International Workshop on Science Gateways

Report of Contributions

Contribution ID: **0** Type: **not specified**

Introduction

Monday, 3 June 2013 12:45 (15 minutes)

Presenter: KUNSZT, Peter (ETH ZURICH)

Session Classification: Introduction and Keynote

Contribution ID: **1** Type: **not specified**

Gateway Challenges and Evolution: Current Trends and Practices in US Science Gateways

Monday, 3 June 2013 13:00 (1 hour)

The past decade has shown that web–based access to HPC resources can be a powerful enabler of scientific research. In the aggregate, Science Gateways across the globe have provided access for tens of thousands of researchers and students, and have enabled thousands of publications. Providing open access to such powerful resources is clearly beneficial to the progress of science, but a number of important challenges remain. Many of these challenges are in the technical/logistical realm. For example, how can we create software that better meets the needs of a given user population or makes Gateway construction simpler? However, challenges of equal magnitude remain in the areas of sustainability and public policy. For example, how can important Gateways be funded in the long term, and what users should be granted access to which resources? This talk will survey current US Gateway development strategies in the context of these challenges, and describe the continuing evolution of the CIPRES Science Gateway as a worked example of one group's attempt to manage these challenges.

Presenter: Dr MILLER, Mark (San Diego Supercomputing Center SDSC)

Session Classification: Introduction and Keynote

Contribution ID: **2** Type: **not specified**

Improving the Swiss Grid Proteomics Portal

Monday, 3 June 2013 14:00 (25 minutes)

We have received feedback from our users and supporters on the functionality and usability of the Swiss Grid Proteomics Portal during its first year of operation. We have also realized which aspects of the portal could be improved upon through daily interaction with the production system under heavy use. In a second, highly upgraded version of the Swiss Proteomics Portal, called iPortal, we have introduced several new concepts based on this feedback and both user and supporter experience. In this paper we detail the requirements and the improvements we have made, and also give an outlook on future possible improvements.
Peter Kunszt, Lorenz Blum, Béla Hullár, Emanuel Schmid, Witold Eryk Wolski, Bernd Rinn, Chandrasekhar Ramakrishnan, Andreas Quandt and Lars Malmström, ETH Zürich, Switzerland

Presenter: BLUM, Lorenz (ETH Zurich)

Session Classification: Paper Session 1

Contribution ID: **3** Type: **not specified**

The Urban Research Gateway for Australia: Development of a Federated, Multi-disciplinary Research e-Infrastructure

Monday, 3 June 2013 14:25 (25 minutes)

The $20m Australian Urban Research Infrastructure Network (AURIN) project (www.aurin.org.au) began in July 2010. AURIN is developing a secure, web-based virtual environment (e-Infrastructure) - a lab-in-a-browser - offering access to diverse, distributed and extremely heterogeneous data sets together with an extensive portfolio of targeted analytical and visualization tools. This is being provisioned for Australia-wide urban and built environment researchers –itself a highly heterogeneous collection of research communities with diverse demands. This paper describes these demands and their associated needs and expectations on the e-Infrastructure and illustrates through a range of working examples how the e-Infrastructure allows inter-disciplinary research collaborations to take place. An overview of the e-Infrastructure itself is provided and how it allows tackling these demands.
Richard Sinnott, Christopher Bayliss, Andrew Bromage, Gerson Galang, Guido Grazioli, Philip Greenwood, Angus Macauley, Damien Mannix, Luca Morandini, Marcos Nino-Ruiz, Christopher Pettit, Martin Tomko, Muhammad Sarwar, Robert Stimson, William Voorsluys and Ivo Widjaja

Presenter: SINNOTT, Richard (University of Melbourne)

Session Classification: Paper Session 1

Contribution ID: **4** Type: **not specified**

From the desktop to the grid: conversion of KNIME Workflows to gUSE

Monday, 3 June 2013 14:50 (25 minutes)

The Konstanz Information Miner is a user-friendly graphical workflow designer with a broad user base in industry and academia. Its broad range of embedded tools and its powerful data mining and visualization tools render it ideal for scientific workflows. It is thus used more and more in a broad range of applications. However, the free version typically runs on a desktop computer, restricting users if they want to tap into computing power. The grid and cloud User Support Environment is a free and open source project created for parallelized and distributed systems, but the creation of workflows with the included components has a steeper learning curve.
In this work we suggest an easy to implement solution combining the ease-of-use of the Konstanz Information Miner with the computational power of distributed computing infras- tructures. We present a solution permitting the conversion of workflows between the two platforms. This enables a convenient development, debugging, and maintenance of scientific workflows on the desktop. These workflows can then be deployed on a cloud or grid, thus permitting large-scale computation. To achieve our goals, we relied on a Common Tool Description XML file format which describes the execution of arbitrary programs in a structured and easily readable and parseable way. In order to integrate external programs into we employed the Generic KNIME Nodes extension.
Luis de La Garza, Jens Krüger, Charlotta Schärfe, Marc Röttig, Stephan Aiche, Knut Reinert and Oliver Kohlbacher, University of Tübingen, Germany

Presenter: DE LA GARZA, Luis (University of Tübingen)

Session Classification: Paper Session 1

Contribution ID: **5** Type: **not specified**

Scaling bio-analysis from computational clusters to grids

Monday, 3 June 2013 16:00 (25 minutes)

Life sciences have moved rapidly into 'big data'thanks to new parallel methods for gene expression, genome-wide association, proteomics and whole genome DNA sequencing. The scale of these methods is growing faster than predicted by Moore's law.
This has introduced new challenges and needs for methods for specifying computation protocols for e.g. Next-Generation Sequencing (NGS) and genome-wide association study (GWAS) imputation analyses and running these on a large scale is a complicated task, due to the many steps involved, long runtimes, heterogeneous computational resources and large files. The process becomes error-prone when dealing with hundreds of samples, such as in genomic analysis facilities, if it is performed without an integrated workflow framework and data management system.
From recent projects we learnt that bioinformaticians do not want to invest much time in learning advanced grid or cluster scheduling tools, preferring to concentrate on their analyses, be closer to old-fashion shell scripts that they can fully control and have some automatic mechanisms taking care of all submission and monitoring details.
We present a lightweight workflow declaration and execution system to address these needs, built on top of the MOLGENIS framework for data tracking.
We describe lessons learnt when scaling running NGS and imputation analyses from computational clusters to grids and show application of our solution, in particular, in the nation-wide "Genome of the Netherlands" project (GoNL, 700TB of data and about 200.000 computing hours)
Heorhiy Byelas, Martijn Dijkstra, Pieter Neerincx, Freerk van Dijk, Alexandro Kanterakis and Morris Swertz

Presenter: BYELAS, Heorhiy (University Medical Center Groningen)

Session Classification: Paper Session 2

Contribution ID: **6** Type: **not specified**

A Data-Centric Science Gateway for Computational Neuroscience

Monday, 3 June 2013 16:25 (25 minutes)

Science gateways provide user interfaces and high-level services to access and manage applications and data collections on distributed resources. They facilitate users to perform data analysis on distributed computing infrastructures (DCIs) without getting involved into the technical details. The e-BioInfra Gateway is a science gateway for biomedical data analysis on a national grid infrastructure, which has been successfully adopted for neuroscience research. Necessary improvements in this gateway motivated the design of a new next generation of e-BioInfra Gateway. In this paper we describe the motivation, requirements and design of this new gateway, which is based on the WS-PGRADE/gUSE SG framework, allowing for support for other types of DCIs. The new gateway has additional generic data and meta-data management facilities to access and manage (biomedical) data servers, and to provide an integrated and data-centric user interaction. Its first prototype is implemented and deployed for the computational neuroscience research community of the Academic Medical Center of University of Amsterdam.
Shayan Shahand, Ammar Benabdelkader, Jordi Huguet, Mahdi Jaghouri, Mark Santcroos, Mostapha Al Mourabit, Paul Groot, Matthan Caan, Antoine van Kampen and Silvia Olabarriaga, Amsterdam Medical Center, Netherlands

