Towards a Collaborative Orthopaedic Research Environment

Y. W. Sim, C. Wang, L. Gilbert, G. B. Wills
Electronics and Computer Science, University of Southampton, UK
{yws01, cw2, lg3, gbw}@ecs.soton.ac.uk

Abstract

Until recently, it has been the case that scientific findings are captured, summarized, and shared through manuscripts. Nevertheless, the practices of science have already been affected dramatically by Web technologies. Virtual learning and research environments have typically been implemented as monolithic systems in the past. Collaborative Orthopaedic Research Environment (CORE) is a Virtual Research Environment (VRE) which is designed based on the concept of Service Oriented Architecture (SOA) and Grid/Web services. It aims to provide an infrastructure that combines clinical, educational and research in one working environment. This paper describes the VRE requirements and discusses the advantages that will be achieved by implementing it as Grid/Web services.

1. Introduction

The research community, dynamic and technology driven, shares its information through approaches initiated with Gutenberg's printing press and conceptually recognizable to scientists in the 18th century [6]. Until recently, it has been the case that scientific findings are captured, summarized, and shared through manuscripts. Nevertheless, the practices of science have already been affected dramatically by Web technologies and, in particularly, by the Internet. For example, the genome sequence in gigabytes magnitude available online [2] means that for biologists, data is something that they can find on the Web, not only in the lab. The rapidly expanding computing and storage capabilities of federated Grid, and advances in optical networks are accelerating the trends of disseminating and sharing scientific findings over the Web.

The Collaborative Orthopaedic Research Environment (CORE) [7] is a JISC funded project, which aims to provide an infrastructure that combines clinical, educational and research in one working environment. The current paradigm of information sharing and resource use in biology and medicine are being challenged on several fronts. Firstly, as the number of investigators, organizations and institutions conducting biomedical research increases, it becomes difficult to track the work and provide infrastructure to support these expansions. Although current information technology supports ready access, it does not address abstraction, integration and interpretation of information. The diverse bio-informatics tools generated to consume and evaluate the data rarely interoperate [4]. Secondly, the very large volume of data generated in modern biomedicine presents a primary challenge to the researcher. To integrate biological data one would want to move seamlessly between biologic and chemical process, organelle, cell, organ, organ system, individual, family, community and population. Such integration generates challenges to information structure as each research community tends to speak its own scientific dialect [9]. Finally, biomedicine's culture is at the nexus of challenge faced by many other scientific fields: the need for collaborative research. The collaborative researchers recognize that many of the technology approaches required in biology and medicine are expensive, beyond the reach of individual investigators, and increasingly challenging the resources reserved of all but a few institutions.

2. The CORE approach

Regarding to the challenges discussed above, the CORE project aims to address them by implementing a Virtual Research Environment (VRE) demonstrator using Service Oriented Architecture (SOA) concept. The CORE is a follow-up project to the EU-funded Virtual University for Orthopaedics (VOEU) [15], which has two main functions: to support the educational process and to aid surgeons in preparing findings for publication. Orthopaedic surgical trials typically run for extended periods (up to two years), with postoperative assessment results being collected regularly. The collated results are then analyzed and discussed by a team of surgeons before being disseminated to the wider orthopaedic community. A detail description of the activities and workflow supported within the VOEU can be found in [5] and [8].

The VRE in VOEU is based on an integrated and tightly coupled architecture, making it difficult to expand as the user requirements change. The Collaborative Orthopaedic Research Environment (CORE) is a VRE, which takes the foundations established in VOEU and decomposes them into modules. These modules will then be developed using Grid/Web services technologies and supported by a SOA, where flexible granular functional components expose service behaviors accessible to other applications via loosely coupled standards-based interfaces. The term, SOA, refers to systems structured as networks of loosely coupled, communicating services [3]. It is a style of design that guides all aspects of creating and using services through their lifecycle (from conception to retirement), as well as defining and providing the information infrastructure that allows different applications to exchange data regardless of the operating systems or programming languages underlying those applications.

Grid services were originally conceived as a method to support virtual communities through sharing of computational and data resources. Access and identity control are fundamental components of the Grid technology, which supports deterministic queries across a distributed, common schema. The technology also supports stateful processes important to the concept of workflow. The developing Open Grid Service Architecture-Data Access Integration framework holds promise for adding semantics to the Grid technology so that computable, semantic interoperability may be achieved.

