AHE and ACD: A Gateway into the Grid Infrastructure for VPH-Share

David C.W. Chang[1], Stefan J. Zasada[1], Ali N. Haidar[1] and Peter V. Coveney[1]

[1]*University College London, UK*

Correspondence: d.chang@ucl.ac.uk; +44 (0)202 7679 4802; Centre for Computational Science, Department of Chemistry, University College London, WC1H 0AJ, UK

Introduction

With an ever increasing amount of data being exploited by computationally intensive workflows, a secure, flexible and scalable infrastructure is required to cope with modern day biomedical research and simulation landscape. The VPH-Share [1] project is developing such an infrastructure which will provide a range of intelligent services that enable users to expose and share their knowledge and data more easily and efficiently. A major goal of the project is to develop new VPH workflows and encourage collaborations between VPH community members.

VPH-Share is developing services to assist researchers with data annotation, inference and assimilation as well as data reduction and representation. VPH-Share is also developing a cloud computing platform and associated user interface component to store data efficiently and run complex multi-scale workflows using that data. The development of tools and services are driven by four existing flagship projects, namely @neurIST [2], euHEART [3], VPHOP [4] and Virolab [5]. In some of the simulation scenarios considered by the project, the cloud platform will not be able meet the computational requirements of some workflow components, meaning that effort is underway in the VPH-Share project to develop a service that will allow seamless interoperation between the clound infrastructure and grids of high performance computing resources. The Application Hosting Environment [6] (AHE) and Audited Credential Delegation [7] (ACD) fulfil this important task of providing grid access from the cloud platform.

VPH-Share

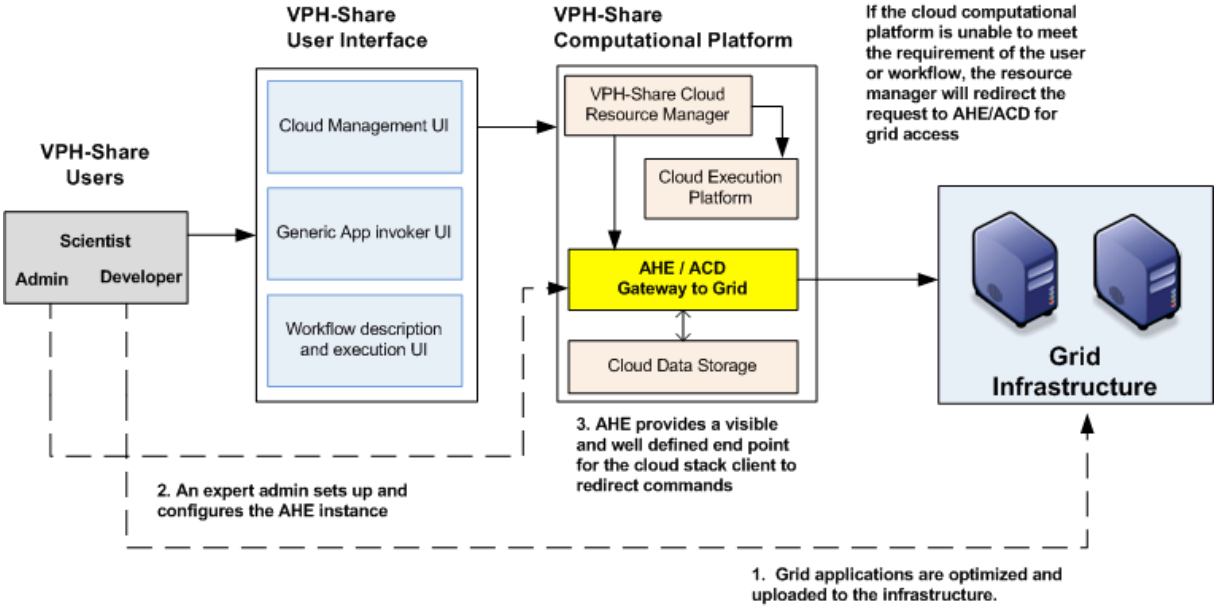

Figure 1: Bridging the gap between grid and cloud computing. A user will access the computational platform through the VPH-Share UI component. The VPH-Share resource manager will then decide which resources are appropriate to execute the user's simulations.

The VPH-Share cloud computing platform consists of several components, shown in figure 1: the Cloud Resource Allocation Management, Cloud Execution Environment, High Performance Execution Environment, Data Access for large Binary Objects, Security framework and Data reliability and integrity service components. All services are registered with the cloud resource allocation management and these services can be invoked as needed. To access the platform, VPH-Share provides a user interface component, which a user can use to execute their workflow. The cloud resource manager will assess the request, and if the Cloud Execution Environment is unable to meet the requirements of a section of the workflow, the Cloud Resource Allocation Management will invoke AHE/ACD to utilize high performance grid infrastructure to execute the computation.