Presenter: SHAHAND, Shayan (Academic Medical Center of University of Amsterdam)

Session Classification: Paper Session 2

Contribution ID: **7** Type: **not specified**

A GEANT4 Web-based Application to Support Intra-Operative Electron Radio-Therapy using the European Grid Infrastructure

Monday, 3 June 2013 16:50 (25 minutes)

Radiotherapy techniques permit to deliver ionizing radiations (X-rays, photons, electrons, protons, etc) inside cancerous tissues to kill the abnormal cells. Radiotherapy related activities, as the optimization of the therapeutic radiation dose to the patient, worker radioprotection, performance controls and technical innovations of linear accelerators, are strongly based on the ability to predict the dose distribution. Monte Carlo simulations are the most accurate tools in this field but, unfortunately, they require large computing power to achieve accurate results in reasonable times. In the last years, advanced cancer treatment clinical and research communities have adopted e-Infrastructures to reduce this gap. The present paper reports on the developments of a computing facility for helping physicians, radiotherapists and medical physicists in using modern R&E networking and distributed computing and storage resources to address some technical and clinical Intra-Operative Electron Radio-Therapy (IOERT) needs (e.g., the design of the linear accelerator collimation system and/or the optimization of the patient therapeutic dose distribution).
Carlo Casarino, Giorgio Russo, Giuliana Candiano, Giuseppe La Rocca, Roberto Barbera, Giovanni Borasi, Cristina Messa, Mariacarla Gilardi and Gianluca Passaro

Presenter: LA ROCCA, Giuseppe (INFN)

Session Classification: Paper Session 2

Contribution ID: **8** Type: **not specified**

User-friendly workflows in quantum chemistry

Monday, 3 June 2013 17:15 (25 minutes)

Quantum chemical workflows can be built up within the science gateway MoSGrid (Molecular Simulation Grid). Complex workflows required by the endusers are dissected into smaller workflows which can be combined freely to larger meta-workflows. General quantum chemical workflows are described here as well as the real use case of a spectroscopic analysis resulting in an enduser desired meta-workflow. All workflow features are implemented via WS-PGRADE and submitted to UNICORE. The workflows are stored in the MoSGrid repository and ported to the SHIWA repository.
Sonja Herres-Pawlis, Akos Balasko, Georg Birkenheuer, Andre Brinkmann, Sandra Gesing, Richard Grunzke, Alexander Hoffmann, Peter Kacsuk, Jens Krüger, Lars Packschies, Gabor Terstyansky and Noam Weingarten

Presenter: HERRES-PAWLIS, Sonja (Ludwig-Maximilians-Universität München)

Session Classification: Paper Session 2

Contribution ID: **9** Type: **not specified**

Easy Development and Integration of Science Gateways with Vine Toolkit in production.

Tuesday, 4 June 2013 08:45 (1 hour)

The advanced web-based graphic and multimedia oriented user interfaces (GUIs) designed for scientists and engineers could change the way users collaborate, share computing experiments and data, and work together to solve day-to-day problems. Moreover, future science and engineering gateways will influence the way users will not only access their data, but also control and monitor their demanding computing simulations using the Internet. To allow users to interact remotely with clusters, supercomputers and large-scale computing environments in a more interactive and visual manner, we present example Science Gateways for different scientific applications used by communities in the PL-Grid infrastructure. The PL-Grid portal has been successfully integrated with Vine Toolkit it offers now users a set of gateways connected to all production grid sites in Poland. In a nutshell, Vine Toolkit is a modular, extensible and easy-to-use tool that can be used as a graphical web front-end for remote job and data management. It also offers high-level Application Programming Interface (API) for various applications, visualization components and building blocks to allow interoperability between HPC and grid technologies, such as QosCosGrid, gLite, Unicore, iRODS, etc. It supports Adobe Flex and BlazeDS technologies to help developers quickly prototype and build advanced and rich web applications similar to many stand-alone GUIs. Additionally, Vine Toolkit has been integrated with well-known open source web frameworks, such as Liferay and Gridsphere. In this presentation, we briefly describe new technological solutions relevant to advanced scientific and engineering portals driven by various requirements defined by experts in chemistry and nanotechnology.

Presenter: GRABOWSKI, Piotr (Poznań Supercomputing And Networking Center)

Session Classification: Keynote

Contribution ID: **10** Type: **not specified**

VisIVO Science Gateway: a Collaborative Environment for the Astrophysics Community

Tuesday, 4 June 2013 09:45 (25 minutes)

VisIVO Science Gateway is a web based, workflow enabled environment wrapped
around a WS-PGRADE/gUSE portal integrating seamlessly large-scale multi-dimensional astrophysical datasets with applications for processing and visualization based on Distributed Computing Infrastructures (DCIs). We present the main tools and services supported including an application for mobile access to the gateway. We discuss issues in sharing workflows and report our experiences in supporting specialised communities. We present a number of workflows developed recently for visualization and numerical simulations and outline future workflows currently under development. Finally, we summarise our work on the gateway with pointers to future developments.
Eva Sciacca, Marilena Bandieramonte, Ugo Becciani, Alessandro Costa, Mel Krokos, Piero Massimino, Catia Petta, Costantino Pistagna, Simone Riggi and Fabio Vitello, INAF

Presenter: SCIACCA, Eva (INAF, Osservatorio Astrofisico di Catania)

Session Classification: Paper Session 3

Contribution ID: **11** Type: **not specified**

EU-Brazil Open Data and Cloud Computing e-Infrastructure for Biodiversity

Tuesday, 4 June 2013 10:10 (25 minutes)

EUBrazilOpenBio is a collaborative initiative addressing strategic barriers in biodiversity research by integrating open access data and user-friendly tools widely available in Brazil and Europe. The project deploys the EU-Brazil cloud-based e-infrastructure that allows the sharing of hardware, software and data on-demand. This e-Infrastructure provides access to several integrated services and resources to seamlessly aggregate taxonomic, biodiversity and climate data, used by processing services implementing checklist cross-mapping and ecological niche modelling. The concept of Virtual Research Environments is used to provide the users with a single entry point to processing and data resources. This article describes the architecture, demonstration use cases and initial experimental results.
Rafael Amaral, Rosa Badia, Ignacio Blanquer, Leonardo Candela, Donatella Castelli, Renato Giovanni, Alex Gray, Andrew Jones, Daniele Lezzi, Pasquale Pagano, Vanderlei Perez-Canhos, Francisco Quevedo, Roger Rafanell, Vinod Rebello and Erik Torres

Presenter: LEZZI, Daniele (Barcelona Supercomputing Center)

Session Classification: Paper Session 3

Contribution ID: **12** Type: **not specified**

Orbital analysis of oxo and peroxo dicopper complexes via quantum chemical workflows in MoSGrid

Tuesday, 4 June 2013 11:05 (25 minutes)

The science gateway MoSGrid (Molecular Simulation Grid) is a valuable tool to submit and process molecular simulation studies on a large scale. An orbital analysis of oxo and peroxo dicopper complexes, which are bioinspired models of tyrosinase, is presented as a real-world chemical example. The orbital analysis is result of a quantum chemical workflow which has been employed on several tyrosinase model complexes as well as on simple {Cu2O2(NH3)x} cores (with x = 4,6). The structures were optimized using Gaussian09 and the orbitals visualized after production of formatted checkpoint files. All meta- and post-processing steps have been performed in this portlet. All workflow features are implemented via WS-PGRADE and submitted to UNICORE.
Alexander Hoffmann, Sonja Herres-Pawlis, Richard Grunzke and Lars Packschies

