A Framework for Providing Research Applications as a Service Using the IOME Toolkit

M.K.Griffiths[1], C.C.Reyes[2] and J.M.Rodenberg[3]

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

2 The Biomedical Engineering Group, Richmond Building, School of Engineering and Informatics, University of Sussex, Falmer, Brighton, East Sussex, BN1 9QT, U.K.

3 Pi-Phi Project, Kroto Institute, North Campus, Broad Lane , The University of Sheffield, Beech Hill Road, Sheffield S3 7HQ, U.K.

1st August 2012

Abstract

This paper presents a unique, multi-purpose toolkit, enabling researchers to easily develop modelling and analysis applications, which can be run as web services and accessed interactively. The development kit is based on a protocol that uses an XML markup called the "Interactive Object Management Environment Markup Language" (IOME ML). The paper describes the IOME ML and its development kit.

We illustrate the capabilities of IOME with two case studies the first case study is based on a medical image processing application (CAIMAN: CAncer IMage ANalysis), offering image analysis tools for life scientists. For the second case study, the Pi-Phi collaboration have developed an inverse imaging method for 'lensless' microscopy a demonstrator is introduced for the Pi-Phi project. For both case studies the application is wrapped as a web service and accessed through a web browser.

The paper concludes with a review of further developments, including refinements to the mark up language and the development of a service factory, enabling a more scalable service provision model through the dynamic invocation of published simulations as IOME web service applications.

Introduction

Researchers developing computational models would benefit considerably from development kits and tools that would enable them to provide simulations with a range of methods that facilitate collaboration. Many researchers are very good at programming for example using C/C++, Matlab or Fortran and implement algorithms related to their disciplines. However, these researchers may find it much more challenging to develop distributed applications. This is partly due to the extra knowledge of communications protocols that is required to develop distributed applications and other the knowledge of other environments such as those related with web-oriented applications. Distributed applications development tools such as *Message Passing Interface* (MPI) or *Parallel Virtual Machine* (PVM) are popular amongst the scientific community but these can be difficult to use for developing distributed heterogeneous applications running across administrative domains. MPI and PVM belong to the same class of "Single Instruction-Multiple Data" model of computing where the computing task is shared by the coordinated distribution and execution of the same code on multiple servers. Such an approach can not lend itself well to developing heterogeneous independent applications that can communicate across administrative domains. Although PVM was developed with heterogeneity in mind, its emphasis is on the coordination of multiple workers and hence the approach addresses a different set of problems to that presented here. IOME can be viewed as a set of tools for developing potentially "Multiple Instructions-Multiple Data" computing models that can span administrative boundaries. The challenging task of implementing a reliable method of communications between heterogeneous and independent applications running across different software&hardware platforms and administrative domains has been implemented via the adoption of the Simple Object Access Protocol "SOAP" standards. SOAP is used as the carrier of all IOME communication messages. This binds

IOME to soap standards and allows future developments based on these standards. By exploiting a mapping between web services and MPI, it is possible to alleviate some of the difficulties in distributed application development for many researchers [1].

The proposed presentation uses two case studies to provide an outline of application development using IOME. We present an application demonstrating how the IOME toolkit is used in practice with a web site offering specific image analysis algorithms for the life sciences. The framework is applied to a further project demonstrator. In the final section we conclude with a summary of our findings and we present further developments required for the IOME toolkit.

We have investigated a diverse range of collaborative techniques, including; workflow development tools such as the *Taverna* engine [2] or computational steering toolkits such as *Gviz* or the reality grid [3, 4]. Taverna provides an approach for consuming a service whereas IOME provides a method for making an application available as a service and provides the tools for developing suitable clients for a given user community. Although some web services are simple to implement e.g. using AXIS [5] one essentially drops web service onto a container. We have also found that web service wrapping tools such as CARMEN and OPAL [6] require significant expertise and knowledge of java accessible web services. For many of these systems there is a rich number of use cases and this leads to greater system complexity making it more difficult for users with less specialised middleware knowledge to implement a solution. IOME distinguishes itself by mediating the gap between service provision frameworks and message passing frameworks for high end computing such as PVM/MPI. In contrast IOME provides a set of web service based tools enabling researchers to tailor build and customise service structures to the needs of a specific problem. Our experience has demonstrated that in many cases the complexity of the toolkit, the accessibility to it, and the distributed nature of the problem add to the challenge for researchers to develop a collaborative solution in a timely fashion. Our work with a community of research computing service users indicates that there is a need for a simplified method of communication between a range of client applications running on different platforms. Using IOME, this requirement can be met by providing a general-purpose description language for modelling and analysis problems. We have named the resulting XML based mark up language IOME-ML.