Grid technology does have its limitations. Despite its developing research maturity [10], Grid is a distant second in commercial application. Web services technology is the preferred choice for the vast majority of information infrastructure support installations, in part because of the greater relative simplicity of the technology. Web services extend the concept of SOA into a vast networking platform that allows the publication, deployment, and discovery of service applications on the Internet scale. It is a straight forward extension of Internet and Web infrastructure familiar to the majority of systems designers and administrators.

However, the security and state awareness of Grid technology are essential for VRE applications. Hence, the CORE will attempt to combine the Grid and Web services technologies under one architecture, i.e. SOA, in order to realize the benefits attached to both technologies.

Wilson *et al.* [16] discuss in detail the advantages of using SOA, but in the CORE project's context the following particularly apply:

- **Modularity**: Due to the nature of loose coupling in SOA, applications can be developed and deployed incrementally. Often, a reasonable subset of the full functionality can be developed quickly, which has obvious time-to-deployment advantages. Additional functionality can readily be added in planned stages until the full feature set has been realised.
- **Interoperability**: As the services specifications defined by standard bodies, such as W3C [2] and OASIS [13], progress toward standardization, it becomes increasingly easy to incorporate services from third party into the VRE as required.
- **Extensibility**: Some developers will find the core services specification sufficient, while other may not. Hence, the core specifications were defined with built-in extensibility points such as security and reliability of services. Such feature will steer the developers away from the danger of technology lock-in.

3. The user requirements

The focus of the generic toolkit components in CORE will be a Virtual Bone Biology Laboratory, incorporating the basic science disciplines of molecular and tissue biology, engineering and computing with allied medical disciplines of pharmacology, prosthetics, trials management and the clinical disciplines concentrating upon musculoskeletal applications in rheumatology and orthopaedics.

The Bone Laboratory aims are to prepare the processes and test methods for the development of new technologies for the information management, and construction of technologies on the nanometric scale. The Core will need to provide Grid services for the simulation and modeling of bone and soft tissue biology, analysis of large scale experiments and the modeling of nanometric tissue units.

In addition, the Core Bone laboratory will provide a post-graduate training experience for researchers and clinicians with a background in bioengineering initially treating musculoskeletal problems with emphasis upon training the next generation of clinicians and scientists. The use of the Grid for distributed computation means that powerful analysis and modelling tools can be made available to staff and students. But there are reasons beyond saleability that Grid services make sense in such learning scenarios. Learning is more than a series of quick transactions and needs a mechanism to orchestrate the services, manage state, and provide security and trust across a range of services.

In order to design a VRE with enhanced features that will tailor to the end users' needs, the authors have conducted a user requirement study with various experts and users (including an investigator working in Bone Biology Laboratory) [12]. The results from the study have reinforced the necessity of developing a Grid/Web serviced based research environment, which is equipped with the ability to evolve with the changing requirements of its user community. For example, there are vast numbers of protocols to choose from when conducting a trial. Although it might be possible to make a small number of generic trial protocols or templates; however, it is unreasonable to predefine and encode every possible experimental procedure in most occasions. Likewise, it is also difficult to predict every type of data that can be collated and analyzed, and every possible dissemination route that users will follow to publish their results to the wider community. Therefore a loosely coupled architecture is essential to allow the flexibility of adding extra services at a later stage when the user requirements change.

A clear priority requirement in the study is the design of a user-oriented portal, aimed at the non-computer specialist, with accessible resources, i.e. scientific data and publications, that are easy to browse, upload and download. In view of this, portal technology is incorporated into the design of the CORE VRE. According to JISC FAQ, a portal is described as "*a network service that brings together content from diverse distributed resources using technologies such as cross searching, harvesting, and alerting, and collate this into an amalgamated form for presentation to the user*" [11]. A portal can provide personalization to users and allow them to access the VRE using Web browsers. Such features make the VRE easy to use for users who may not be computer literate.

Another important finding in the user requirement study is that different types of users would require data to be presented differently. For instance, biologists are generally interested in graphical representations of data while mathematicians might prefer numerical interpretation of the same set of data.