Presenter: HOFFMANN, Alexander (Ludwig-Maximilians-Universität München)

Session Classification: Paper Session 4

Contribution ID: **13** Type: **not specified**

On the equivalence of specific control flow and data flow patterns with use cases

Tuesday, 4 June 2013 11:30 (25 minutes)

Although many of workflow languages use workflow patterns in various aspects such as control or data, the implementation of these patterns are quite different, which makes the workflow language interoperability really difficult.
This paper introduces compositions of specific data and control patterns and then proves their equivalence, which benefits that control structures can be replaced by data patterns and vice versa, independently from the implementation itself.
To confirm the necessity of our results we introduce use cases based on community workflows that apply data patterns to trigger control structures. As WS-PGRADE/gUSE workflow management system became a commonly used general framework for workflow development by various scientific communities, the use cases are created in there, to facilitate using our solutions in other workflows.
Akos Balasko and Peter Kacsuk, MTA SZTAKI, Hungary

Presenter: BALASKO, Akos (MTA SZTAKI)

Session Classification: Paper Session 4

Contribution ID: **14** Type: **not specified**

A Model-based Information Security Risk Assessment method for Science Gateways

Tuesday, 4 June 2013 11:55 (25 minutes)

BACKGROUND: Information Security is important for e-Science research groups and other small organisations that design and operate science gateways and virtual research environments, especially when such environments are being used for (bio)medical research. We propose a novel method to do risk assessments: MISRAM, the Model-based Information Security Risk Assessment Method. It uses an information architecture model, a method to assign values to information assets and IT components, and a method to calculate risks. The output of MISRAM is a ranked list of risks and a list of actionable tasks to solve the main issues.
METHODS: MISRAM was applied as a test case to an e-Science research group at a Dutch research hospital. Meetings and surveys were used to create and evaluate lists of information assets and IT components. One meeting was used to create a list of practical task recommendations.
RESULTS: Good insight into the information architecture and security problems of the IT infrastructure was gained. Also the participating group members confirmed that the identified security issues were realistic.
CONCLUSIONS: Our approach raises awareness about security among the developers and operators of e-Science environments. It also gives insight in how the technical architecture affects information security. Traditional questionnaires are an important part of any risk assessment, and MISRAM's inclusion of such generic questionnaires is an important aspect to create an integrated information security risk assessment.
Evert Mouw, Guido Van 'T Noordende, Baas Louter and Silvia Olabarriaga, Amsterdam Medical Center, The Netherlands

Presenter: MOUW, Evert (Academic Medical Center of University of Amsterdam)

Session Classification: Paper Session 4

Contribution ID: **15** Type: **not specified**

Developing Customised Science Gateways at Various Levels of Granularity

Tuesday, 4 June 2013 14:00 (15 minutes)

Science gateways have the potential to offer transparent and user friendly access to a wide variety of distributed computing resources. These tools hide the complexity of the underlying infrastructure from the scientist end-users and let them concentrate on their scientific research problem instead of requiring a steep and sometimes impossible learning curve in complex computing paradigms.

Many web and desktop based tools have been developed in the last few years that have been labelled as science gateways. However, close examination of these tools reveal that the level of granularity how end-users can access the applications is rather varied. There are solutions which do not aim to hide the details of the original command line interface and simply provide web based access to the application. On the other extreme, there are custom built portals supporting a single or a small family of applications and providing highly intuitive graphical user interfaces incorporating visualization tools, for example.

This presentation will analyse and illustrate via the example of a molecular docking gateway "family", how customised science gateways at different levels of granularity regarding their user access can be developed using a generic science gateway framework.

When compared to completely customised "from scratch"development, generic science gateway frameworks provide services readily available, significantly decreasing development time and effort. On the other hand, if these frameworks are extended with customisation methodologies then highly specific gateways can be built on top of them with fraction of the cost.

A molecular docking gateway supporting several user scenarios (random blind docking and virtual screening) has been implemented on top of the generic WS-PGRADE/gUSE [1] gateway framework. The presentation analyses and compares four different levels of gateway granularity. Although these implementations are almost identical regarding their performance, significant difference exists regarding the required development effort and also from the perspective of usability.

At the lowest level, the generic framework can be deployed "out of the box"and offered directly for end-users who can build their workflow applications. Although the deployment of the portal is simple and does not require any specific modification or customisation, significant effort and learning curve is needed from the end-users'perspective. In the second scenario, the concrete workflows are developed by specialist workflow developers and published in the WS-PGRADE workflow repository making them accessible for scientists. The third solution demonstrates the end-user interface of WS-PGRADE where workflow developers can create custom web-forms on top of the concrete workflows just by dragging and dropping. Finally, a highly specific set of portlets have been developed using the Application Specific Module (ASM) API of WS P-GRADE incorporating visualisation tools and highly specific user interfaces.

ACKNOWLEDGMENT

The research leading to these results has received funding from the SCI-BUS project, supported by the European Commission Seventh Framework Programme (FP7) (grant agreement no. RI-283481).

REFERENCES

[1] Kacsuk P, P-GRADE portal family for Grid infrastructures, Concurrency and Computation: Practice and Experience journal, Volume: 23, Issue: 3, 2011, pp. 235-245.

Tamas Kiss, Peter Borsody, Attila Sasvari, Gabor Terstyanszky, Zoltan Farkas and Peter Kacsuk

Presenter: KISS, Tamas (University of Westminster, London, UK)

Session Classification: Lightning Talks Session 1

Contribution ID: **16** Type: **not specified**

Statistical Seismology Science Gateway

Tuesday, 4 June 2013 14:15 (15 minutes)

Statistical Seismology aims to bridge the gap between physics and statistics based models. Statistical methods are useful for characterizing seismic hazard because earthquakes are random phenomena. They provide additional insights to the seismic hazard or risk problem through effective modeling.
The Statistical Seismology Science Gateway provides robust and effective statistical methods in order to obtain more precise estimates, to reduce uncertainties and to conduct more reliable seismic hazard analysis. With this gateway skilled users can develop their own seismology models simply by accessing the gateway services through programming or even workflow development, while novice users can easily use existing applications. It aims to implement the following foremost statistical seismology functions, reflecting the latest advances in the literature, for the international seismology community:
- Integration of multi-source data for increased reliability and quality
In order to construct a consistent and reliable earthquake database, the catalogs for related regions are compiled from different sources by considering both historical and instrumental seismicity. While entries in the catalogues are crosschecked to ensure that repetitions are not included by using new statistical techniques, completeness analysis is performed, too.
- Determination of probability distributions and their input parameters
Due to uncertainties in magnitudes, source to site distances, holding times between earthquakes, determination of their probability distribution functions are very important. Our studies have indicated that sometimes probability distributions can be different than the ones available in the literature. Here, methods are provided for the determination of such probability distributions and estimation of their input parameters.
- Robust techniques for the parameter estimation in models and relations
Robust estimation techniques like Huber's M, modified maximum likelihood and adaptive modified maximum likelihood for the estimation of earthquake hazard parameters are implemented.
- Complex predictive modeling of earthquake phenomena in time-space domains
The most widely used and powerful spatial, temporal and spatiotemporal models for earthquake occurrences are served. Statistical comparisons among them are also carried out by considering their advantages and defects.
- Ground motion prediction equations
A critical input to seismic hazard analysis is the ground motion prediction equation. Next-generation ground motion prediction equations are more realistic and allow us to better quantify source, propagation, and site effects. The most commonly used and next generation equations are provided together with robust regression techniques to estimate the parameters of these equations.
- Complex logic-tree and sensitivity analysis in probabilistic seismic hazard assessment
The logic tree and sensitivity analysis are statistical analyses used to identify the input parameters that have the greatest impact on hazard assessment and its uncertainty. We propose objective and statistically supportive criteria for assessing weights to the branches of the logic tree.
- Statistical calculations for seismic risk
The statistical calculations are performed for seismic risk to determine the likelihood of certain events occurring and magnitude of their possible consequences.
The gateway executes these functions on high-capacity distributed computing infrastructures such as clusters or grids.