IOME is a toolkit that enables researchers to use a Software as a Service (SAAS) model to enable access to research applications which may exploit resources provided through cloud computing. Since the IOME server is a standalone self contained web server it is easier to administer and maintain than solutions implemented using taverna/condor/AXIS. The development kit based on a protocol that uses an XML markup called the "Interactive Object Management Environment Markup Language" (IOME ML). 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, researchers can develop collaborative research tools without requiring an in depth knowledge of the web service protocols. The IOME toolkit has been developed for a range of popular development tools including Matlab, Scilab, python, php, C++ and FORTRAN.

The CAIMAN and Pi-Phi Case Studies

CAIMAN (CAncer IMage ANalysis) [7,8] is an Internet-based Image Analysis project where images produced by several cancer-related experiments (scratch-wound assays, vasculature tracing, immunohistochemstry) are processed automatically and returned to users from Life Sciences without the need to do any programming. Since it was launched, CAIMAN has processed more than 1,500 images; users in several countries have processed different images. Some users have processed more than 100 images each, which denotes the usefulness of CAIMAN for their research. For the CAIMAN model it was realised that what was required was a single user interface for accessing a particular service. The front-end of CAIMAN is a PHP-based website where the user can access the current image analysis algorithms. 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. The actual components participating during the submission of an image processing task are shown in Fig. 1.

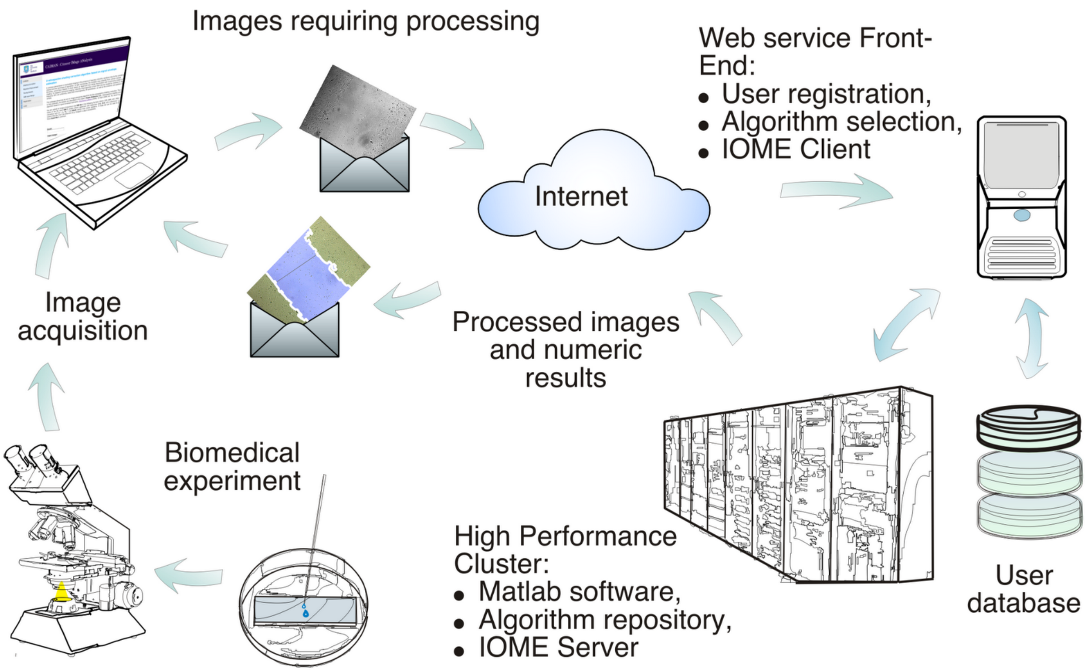

Fig. 1 Graphical description of CAIMAN

The CAIMAN case study presented here demonstrate that researchers using IOME can easily develop modelling and analysis applications that are able to communicate useful data and results between geographically separated researchers. It has been shown that researchers can develop practical web interfaces, which allow collaboration partners to run simulation and analysis applications. The main benefit is that researchers can set up these models without the need for a detailed understanding of web services and protocols such as SOAP.

