IOME, A Toolkit for Distributed and Collaborative Computational Science and Engineering

M.K.Griffiths[1], C.C.Reyes[2], D.Savas[1], and T.Greenfield[3]

1 Corporate Information and Computing Services, The University of Sheffield, 285 Glossop Rd., Sheffield, S10 2HB, U.K.

2 Cancer Research UK Tumour Microcirculation Group, Department of Oncology, School of Medicine & Biomedical Sciences, The University of Sheffield, Beech Hill Road, Sheffield S10 2RX, U.K.

3 ISRU, School of Maths and Stats, Herschel Building, Newcastle University, Newcastle upon Tyne, NE1 7RU

1. Introduction

The internet provides a media rich communications platform enabling communities to share content. This paper presents a toolkit enabling researchers to easily develop novel mechanisms of information sharing. Two case studies are presented, one illustrating a web site offering medical professionals image analysis tools and another providing a collaborative visualisation of a tsunami simulation based on a shallow water wave model. Researchers developing computational models require development kits and tools enabling them to provide simulations with a range of methods that facilitate collaboration. Alongside the increased activity in collaborative work, recent developments on workflow tools are now enabling researchers from different disciplines to collaborate in a unique way leading to new knowledge discovery mechanisms. This occurs by the way in which data is shared and results generated for example between large multi-disciplinary, optimization problems. This paper presents a novel, multi-purpose toolkit and protocol enabling researchers to easily develop simulations which can be run as web services, accessed interactively and enabled to annotate data sets automatically. The simulations can be controlled by client applications such as visualisation tools, web accessible portlets, popular web API tools and other bespoke clients. By using the IOME toolkit scientists and engineers can develop collaborative computational models without requiring an in depth knowledge of the web service protocols. The IOME toolkit has been developed for a range of popular modelling tools including Matlab, Scilab, python, php, C++ and FORTRAN. IOME provides researchers with a range of benefits by enabling

- data sharing between heterogeneous applications and platforms which can be running within different administrative domains,
- publication of computational models as web services and provides tools for building clients,
- automated generation of metadata which may be used for managing distributed data collections and for automated laboratory notebook generation,
- applications which use coarse grained parallelism and exploit multi-core architectures.

The next section provides an overview of the architecture of the IOME toolkit. Section 3 provides example applications demonstrating how the IOME toolkit is used in practice. We demonstrate the model as a web service use case, using as an example an application from the on line algorithm repository for cancer image analysis. We also demonstrate a collaborative and computational steering example of a study of tsunamis using a shallow water wave propagation model with Scilab or Matlab. Section 4 gives an outline of application development using IOME. In section 5 we provide conclusions and review some of the further developments that would be required in a collaborative research computation facility.

2. Design of the IOME Toolkit

Ideally, it would be desirable to make use of high performance distributed computing standards such as OpenMP, MPI and PVM. However these are not designed for developing applications that run over administrative domains, thus only the service oriented architectures such as web services enable us to handle this complexity. To facilitate collaborative working, it is necessary to communicate information about a simulation between resources and collaborating researchers. IOME achieves this by providing a system for efficiently managing a collection of data. The IOME server contains the data collection whilst the clients make requests to modify the data store. A range of operations are available for adding and removing data items from the store. There are also operations for inspecting, modifying and listing the contents of the data store. IOME uses a protocol based on an XML markup called IOME ML, this is the Interactive Object Management Environment Markup Language. The paper describes a range of tools are for developing web service clients using the IOME toolkit. This section explores the architecture in more detail before examining the way IOME is used.

3. Example Simulations

3.1 An on-line algorithm repository for Cancer Image Analysis

The need for image analysis is ever growing in many fields and cancer diagnostics is not an exception. With the advent of new imaging techniques such as intravital, confocal and multiphoton microscopy, just to mention a few, researchers can visualise physiological and pharmacological processes together with the traditional anatomical images, such as Computed Tomography (CT) or Magnetic Resonance Imaging (MRI), which have been improving resolution. In some cases, academic collaboration between areas like mathematics, computer science and physics with clinicians and biologists has lead to development of mathematical models that describe biological processes, but in many cases, groups work isolated in an isolated fashion and there is a lack of communication between interested parties.

CAIMAN (CAncer IMage ANalysis) is an Image Analysis internet-based project that combines the strength of open-source web-based scripting languages, the powerful high-level technical computing language MATLAB, and the vast literature on image analysis and computer vision to provide a user-friendly web-page where any person can upload cancer-related images and execute analysis algorithms and obtain quantitative measurements related to their images. The user will select the appropriate webpage and algorithm to apply to the images at the front end and will proceed to upload the image (fig. 2).

Once the images have been uploaded to the web server , the IOME php toolbox is used to make a web service request to the image analysis service this request is made using IOME-ML. The process is summarised in Fig. 3. Details of the methodology are discussed in the full paper.