Çelebi Kocair, Ayşen Akkaya and Cevat Şener

Presenter: AKKAYA, Aysen (METU)

Session Classification: Lightning Talks Session 1

Contribution ID: **17** Type: **not specified**

AEGIS CMPC Scientific Gateway

Tuesday, 4 June 2013 14:30 (15 minutes)

Numerical simulations in the condensed matter physics deploy a broad range of algorithms, such as solving of nonlinear partial differential equations, classical and quantum Monte Carlo techniques, including solving of Bose-Hubbard and Fermi-Hubbard models, exact diagonalization techniques for strongly correlated systems, etc. Whichever is chosen, typically it will require large-scale computing resources for simulations of relevant physical systems.
Within the Serbian Condensed Matter Physics Community (CMPC), the use of computing resources is mainly related to the three applications: GP-SCL [1], SPEEDUP [2,3] and QSPEEDUP [4]. The first one solves Gross-Pitaevskii equation, the effective equation governing the dynamics of Bose-Einstein condensates and nonlinear optical systems. The application solves both time-dependent and time-independent Gross-Pitaevskii equation in one, two, and three spatial dimensions using imaginary-time and real-time propagation. The SPEEDUP application uses Monte Carlo, while QSPEEDUP uses quasi-Monte Carlo-based path integral algorithm for calculation of quantum mechanical transition amplitudes and partition functions for 1D models.
Recently, within the framework of SCI-BUS project and over the Academic and Educational Grid Initiative of Serbia (AEGIS) infrastructure [5], we have provided Scientific Gateway on top of these three applications. The development relies on a widely utilized Liferay-based WS-PGRADE/gUSE portal, which was extended with the application-specific workflows, and a set of modules that enable specification of the physical system to be studied numerically, as well as exporting of the results of simulations.
In this presentation we explain the development and implementation of AEGIS CMPC Scientific Gateway and main technologies used. We also briefly describe main workflows and present plans for further development.

Acknowledgments
The research leading to these results has received funding from the SCI-BUS project, supported by the European Commission Seventh Framework Programme (FP7) (grant agreement no. RI-283481), and from the national research project ON171017, supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

References

[1] D. Vudragovic, I. Vidanovic, A. Balaz, P. Muruganandam and S. Adhikari: "C Programs for Solving the Time-dependent Gross-Pitaevskii Equation in a Fully Anisotropic Trap", Comput. Phys. Commun. 183 (2012) 2021
[2] A. Balaz, I. Vidanovic, D. Stojiljkovic, D. Vudragovic, A. Belic and A. Bogojevic: "SPEEDUP Code for Calculation of Transition Amplitudes Via the Effective Action Approach", Commun. Comput. Phys. 11 (2012) 739
[3] A. Balaz, A. Bogojevic, I. Vidanovic and A. Pelster: "Recursive Schroedinger Equation Approach to Faster Converging Path Integrals", Phys. Rev. E 79 (2009) 036701
[4] D. Vudragovic, A. Balaz, A. Belic and A. Bogojevic: "QSPEEDUP: Quasi-MC Implementation of the SPEEDUP Path Integral Code", Proceedings of the INFOTEH 2010, 17-19 March 2010, p. 73, A-15, Jahorina, Bosnia and Herzegovina
[5] D. Vudragovic, A. Balaz, V. Slavnic and A. Belic: "Serbian Participation in Grid Computing Projects", Proceedings of the NEC2009, 7-14 September 2009, p. 286-293, Varna, Bulgaria

Petar Jovanovic, Nikola Grkic, Dusan Vudragovic and Antun Balaz

Presenter: JOVANOVIC, Petar (Institute of Physics Belgrade)

Session Classification: Lightning Talks Session 1

Contribution ID: **18** Type: **not specified**

WRF4SG: A Scientific Gateway for WRF community

Tuesday, 4 June 2013 14:45 (15 minutes)

Weather Research Forecasting (WRF) model is a public domain software with a world wide spread community of users. This community is heterogeneous both in terms of application domains and Distributed Computing Infrastructure (DCI) profiles. WRF researches are physicists, chemists, mathematicians and engineers who demand a huge variety of DCIs in order to tackle climate experiments such as operational weather forecast, chemistry applications, regional projections of climate change scenarios or atmospheric physics/parametrization research by using WRF.
WRF modeling system is composed of several components which are executed sequentially. Myriad of input and output files have to be managed and monitored in order to run a WRF simulation. Therefore, when a experiment requires executing more than one simulation, its complexity rises almost exponentially demanding an enormous effort for a researcher.
This work proposes to develop a scientific gateway which is going to provide the services needed to perform WRF experiments on heterogeneous DCIs. In addition, it will offer a set of saving time tools in order to design, monitor and manage complex climate experiments such as weather forecast, extreme weather case studies, future climate projections or re-forecasts. This scientific gateway, called WRF4SG (WRF for Scientific Gateway), consists in porting WRF4G application to the WS-PGRADE/gUSE framework, which provides a wide access to DCIs and the possibility to design more complex experiments by creating specific portlets.
Finally, the objective of WRF4SG is able to divide into two goals. The first goal is to bring a more comprehensive interface to WRF community in order to manage more ambitious climate experiments. The second goal is to assist researchers in order to run their experiments concurrently on heterogeneous DCIs.
Jose Carlos Blanco, Antonio S. Cofiño and Valvanuz Fernández-Quiruelas

Presenter: BLANCO, Jose Carlos (University of Cantabria)

Session Classification: Lightning Talks Session 1

Contribution ID: **19** Type: **not specified**

Application of the Science Gateway Portal on the Basis of WS-PGRADE Technology for Simulation of Aggregation Kinetics and Molecular Dynamics Simulations of Metal-Organic Nanostructure

Tuesday, 4 June 2013 15:00 (15 minutes)

Nowadays new materials are of great interests that have nanoscale structure (nanomaterials) and unique properties, such as metal nanocrystals, metal nanorodes, and nanoscale non-metallic (organic) materials like carbon nanotubes (CNT, graphene, etc.) and their complicated ensembles. Molecular dynamics (MD) and Monte Carlo (MC) simulations of nanoscale processes in the wide range of physical parameters are very promising, because of possibility for parameter sweeping in a brute force manner in the available heterogeneous distributed computing infrastructure (DCI) on the basis of Desktop Grids (DGs), Service Grids (SGs), clusters, cloud resources, etc. The recent advances in computing algorithms and infrastructures, especially in development of DCIs, allow us to use the efficient methods for solving these tasks without expensive scaling-up. DCIs on the basis of the BOINC SZDG, XtremWeb-HEP, OurGrid, EDGeS platforms for high-performance distributed computing are very promising way to use the heterogeneous computing resources, especially by means of the science gateway technology on the basis of WS-PGRADE platform. The main objective of the paper is to demonstrate the capabilities of the proposed specific science gateway (SSG) on the basis of WS-PGRADE platform for MC and MD simulation and data processing on the basis of LAMMPS and other packages (R, MolTemplate, etc.). The several typical workflows were created for simulation of several physical processes with various demands for the computing resources: tension of metal nanocrystals under different physical conditions, tension of ensemble of metal nanocrystals under the same conditions, manipulations with complex nanostructures like Si-CNT-Si, CNT-graphene-water, etc.