The demonstrator developed for the pi-phi project is an adaptation of the framework used for the CAIMAN project. The pi-phi project relates to an inverse imaging method for 'lensless' transmission microscopy, applicable to all wavelengths (light, X-ray and high-energy electrons). This has been developed by a consortium of 7 UK Universities (the Pi-Phi project, see http://www.pi-phi.org). The on-line demonstrator shows how objects can be imaged by detecting scattered intensity alone (no lenses). The waves scattered by the object are computationally reconstructed using a very efficient phase-retrieval algorithm [9].

Conclusion

IOME provides researchers with a range of benefits by enabling:

- publication of simulation and analysis applications as web services it also provides tools for building clients,
- data sharing between heterogeneous applications and platforms, which can be running within different administrative domains,
- automated generation of metadata, which may be used for managing distributed data collections and for automated laboratory notebook generation,

The CAIMAN a Pi-Phi case studies presented here demonstrate that researchers using IOME can easily develop modelling and analysis applications that are able to communicate useful data and results between geographically separated researchers. It has been shown that researchers can develop practical web interfaces, which allow collaboration partners to run simulation and analysis applications. The main benefit is that researchers can set up these models without the need for a detailed understanding of web services and protocols such as SOAP.

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. In this model we have a continuously running service factory and registry that can start an instance of an IOME service. It is therefore necessary to provide a system enabling the dynamic invocation of web services and a registry service providing information about the currently available services [11,12].

References

1. Puppin, D., N. Tonellotto, and D. Laforenza, *How to run scientific applications over Web services.* 2005 International Conference on Parallel Processing Workshops, Proceedings, 2005: p. 29-33
2. Hull, D., et al., *Taverna: a tool for building and running workflows of services.* Nucleic Acids Research, 2006. **34**: p. W729-W732.
3. Brodlie, K.W., et al., *Distributed and collaborative visualization.* Computer Graphics Forum, 2004. **23**(2): p. 223-251.
4. Brooke, J.M., *Computational Steering in RealityGrid*, in *e-Science All Hands Meeting*, S.J. Cox, Editor. 2003: Nottingham.
5. Avar, A. *The AXIS Project*. 2003 [cited; Available from: http://ws.apache.org/axis/.
6. Krishnan, S., *Opal: SimpleWeb Services Wrappers for Scientific Applications*, in *IEEE International Conference on Web Services (ICWS'06)*. 2006.
7. Reyes-Aldasoro, C.C., et al., *CAIMAN: An online algorithm repository for Cancer Image Analysis.* Comput. Meth. Prog. Bio., 2010. **in press, doi:10.1016/j.cmpb.2010.07.007**.
8. Reyes-Aldasoro, C.C., et al., *Measuring cellular migration with image processing.* Electron. Lett., 2008. **44**(13): p. 791-793.
9. Maiden AM and Rodenburg JM, Ultramicroscopy **109** (2009) 1256-1262
10. van Engelen, R.A. and K.A. Gallivan, *The gSOAP toolkit for web services and peer-to-peer computing networks.* Ccgrid 2002: 2nd IEEE/Acm International Symposium on Cluster Computing and the Grid, Proceedings, 2002: p. 128-135.
11. Watson, P., et al., *Dynamically deploying Web Services on a Grid using Dynasoar.* Ninth IEEE International Symposium on Object and Component-Oriented Real-Time Distributed Computing, Proceedings, 2006: p. 151-158.
12. Moreau, L. *GRIMOIRES, a UDDI-compatible web service registry with added capabilities*. 2008 [cited; Available from: http://www.grimoires.org/.