The authors learned from the study that the users in Bone Biology Laboratory need the infrastructure to run simulations and analyze large scale trial data. Hence, it makes sense to include Grid services in the CORE VRE since the services can provide secure and managed access to distributed computational power.

4. The CORE infrastructure

The CORE VRE is being implemented as a toolkit of generic components. Figure 1 illustrates the CORE framework concept, in the context of an e-science community Web site and integrated Grid/Web-based services.

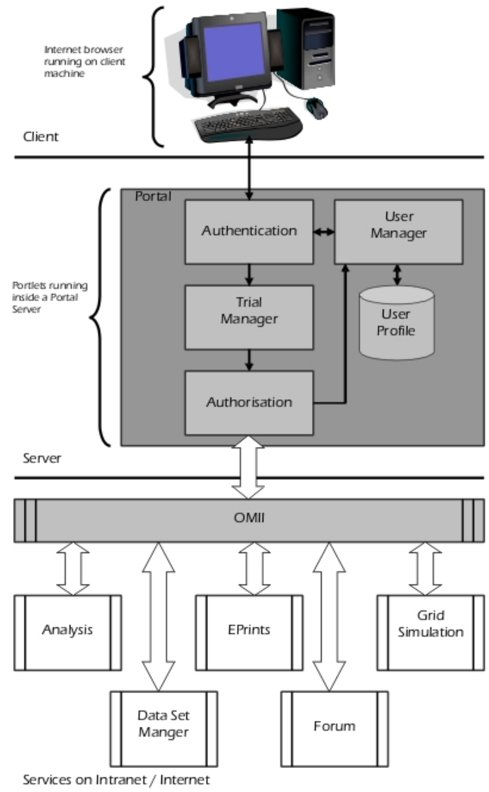

Figure 1: Overview of the VRE infrastructure

OMII is a collection of tested, documented and integrated software components that provides a standard platform for integrating e-Science middleware. It also acts as a simple, secure web service-based Grid infrastructure for e-Scientists [14]. Hence, by using the OMII middleware, end users of the CORE toolkit can access Grid resources and applications in a trusted and secure environment.

Four Web services are factorised based on the functionalities provided by the VOEU toolkit. The authors first decomposed the toolkit into processes and then identified which ones are potentially to be reused in other VREs. These reusable processes are *Analysis*, *Data Set Manager*, *EPrints* and *Forum*, which are the

Web services depicted in Figure 1. These services will enable users to perform tasks such as formalising trial protocol, storing and analysing data, submitting and reviewing articles, and discussing research findings in a forum. Another Grid service, named *Grid Simulation* is also included in the VRE infrastructure. It will provide users with functionalities, such as job submission, file transfer, and credential management, in running their simulations.

A portal is used to facilitate the sharing of research resources in the CORE VRE. The main purpose of a portal in this infrastructure is to act as a presentation layer which aggregates, integrates, personalises and presents information, transactions and applications to the user according to their role and preferences. It provides a persistent state for individual user or a group of collaborators. In the context of personalization and embedding, portals can achieve this through creating distinct pluggable components of functionality and offering them to the users as visible components. Each of the components can be constructed by utilizing a new emerging technology named Portlet. JSR-168 is the Java Portlet Specification adopted by majority of Portal vendors in the market [1]. The specification aims to standardize the pluggable portal components so that they are independent of the portal vendor solution. Hence, the authors will design and develop the components of the CORE portal based on the JSR-168 Specification.

5. Conclusions and future work

The design of the CORE project's VRE described in this paper is underpinned by Grid/Web services technology based on the concept of SOA. The SOA theme enables reuse via shared services, where flexible granular functional components expose service behaviors accessible to other applications via loosely coupled standards-based interfaces. There is another common way of integrating systems, which is to integrate at the user interface level using portals. However, the SOA approach does not preclude using portals, and is in fact agnostic about how the rest of the applications in an organization are configured. Thus, SOA is a good approach for constructing framework.