It is shown that the physical characteristics evaluated on the basis of MD simulations in the proposed SSG on the basis of LAMMPS are in satisfactory agreement with the experimental data and allowed to discover the new aspects of operation and manipulation of nanomaterials. Porting MD-applications to heterogeneous DCI through SSG is easy and efficient, if WS-PGRADE platform is used; parameter decomposition and sweeping parallelism are possible; message passing is localized at worker side.

The work presented here was partially funded by EU FP7 SCI-BUS (SCIentific gateway Based User Support) project, No. RI-283481, and partially supported in the framework of the research theme "Introduction and Use of Grid Technology in Scientific Research of IMP NASU"under the State Targeted Scientific and Technical Program to Implement Grid Technology in 2009-2013.

Olexandra Baskova, Olexander Gatsenko, Lev Bekenev, Elena Zasimchuk and Yuri Gordienko

Presenter: GORDIENKO, Yuri (G.V.Kurdyumov Institute for Metal Physics, National Academy of Sciences)

Session Classification: Lightning Talks Session 1

Contribution ID: **20** Type: **not specified**

HelioGate: A portal for Heliophysics

Tuesday, 4 June 2013 15:15 (15 minutes)

Heliophysics is the branch of physics that investigates the interactions among different events across the Solar System. This investigation usually takes place by finding the relevant events and then modelling their interactions through physical and mathematical models.
Events can be found by querying pre-compiled catalogues of metadata or by extracting relevant features directly from instrument data. The analysis of the interactions and correlation among different events is performed with propagation models: mathematical abstractions that describe and predict how physical events move across the Solar System.
Most of these operations can be time consuming and repetitive; in order to offer to the community faster and more intuitive tools for their research, we are building a portal for Heliophysics; an effort supported by the SCI-BUS (https://www.sci-bus.eu/) project.
Although the HelioGate portal is being developed within the SCI-BUS project, it is strongly connected to other two projects: HELIO (http://www.helio-vo.eu/) and ER-FLOW (http://www.erflow.eu/).
The SCI-BUS project, with its capabilities of handling multiple computational back-ends and its multi-layered architecture spanning from GUIs to the computational resources, offers an ideal technology for the development of the portal.
The HELIO project offers the most comprehensive integrated information system in this domain and provides access and coordination to services to mine and analyze data. Furthermore, HELIO also provides a Propagation Model called SHEBA
Finally, ER-FLOW is a Coordination and Support Action that focuses on porting workflows to the SHIWA (http://www.shiwa-workflow.eu/) platform. This will offer various benefits to the Heliophysics community. First it allows for workflows written in different platforms to interoperate, second it allows for workflows to be combined in meta-workflows, third it allows the execution of workflows through simplified interfaces.
The first functionality developed in the portal is an extension of the SHEBA propagation model that allows a finer selection of the governing parameters by using event catalogues to validate results. The extended propagation model is now further enriched by workflows (developed within the HELIO project) that are being adapted and ported in the SHIWA platform (and hence accessible via the SCI-BUS technology) in the ER-FLOW project.
These workflows execute sophisticated orchestrations of HELIO services including catalogues of events and features, location and capabilities of instruments, numerical data evaluation and the propagation model.
Gabriele Pierantoni and David Pérez-Suárez

Presenter: PIERANTONI, Gabriele (TCD)

Session Classification: Lightning Talks Session 1

Contribution ID: **21** Type: **not specified**

Data Bridge: solving diverse data access in scientific Applications

Tuesday, 4 June 2013 16:00 (25 minutes)

The nature of data for scientific computation is very diverse in the age of big data. First, it may be available at a number of locations, e.g. the scientist's machine, some institutional filesystem, a remote service, or some sort of database. Second, the size of the data may vary from a few kilobytes to many terabytes. In order to be available for computation, data has to be transferred to the location where the computation takes place. This requires a diverse set of middleware tools that are compatible both with the data and the compute resources. However, using this tools requires additional knowledge and makes running the experiments an inconvenient task. In this paper we present the Data Bridge, a high-level service that can be used easily in scientific computations to perform data transfer to and from a diverse set of storage services. The Data Bridge not only unifies access to different types of storage services, but it can also be used at different levels (e.g., single jobs, parameter sweeps, scientific workflows) in scientific computations.
Zoltan Farkas, Peter Kacsuk, Mark Santcroos, Silvia Olabarriaga, Akos Balasko and Krisztian Karoczkai

Presenter: FARKAS, Zoltan (MTA SZTAKI)

Session Classification: Paper Session 5

Contribution ID: **22** Type: **not specified**

Rapid Prototyping of Science Gateways in the Brazilian National HPC Network

Tuesday, 4 June 2013 16:25 (25 minutes)

Arguably, an important amount of scientific software development time is likely to be employed on user interfaces. In particular, science gateways have gained increasing interest from the e-Science community because of their convenience to hide the complexity of the underlying resources that give support to the management of scientific data and to the execution of scientific applications. Based on our previous experience with the development of science gateways for diverse application domains in the Brazilian National HPC Network (SINAPAD), we have devised a rapid prototyping strategy to lower the barrier for scientific application developers to launch new science gateways. In this paper we present such strategy, which is based on two main tools. The first tool implements a gateway engine that can be configured by a small set of XML files. Such files completely define the desired functionality of an specific science gateway in such an engine. The gateway engine also offers other features not commonly found in related technologies, such as file sharing, data provenance tracking, and restricted anonymous access to underlying computational resources. The second tool implements both an editor and a packager for the aforementioned engine, allowing the developer to rapidly deploy and launch a new science gateway in ordinary Web application containers. In this paper we present our results with the use of both tools in the SINAPAD network. We also discuss about the current limitations of such tools, as well as how we have been dealing with such limitations to provide a more comprehensive toolset to developers.
Bruno Fernandes Bastos, Vinicius Macedo Moreira and Antonio Tadeu Azevedo Gomes

Presenter: FERNANDES BASTOS, Bruno (Brazilian National System for High-Performance Computing - SINAPAD)

Session Classification: Paper Session 5

Contribution ID: **23** Type: **not specified**

ICAT Job Portal: a generic job submission system built on a scientific data catalog

Tuesday, 4 June 2013 16:50 (25 minutes)

The value of metadata to the scientist is well known: with the right choice of metadata, data files can be selected very quickly without having to scan through huge volumes of data. The ICAT metadata catalog (which is part of the ICAT project) allows the scientist to store and query information about individual data files and sets of data files as well as storing provenance information. This paper explains how a generic job management system, exposed as a web portal, has been built on top of ICAT. This gives the scientist easy access to a high performance computing infrastructure without allowing the complexities of that infrastructure to impede progress.
The aim was to build a job and data management portal capable of dealing with batch and interactive work that would be simple to use and that was based on tried and tested, scalable, and preferably open source technologies. For the team operating the portal, it needed to be generic and configurable enough so that they can, without too much effort, modify their software to run within the portal, add new software, and create new dataset types and parameters. Modifications to existing software should be limited to saving and loading their datasets in a slightly different way so that instead of just being saved to disk, they are registered within the system along with recording any provenance information.
Stephen Fisher, Kevin Phipps and Daniel Rolfe

Presenter: FISHER, Stephen (RAL STFC)

