continue.
- Modify a workflow either by changing existing parameter values or specifying new parameters. The IXI Workflow Service will then determine how the pipeline was affected by this change and recompute any stages as necessary.
- Terminate execution of the workflow.

We have developed a flexible web based portal that permits end users to create custom workflows, browse existing workflows and also submit workflows to the IXI Workflow Service for execution. Full details can be found in [4].

2.3 IXI Data Services

Execution of each workflow inevitably results in the production of not only the final results, but also of a number of intermediate files that may be of value to the end user. These files have numerous uses, for example they may be used to save computation time when executing a workflow or may constitute part of a data provenance system. In the course of a normal execution the various results produced will tend to be scattered amongst several distributed GridFTP servers – manual handling of this data by the end user is cumbersome so to effectively manage this workflow output, we provide a data management service, the IXI Insertion Service. The IXI Insertion Service interacts with the IXI Workflow Service via the OGSI notification mechanism. Upon successful completion of a workflow a notification is published containing a list of URLs for all results/files produced during the lifetime of the workflow plus details of how these results relate to the original workflow. The IXI Insertion Service acts upon this notification and determines an appropriate mapping (from GridFTP url to database insert). The user is then directed to a web front end, where the final selection on which results should persist in the database is made. Finally the IXI Insertion service issues the necessary insert commands to the relevant database.

Conclusion

The development of high-level Grid services is essential if Grid computing is to attain widespread use in the wider medical community. We propose a framework facilitating the exposure of legacy image processing algorithms as Grid services, the ability to compose these services into workflows and a means to connect these workflows to data sources in a flexible, secure and co-ordinated way across a distributed set of resources. Whilst much research remains to be done, we have successfully developed a working prototype and integrated this with the algorithms and data of four different medical imaging departments / institutions. Early results are promising and indeed indicate that Grid computing can add real value to many areas of medical imaging such as drug discovery and decision support in healthcare.

Acknowledgements

The group would like to acknowledge the assistance of the UK e-Science program for their assistance in funding the *Information eXtraction from Images* (IXI) project of which this work forms a part. We would also like to thank GlaxoSmithKline, Philips Medical Systems and the Dunhill Charitable Trust for their additional support and funding.

References

[1] Non-rigid registration using freeform deformations: application to breast MR images. Rueckert D, Sonoda LI, Hayes C, Hill DLG, Leach MO, Hawkes DJ. IEEE Trans. Medical Imaging 1999; 18(8):712-721

[2] Fast robust automated brain extraction. Smith S. Human Brain Mapping 2002, 17(3): 143-155.

[3] Medical Image Registration. Joseph Hajnal, David Hawkes and Derek Hill.

[4] Image Processing Workflows Using a Grid Enabled Database. A.L. Rowland, M. Burns, T. Hartkens, J.V. Hajnal, D. Rueckert, D.L.G. Hill. Submitted to Medical Imaging Computing and Computer Assisted Intervention (MICCAI) 2004.