The CORE VRE supports sharing and dissemination of findings, i.e. scientific data and publications, in research activities. This is important as any scientific effort will have diachronic features, i.e. collaborative research activities may extend through time, enabling the influence of these activities to carry on beyond their timescale and disseminating research findings beyond the original boundaries of collaboration. The CORE project should have a major impact on a number of areas including:

- Being able to keep track of the research administration: trial protocol, ethical approval, and workflow as the trial progresses,
- Enabling access to research data from various trials and in formats that allow analysis of the data,
- Allowing easier meta-analysis or thematic reviews,
- Monitoring the effectiveness of surgical interventions,
- Enabling a consortium to write appropriate documents for dissemination (medical reports, journal articles, etc)
- Producing up to date learning and teaching material.

The CORE project will build a VRE demonstrator based on the CORE infrastructure presented in this paper. Issue such as semantic Grid/Web services will also be investigated in order to enhance the functionalities of the proposed VRE, particularly in areas of automation and organization. For example, it may be possible to automate a feasibility check of possible trials by looking at the availability of resources, i.e. laboratory timetable, datasets and staff. In addition, evaluation of the demonstrator will be conducted qualitatively (using workshops and focus groups) and quantitatively (using experiments) once it is implemented. The evaluation will focus on the usability aspects of the demonstrator (e.g., can it be used simply and effectively?).

6. Acknowledgements

The CORE project is funded by the Joint Information Systems Committee under the Virtual Research Environment program.

7. References

[1] R. Allan, C. Awre, M. Baker, and A. Fish, "Portals and Portlets 2003", National e-Science Centre, 2004, available from: http://www.nesc.ac.uk/technical_papers/UKeS-2004-06.pdf

[2] A. Bairoch, and R. Apweiler, "The SWISS-PORT Protein Sequence Database and its Supplement TrEMBL in 2000", *Nucleic Acids Research*, 28(1), 2000, pp. 45-48.

[3] D. Booth, M. Champion, C. Ferris, F. McCabe, E. Newcomer, and D. Orchard, "Web Services Architecture (W3C Working Group Note 11 February 2004)", World

Wide Web Consortium, 2004, available from: http://www.w3.org/TR/ws-arch/

[4] K. Buetow, “Cyberinfrastructure: Empowering a ‘Third Way’ in Biomedical Research”, *Science*, 308(5723), May 2005, pp. 821-824.

[5] L. Carr, T. Miles-Board, G. Wills, G. Power, C. Bailey, W. Hall, and S. Grange, “Evolving a Digital Library Environment to the Changing Needs of its Users”, In *Proceedings of Healthcare Digital Libraries Workshop 2004 (HDL 2004)*, Bath, UK, 2004.

[6] M. Castells, *The Internet Galaxy: Reflections on the Internet, Business, and Society*, Oxford University Press, Inc., New York, USA, 2001.

[7] CORE Project Proposal, University of Southampton, available from: http://www.core.ecs.soton.ac.uk/overview/

[8] S. Grange, G. Wills, G. Power, T. Miles Board, L. Carr, and W. Hall, “Building a dynamic review journal (DRJ) - Extending the role of the Virtual Orthopaedic University”, In *Proceedings of the 3rd Annual Meeting of the International Society for Computer Assisted Surgery (CAOS International)*, Marbella, Spain, June 2003, pp. 122-123.

[9] T. Hey, and A. Trefethen, “Cyberinfrastructure for e-Science”, *Science*, 308(5723), May 2005, pp. 817-821.

[10] The LHC Computing Grid Web site: http://lcg.web.cern.ch/LCG/

[11] “Portals: Frequently Asked Questions”, JISC, available at http://www.jisc.ac.uk/index.cfm?name=ie_portalsfaq

[12] Y.W. Sim, C. Wang, L. Gilberts, G. B. Wills. “User Requirement Study for a Virtual Research Environment”, Technical Report, University of Southampton, available from: http://eprints.ecs.soton.ac.uk/10969/

[13] OASIS SOA Reference Model TC Web site at: http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=soa-rm

[14] OMII Web site: http://www.omii.ac.uk/

[15] Virtual Orthopaedic University Web site: http://www.voeu.ecs.soton.ac.uk/

[16] S. Wilson, K. Blinco and D. Rehak. “Service-Oriented Frameworks: Modelling the infrastructure for the next generation of e-Learning Systems”, Joint Information Systems Committee (JISC), 2004, article available from: http://www.jisc.ac.uk/uploaded_documents/AltilabServiceOrientedFrameworks.pdf