Session Classification: Paper Session 5

Contribution ID: **24** Type: **not specified**

Bringing Scientific Workflow to the Masses via Pegasus and HUBzero

Tuesday, 4 June 2013 17:15 (25 minutes)

Scientific workflow managers are powerful tools for handling large computational tasks. Domain scientists find it difficult to create new workflows, so many tasks that could benefit from workflow automation are often avoided or done by hand. Two technologies have come together to bring the benefits of workflow to the masses. The Pegasus Workflow Management System can manage workflows comprised of millions of tasks, all the while recording data about the execution and intermediate results so that the provenance of the final result is clear. The HUBzero platform for scientific collaboration provides a venue for building and delivering tools to researchers and educators. With the press of a button, these tools can launch Pegasus workflows on national computing infrastructures and bring results back for plotting and visualization. As a result, the combination of Pegasus and HUBzero is bringing high-throughput computing to a much wider audience.
Michael McLennan, Steven Clark, Ewa Deelman, Mats Rynge, Frank McKenna, Derrick Kearney and Carol Song

Presenter: MICHAEL, McLennan (Purdue University)

Session Classification: Paper Session 5

Contribution ID: **25** Type: **not specified**

Recipes 2.0: Building for Today and Tomorrow

Tuesday, 4 June 2013 17:40 (25 minutes)

The history of science gateway development has, in many ways, been a story of the "Haves"vs. the "Have-nots."Large infrastructure projects led the way, building thick client portals to provide coherent interfaces to an incoherent environment. Contrast this with the way the modern web is designed using light, front end components and outsourcing much of the heavy lifting to a mash-up of REST APIs, and it is easy to see why modern web applications can be prototyped and refined into stable products in the time it previously took thick client portals to do an initial release. This paper argues that a "build for today"philosophy can lead to the rapid development of science gateways to serve the "Have-nots."Also presented is a set of responsive front end components built on top of the iPlant Foundation API that provide the boilerplate for rapid development of lightweight science gateways using only HTML, JavaScript, and CSS. Using these components, developers can easily stand up new gateways or quickly add new functionality to existing ones.

Presenter: DOOLEY, Rion (University of Texas at Austin/ Texas Advanced Computing Center)

Session Classification: Paper Session 5

Contribution ID: **26** Type: **not specified**

Introducing the Neuroscience Gateway

Wednesday, 5 June 2013 09:00 (25 minutes)

The last few decades have seen the emergence of computational neuroscience as a mature field where researchers are interested in modeling complex and large neuronal systems and require access to high performance computing machines and associated cyberinfrastructure to manage computational workflow and data. The neuronal simulation tools, used in this research field, are also implemented for parallel computers and suitable for high performance computing machines. But using these tools on complex high performance computing machines remain a challenge due to issues with acquiring computer time on these machines located at national supercomputer centers, dealing with complex user interface of these machines, dealing with data management and retrieval etc. The Neuroscience Gateway is being developed to alleviate all of these barriers to entry for computational neuroscientist. It hides or eliminates, from the point of view of the users, all the administrative and technical barriers and makes parallel neuronal simulation tools easily available and accessible on complex high performance computing machines and handles the running of jobs and data management and retrieval. This paper describes the architecture it is based on, how it is implemented, and how users can use this for computational neuroscience research using high performance computing at the back end.
Subhashini Sivagnanam, Amit Majumdar, Kenneth Yoshimoto, Vadim Astakhov, Anita Bandrowski, Maryann Martone and Ted Carnevale

Presenter: MAJUMDAR, Amitava (San Diego Supercomptuer Center, Univ. of California San Diego)

Session Classification: Final Paper Session

Contribution ID: **27** Type: **not specified**

A Customised and Extendable Science Gateway Supporting Diverse Local User Communities

Wednesday, 5 June 2013 09:25 (15 minutes)

Universities and research institutes regularly offer local high-performance or high-throughput computing resources supporting their diverse internal user communities. These local users can range from researchers to students at various levels of expertise, and can span different disciplines and application areas, depending on the individual institution. The provided computational and data resources can also be rather diverse, including local clusters, desktop grid resources or private cloud computing infrastructures. These resources are typically provided free of charge but restricted to local users.

Supporting localised computing infrastructures with customised science gateways raise interesting new challenges due to the diverse nature of the targeted user communities. While public science gateways are typically related to virtual organisations and user communities of a specific discipline, local gateways may require supporting as diverse user communities as for example bioscientists, social scientists, or computer animation students, all at the same time. Most of these users are also newcomers to distributed computing infrastructures gaining their first ever experience with high-performance or high-throughput computing locally.

Deploying a generic science gateway framework for these users, such as WS-PGRADE [1], could be an easy solution on top of these local infrastructures. Although a generic science gateway framework simplifies access to the computing resources, it still requires a learning curve and represents significant hurdle to newcomers.

This presentation explains how an extendible and fully customised science gateway has been built at the University of Westminster that provides custom views to users from diverse disciplines. The University of Westminster Desktop Grid Portal [2], built on top of the WS-PGRADE framework, is connected to a local BOINC [3] based desktop grid infrastructure that connects over 1,800 laboratory PCs of the university into a powerful computing resource. The gateway currently supports two user communities. Computer animation and games students can use it to render their coursework utilising the MentalRay commercial rendering engine. On the other hand, researchers and students from the School of Life Sciences run molecular docking simulations on the infrastructure. Both user communities have access to custom end-user interfaces in the form of specialist portlets, plugged into the generic WS-PGRADE portal.

It is envisaged, that further custom portlets for other local user communities will be developed and deployed in the near future. Due to its easily extendable Liferay-based framework, the local gateway can accommodate this extension easily.

ACKNOWLEDGMENT

The research leading to these results has received funding from the SCI-BUS project, supported by the European Commission Seventh Framework Programme (FP7) (grant agreement no. RI-283481).

REFERENCES

[1] Kacsuk P, P-GRADE portal family for Grid infrastructures, Concurrency and Computation: Practice and Experience journal, Volume: 23, Issue: 3, 2011, pp. 235-245.

[2] The University of Westminster Desktop Grid portal, https://dg-portal.wmin.ac.uk/liferay-portal-6.1.0/

[3] Anderson D.P., BOINC: A System for Public-Resource Computing and Storage, 5th IEEE/ACM International Workshop on Grid Computing, November 8, 2004, Pittsburgh, USA, pages 4-10, ISSN

1550-5510.
Tamas Kiss, Peter Borsody, Attila Sasvari, Dario Ferrer and Gabor Terstyanszky

Presenter: BORSODY, Peter (University of Westminster)

Session Classification: Lightning Talks Session 2

Contribution ID: **28** Type: **not specified**

Deploying astronomical workflows in heterogeneous distributed computing infrastructures

Wednesday, 5 June 2013 09:40 (15 minutes)