The Application Hosting Environment 3.0 (AHE) and Audited Credential Delegation (ACD) constitute the grid gateway service for VPH-Share, allowing grid infrastructure to be accessed similarly to an Infrastructure as a Service (IaaS) cloud resource. AHE and ACD are deployed based on the Software as a Service (SaaS) model. AHE is responsible for handling the execution lifecycle of virtualized application on the grid, while ACD bridges the gap between different security infrastructures used by the execution platform and the grid resources. The development of AHE and ACD is motivated by the complexity of interacting with the underlying grid infrastructure. Many grid middleware toolkits impose a steep learning curve on the majority of end users, who do not possess the technical expertise to compile, optimize, install, debug and finally launch their application; they simply want to run their jobs and focus on their scientific endeavours. AHE and ACD aim to simplify this experience by providing a lightweight middleware layer between the user and the grid. An expert user configures ACD and AHE once, and the functionalities of the grid hosted applications can then be exposed through a simple and concise set of RESTful web service commands.

AHE

The Application Hosting Environment 3.0 is a re-implementation of AHE 2.0 [6] in Java. AHE 3.0 adds additional features including a workflow engine, a RESTful web service interface, a Hibernate Object Relational Mapping framework and additional enhancements to usability and reliability.

AHE is based on two key concepts, *Application Virtualization* and the *Community Model*. Virtualization in computer science refers to the abstraction of resources, such as an underlying hardware or software platform, from the user. In AHE, application virtualization refers to the abstraction of the application lifecycle from compilation to execution, which is predominantly hidden from the users, with required functionality exposed through a small set of web services. To achieve this, AHE employs the community model user workflow: expert users configure AHE with their domain knowledge concerning the grid platform being used, as well as details of the application to be executed. Once this process is complete, the expert user can share the AHE web service with the end user, allowing them to perform their scientific investigations.

The RESTful web service interface of AHE 3.0 allows the AHE server to expose its functionalities via simple operations on URIs. AHE 3.0 also incorporates a new workflow engine using JBoss's JBPM workflow engine. This allows AHE to model persistent user workflows and provides an easier mechanism to introduce more complex workflows in the future, such as error recovery, or implement additional functionalities such as SPRUCE [8] urgent computing functionalities into AHE.

ACD

Audited Credential Delegation (ACD) is a security management web application. ACD is able to simplify grid security by managing virtual organizations and their membership. In ACD, a virtual organization maps to a group security credential. This credential can be used by each member of the virtual organization to generate a proxy certificate, which will allow the user to have temporary access the grid infrastructure. Access to the credential is granted using security mechanisms more familiar to end users, such as username/password authentication. This mechanism hides the usual certificate

security mechanism from the end user, simplifying the process of accessing the grid or any resource that deploys similar security policies.

ACD functionalities include user authentication and authorization including basic username/password combinations from a local database, OpenID or Shibboleth based authentication. ACD also contains auditing functionality, with all commands issued to ACD being recorded and any proxy certificate generated associated with the requestor, allowing ACD to track all user operations.

Conclusion

As the amount of biomedical data and the complexity of biomedical simulation workflows increase, a secure, flexible and scalable infrastructure is required to support their computational requirements. The VPH-Share project is developing such an infrastructure, which will allow end users to access a range of intelligent services that will simplify the development and deployment of complex multi-scale workflows. These services range from tools to manage data related operations to components to simplify access to computational storage and execution platforms.

AHE and ACD constitute an important cornerstone of the VPH-Share computational platform, providing access to high performance computational resources from the cloud computing platform. ACD and AHE allow the grid to be accessed in a similar manner as an IaaS cloud resource. This is achieved by virtualizing grid applications and exposing their functionality as simple RESTful web services, and by abstracting the security mechanism of the grid middleware through ACD.

AHE also provides a number of additional capabilities including a workflow engine that allows complex grid simulations to be created, including coupled simulations where data is automatically transferred from one application to the other. ACD provides a security suite that includes support for OpenID and Shibboleth authentication, as well as user auditing. ACD support virtual organization management and is able to provide access to grid proxy credentials through RESTful web services.

References

[1] "VPH-Share," http://vph-share.org/.

[2] J. Iavindrasana, L. Lo Iacono, H. Muller *et al.*, “The @neurIST project,” *Stud Health Technol Inform,* vol. 138, pp. 161-4, 2008.

[3] N. Smith, A. de Vecchi, M. McCormick *et al.*, “euHeart: personalized and integrated cardiac care using patient-specific cardiovascular modelling,” *Interface Focus,* vol. 1, no. 3, pp. 349-364, June 6, 2011, 2011.

[4] R. Thiel, K. A. Stroetmann, V. N. Stroetmann *et al.*, “Designing a socio-economic assessment method for integrative biomedical research: the Osteoporotic Virtual Physiological Human project,” *Stud Health Technol Inform,* vol. 150, pp. 876-80, 2009.

[5] P. Sloot, P. V. Coveney, M. T. Bubak *et al.*, “ViroLab: A collaborative decision support system in viral disease treatment,” *Reviews in Antiviral Therapy,* vol. 3, pp. 4 - 7, 2008, 2008.

[6] S. J. Zasada, and P. V. Coveney, “Virtualizing access to scientific applications with the Application Hosting Environment,” *Computer Physics Communications,* vol. 180, pp. 2513-2525, 2009.

[7] A. N. Haidar, S. J. Zasada, P. V. Coveney *et al.*, “Audited Credential Delegation: A Usable Security Solution for the Virtual Physiological Human Toolkit,” *Interface Focus*, 2011.

[8] P. Beckman, S. Nadella, N. Trebon *et al.*, "SPRUCE: A System for Supporting Urgent High-Performance Computing." pp. 295-316.