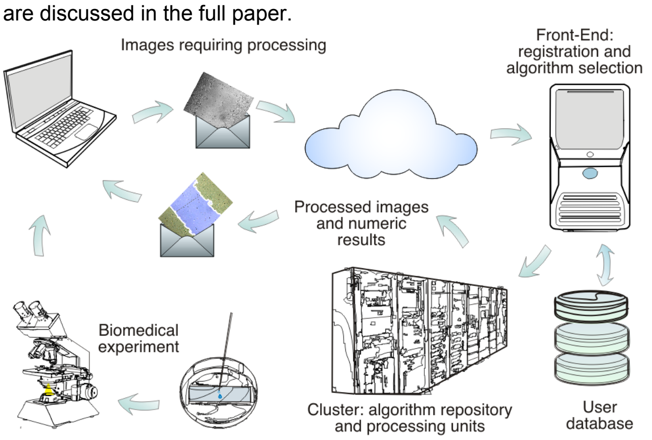

Fig 3 Graphical description of CAIMAN

3.2 Collaborative Visualisation and Computational Steering

Using the IOME tool kit we have developed an application enabling a researcher to develop a steered simulation or collaborative visualisation using the popular visualisation tool IBM data explorer. One of the reasons for IBM data explorer's popularity is because of its visual programming interface enabling researchers to rapidly develop applications by dragging and dropping components from a toolbox and connecting those components together. The IOME visualisation tool for IBM data explorer uses the interface tool-kit provided with data explorer to enable data explorer applications to move data between a custom application and a visual program. We describe the development of a collaborative visualisation of a model used for simulating a tsunami using the shallow water equations. Most sophisticated methods make use of adaptive mesh techniques here we present a simplified approach using the Crank-Nicholson method to solve the shallow water equations. The steps to develop our collaborative simulation are described in the full paper.

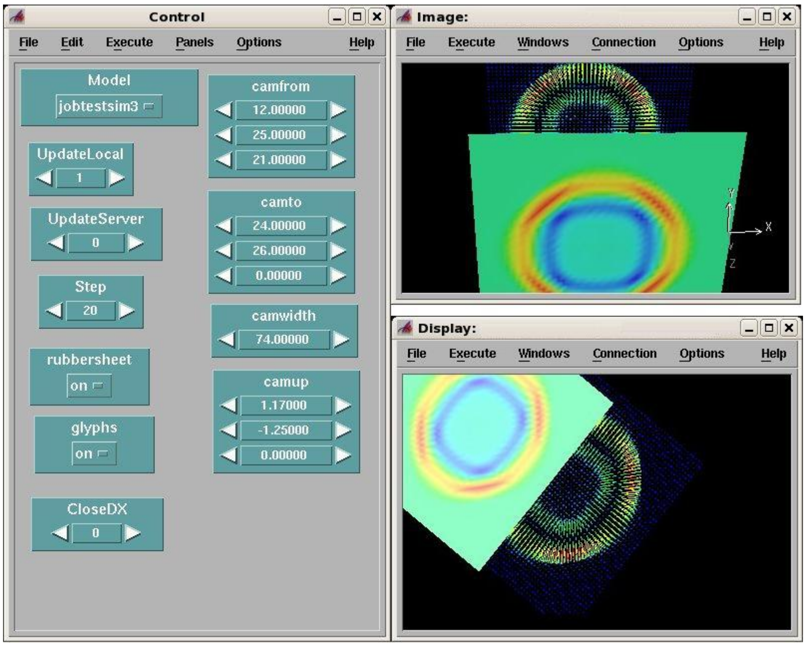

Figure 7. Collaborative Visualisation of a Tsunami Simulation Using IBM Open Data Explorer.

4. Application Development Using IOME

This section will provide further detail concerning application development using the IOME toolkit.

5. Conclusion and Future Developments

The examples presented here demonstrate that researchers using IOME can easily develop modelling, visualisation and analysis applications that are able to communicate useful data and results between geographically separated researchers. It has also been demonstrated that researchers can develop practical web interfaces enabling colleagues to run simulations. The main benefit is that researchers can set up these models without a detailed understanding of web services and protocols such as SOAP. These techniques can be built into diverse applications, running across heterogeneous platforms. It has been seen that the IOME tool-kit provides the benefits listed below ;

- allows data sharing between client applications running in different locations for collaborative working or in a computational steering scenario,
- allows researchers to publish a simulation as a web accessible service,
- enables the automated generation of simulation metadata which may be used to manage distributed data collections.

The current implementation of IOME assumes a service provision model for which the services are static and stateless. In a service model with multiple applications a more appropriate approach is to make use of a stateful service model we discuss future implementations of IOME which would enable the creation of stateful services.