Astronomy is facing an exponential increase of the data generated by the state of the art of the instruments (e.g. ALMA, LOFAR, SKA precursors, etc.). This copious data flux involves a technological challenge not only for streaming the data from the instruments to the data centres but also for analysing them and for extracting new science knowledge from them.
The AMIGA group (http://amiga.iaa.es) leads AMIGA4GAS (AMIGA for GTC, ALMA and SKA pathfinders, AYA2011-30491-C02-01), a coordinated project between the Instituto de Astrofísica de Andalucía –CSIC (IAA-CSIC) and the Foundation of Supercomputing Centre of Castilla and León (FCSCL).
This project aims to build a federated layer which facilitates astronomers launching their workflows in heterogeneous distributed computing infrastructures. This will allow the user to be unaware of technical aspects and also will smartly distribute the workflow tasks to the infrastructures in function of its characteristics such as input data, available resources status and data transfer latency.
The developments have been driven by a generic use case, in which datacubes are analysed to get kinematical models of galaxies. The tasks of this use case are being implemented as web services which run on computing infrastructures, specifically on a Sun Grid Engine cluster, on a cloud computing system, and on the European Grid Infrastructure (EGI) through the IBERGRID Virtual Organisation phys.vo.ibergrid.eu. One of the pieces of these implementations is COMPSs, a programming framework developed by the Barcelona Supercomputing Centre (BSC) which facilitates the execution of applications on distributed infrastructures,optimizing the use of them.
In this talk we will present the architecture of the federated layer, the role of COMPSs, how the web services are connected to the infrastructures, the problems found, and the next steps.
Susana Sanchez Exposito, Pablo Martin, Jose Enrique Ruiz, Elena Ortega, Lourdes Verdes-Montenegro, Antonio Ruiz-Falco, Julian Garrido, Juan De Dios Santander-Vela, Raul Sirvent, Enric Tejedor, Rosa M. Badia and Jesus Lorenzana

Presenter: SANCHEZ EXPOSITO, Susana (Instituto de Astrofísica de Andalucía - CSIC)

Session Classification: Lightning Talks Session 2

Contribution ID: **29** Type: **not specified**

Gateways to analyse and publish High Content Screening imaging data in SystemsX.ch

Wednesday, 5 June 2013 09:55 (15 minutes)

RNAi-based High Content Screening (HCS) is a recent imaging technology used in systems biology to study cellular phenomena on a large scale. HCS experiments produce very large amounts of imaging data that require nontrivial automated analysis procedures, with reliable and traceable processing and data management capabilities. Due to the large data volumes and millions of image files produced, HCS presents significant IT challenges to researchers and platform operators. Our software system for RNAi HCS integrates automated analysis and customizable data management. It enables robust and complex parallel processing on large computing infrastructures and assures reliable storage of primary and resulting data, linking the results with the raw imaging data by making use of the open biology information system OpenBIS. Our image processing gateway, called iBrain2, implements a state machine, which is described in an interchangeable XML format. An image processing expert can set up several customizable pipelines to work through the individual image processing steps. The gateway interface is very easy to use by the end-user, who can simply choose a predefined image processing pipeline with her dataset to start processing. The gateway provides also a monitoring view to show details of the status of the jobs in the pipeline. We propose a lightning talk and demo to show the iBrain2 processing gateway as well as the OpenBIS data browsing and visualization interfaces for HCS data as used in the SystemsX.ch InfectX project (see www.infectx.ch and www.infectome.org).
Mario Emmenlauer, Pauli Rämö, Eva Pujadas, Béla Hullár, Bernd Rinn and Peter Kunszt

Presenter: EMMENLAUER, Mario (University of Basel)

Session Classification: Lightning Talks Session 2

Contribution ID: **30** Type: **not specified**

A gateway for the determination of the 3D structure of paramagnetic biological macromolecules

Wednesday, 5 June 2013 10:10 (15 minutes)

High-resolution Nuclear Magnetic Resonance (NMR) spectroscopy is one of two main techniques that allow determining three-dimensional (3D) structures of biomacromolecules, such as proteins, RNA, DNA, and their complexes, at atomic resolution. Knowledge of the 3D structure of macromolecules is vital for understanding their function and mechanism of action, and can guide the design of further experimental studies such as rational drug design.
Structural Biology, in general, and NMR in particular, have always been associated with advanced computing. Determining the 3D structure of a protein by NMR typically involves several steps during which the raw data have to be processed to identify the frequency of resonance of each atom. This information is then used to identify relationships between atoms, such as proximity in space, ultimately providing structural restraints. These include inter-atomic distance restraints, dihedral angle restraints, and orientation restraints. Often, proteins coordinate one or more metal ions, which can be paramagnetic. The presence of a paramagnetic metal ion within a protein notably affects its NMR spectra. Most importantly, the perturbations induced by the paramagnetic metal on the chemical shifts and relaxation rates of nuclear spins contain structural and dynamic information. Paramagnetic effects can thus be exploited to obtain paramagnetism-based restraints. All the aforementioned structural restraints are used together to calculate the 3D structure of the protein under investigation. Different modeling approaches can be used, typically as a pipeline of software tools. The result is an ensemble of different conformations in equal agreement with the restraints.
Our aim was to build a uniform environment[1] allowing users to apply a variety of computational tools to their NMR data in a seamless manner. In particular, we developed application web portals to assist the determination and refinement of macromolecular structures, using the programs XPLOR-NIH[2] for structure calculations and AMBER[3,4] for energy refinement. These web portals provide a personal workspace for the creation and management of calculations. This framework allows users to create a new single calculation or a so-called project in which it is possible to store a number of related calculations (e.g. determination of a protein structure with different ensembles of restraints). We built an integrated workflow pipeline, incorporating algorithms to perform the initial data analysis, tools allowing different programs to interact and exchange data, and mechanisms taking care of Grid submission and monitoring. The interaction with the Grid is transparent to the user. All output data are stored and available for browsing and downloading in the personal workspace of the user. Our portals are available via the WeNMR project web site at http://www.wenmr.eu/wenmr/nmr-services
[1] T.A. Wassenaar, et al. J. Grid Comp., 10:743-767 (2012)
[2] Schwieters, C.D., Kuszewski, J.J., Tjandra, N., Clore, G.M.: J. Magn. Reson. 160:65-73 (2003)
[3] Case, D.A., et al. J. Comput. Chem. 26:1668-1688 (2005)
[4] Bertini, I., Case, D.A., Ferella, L., Giachetti, A., Rosato, A. Bioinformatics 27:2384-2390 (2011)
Lucio Ferella, Andrea Giachetti and Antonio Rosato

Presenter: FERELLA, Lucio (Magnetic Resonance Center (CERM), University of Florence)

Session Classification: Lightning Talks Session 2

Contribution ID: **31** Type: **not specified**

A Catch-All Scientific Gateway Portal for Malaysia Research Communities

Wednesday, 5 June 2013 11:00 (15 minutes)

Academic Grid Malaysia provides an open access Distributed Computing Infrastructure (DCI) based on Grid Computing technology to the user communities from various domains (e.g. life science, engineering, multimedia) to run their existing applications more quickly and efficiently, and also to create ambitious new applications, without investing in extra hardware, and software resources. This DCI is built based on the EMI Grid Middleware software. However, many of the Academic Grid Malaysia's users are not familiar with EMI's job scripting and command line. Moreover, converting an application to be a DCI-enabled application and creating a new application requires significant technical knowledge, which most of the Academic Grid Malaysia's users do not have. To solved these problems, we developed a "catch-all" Scientific Gateway portal that based on WS-PGRADE/gUSE platform and the Application Specific Module that is supported by the Scientific Gateway Based User Support (SCI-BUS) project (an EU FP7 Capacities Program) in collaboration between NVG Scientific Sdn Bhd and Academic Grid Malaysia. By using the Scientific Gateway that we developed, the users can submit their simulation/application job to the Academic Grid Malaysia's infrastructure easily without having to "struggle" with EMI's job scripting and command line. It is hoped that by having this Scientific Gateway, it will increase the usage of the Academic Grid Malaysia's infrastructure and also the adoption of DCI technology among Malaysia research communities.

Elizabeth Pek Iee, Mohammad Yaser Shafazand and Muhammad Farhan Sjaugi

Presenter: FARHAN, Muhammad (Malaysia)

Session Classification: Lightning Talks Industry Session

Contribution ID: **32** Type: **not specified**

SCI-BUS and the CloudBroker Platform: Extending Science Gateways to Clouds

Wednesday, 5 June 2013 11:15 (15 minutes)

Cloud computing [1] is currently a hot topic for all areas of computational science. It allows to access computer infrastructure, platform and software as a service through the internet on demand and in a scalable and pay-per-use fashion. Public, community, private and hybrid clouds fulfill various needs regarding flexibility and privacy. This makes clouds very interesting for science gateways [2], to utilize cloud computing resources for the services they offer, to run in the cloud themselves, or to adopt the cloud business model. Here we will show an example how science gateways can use clouds as Distributed Computing Infrastructure (DCI) backends for running their compute-intensive application jobs.
The SCI-BUS (SCIentific gateway Based User Support) EU FP7 project [3] develops gateway technology and community gateways to provide researchers seamless access to major computing, data and networking infrastructures and services, with focus on scientific workflows. SCI-BUS is a collaboration of 15 consortium members and six subcontractors, supporting 17 gateways in various disciplines, including proteomics, molecular science, seismology, business processes, medicine, rendering, astrophysics, SMEs, software testing, citizen web and heliophysics.
As its backend interface to public and private clouds, SCI-BUS utilizes the CloudBroker Platform [4]. The CloudBroker Platform is an application store for scientific and technical software in the cloud. It facilitates deploying, offering and running applications as Software as a Service (SaaS) on various Infrastructure as a Service (IaaS) clouds such as Amazon EC2 and S3 and IBM SmartCloud Enterprise. The platform enables automated software, resource, job and data management, dynamic queuing, user authentication, authorization and transport layer security, as well as transparent accounting and billing. It exposes its functionalities as web UI, CLI, REST web service and Java client APIs, which make it easy to use and integrate. Its marketplace combines free access to open clouds and software with pay-per-use for commercial clouds and software, providing new business models compared to grids.
In this presentation, we will show which features have been added to the CloudBroker Platform during the SCI-BUS project that make it attractive as interface to cloud infrastructures for science gateways. In particular, novel adapters to open source cloud systems such as OpenStack, Eucalyptus and OpenNebula have been implemented. These allow SCI-BUS to utilize private clouds for example at the University of Zaragoza (BIFI) and at the Hungarian Academy of Sciences (MTA SZTAKI). We will summarize the technical integration via WS-PGRADE/gUSE [5] and give an overview about the status and plans regarding porting and using of applications in the cloud within the SCI-BUS project.

ACKNOWLEDGMENT
The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7) Capacities Programme under grant agreement no. RI-283481.

REFERENCES
[1] The NIST Definition of Cloud Computing, http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf
[2] EGI-InSPIRE Science Gateway Primer, https://documents.egi.eu/document/1463
[3] SCI-BUS project, http://www.sci-bus.eu
[4] CloudBroker Platform (public version), https://platform.cloudbroker.com
[5] WS-PGRADE/gUSE, http://www.guse.hu
Wibke Sudholt and Nicola Fantini, CloudBroker GmbH Switzerland

Presenter: SUDHOLT, Wibke (CloudBroker GmbH)

Session Classification: Lightning Talks Industry Session

Contribution ID: **33** Type: **not specified**

Eliminating papers and creating electronic archives by using clouds

Wednesday, 5 June 2013 11:30 (15 minutes)

Based on the global economic crisis more and more sounds could have been heard with the following words: „Costs have to be reduced”! Perhaps, that is the reason why the vision of paperless office got again into the focus (not just in Hungary). Several legal regulations exist that aim to provide background of applying electronic business cases, workflows. The most important among these is the Hungarian „Decree No. 13/2005. (X. 27.) of the Ministry of Information and Communications on the Rules of Digitization of Paper-based Documents”, which describes the requirements of digitizing paper-based documents and keeping them in archives in a secure way. This is a good choice for business companies and public administration to immediately reduce their operating costs (e.g. in public administration electronic public archives, or hybrid mailing and inverse hybrid mailing service of Magyar Posta –Hungarian Post –are built upon these regulations).
The implementation of such solutions that are based on these relevant legislations is not easy. In the presentation we will demonstrate a sample workflow, identify the critical points and introduce services that can be used in a cloud system.
The core workflow of creating authentic electronic archives of paper-based documents is the following:
• the digitization (scanning) of paper-based documents,
• the character recognition (OCR) of digitized (scanned) documents,
• the indexing and metadata retrieving, labeling, classification of digitized (scanned) documents,
• the signing (using XML-based XAdES electronic signature technology) of digitized (scanned) documents.
The indexing and content processing of documents by applying several relevancy algorithms, regular expressions and dictionaries are computation intensive tasks. The execution of these steps needs the support of cloud infrastructure.
ACKNOWLEDGMENT
The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7) Capacities Programme under contract nr RI-283481.
Aron Szabo and Peter Tihanyi, EGroup, Hungary

Presenter: SZABO, Aron (E-Group)

Session Classification: Lightning Talks Industry Session

Contribution ID: **34** Type: **not specified**

Quality assurance in SCI-BUS project by applying agile testing practices

Wednesday, 5 June 2013 12:00 (15 minutes)

Agile testing is a software testing practice that follows the principles of agile software development. Agile development recognizes that testing is not a separate phase, but an integral part of software development, along with implementation. Testing and implementation are done incrementally and iteratively, building up each feature until it provides enough value to release into production. As the name refers, agile testing implies something to do very quickly, acting rapidly to changing requirements. Agile tests validate the business requirements as soon as possible. As the build is ready, testing is expected to get started and report the bugs quickly.
SCI-BUS is an eInfrastructure project sponsored party by the European Commission having as coordinator MTA SZTAKI (http://www.lpds.sztaki.hu/) and involving e-Scientiest representing different e-Science communities and a couple of industrial partners from different European countries. The SCI-BUS project aims at easing the life of e-Scientists by creating a new science gateway customisation methodology based on the generic-purpose gUSE/WS-PGRADE portal family that provides seamless access to major computing, data and networking infrastructures and services. The project creates and maintains a Liferay portlet repository that enables the quick creation of user specific customised gateways on top of the generic-purpose gateways providing seamless access to major European DCIs (Distributed Computing Infrastructures) including clusters, supercomputers, grids, desktop grids, academic and commercial clouds.
Having the WS-PGRADE framework the communities behind the customised gateways did (and do) not have to start to build their gateway from scratch. Rather they could use all the thoroughly tested internal services of P-GRADE, and after a relatively easy adaptation and extension procedure they were able to successfully create, deploy and operate their own customised gateways as a production service for their local and public DCI platforms. In that sense, the QA team of the project has a really challenging task. Particularly, the QA team has to adopt a customised agile methodology and the associated toolset to be applicable for a geographically distributed team coming from different areas (academic and business) developing distributed systems based on generic gateway service (gUSE/WS-PGRADE) with underlying grid and cloud infrastructures.
In this presentation, we will show how in SCI-BUS a 'lightweight'development and release process is used, taking into account the research and development nature of the project, the homogeneity of the portals to be implemented, and also the homogeneity of partners (academic and industrial). Agile concepts such as early builds and automated tests, continuous integration, collaboration facilities of project partners and managers and end-users are of particular importance. To realise these concepts, continuous integration processes and supporting tools are put in place. An advanced software build and test portal (etics3.4dsoft.hu) serves as a unified environment for building, testing, software packaging for the SCI-BUS partners'needs. Automated tests are also crucial for a distributed project where more than ten portal developers rely on a common framework (gUSE/WS-PGRADE portal).
4D Soft performs automated testing for SCI-BUS using a generic test automation framework for system and acceptance testing. The framework is based on the Robot Framework (http://code.google.com/p/robotframework/). This framework can be used for testing distributed, heterogeneous applications, where verification requires touching several technologies and interfaces. Users can create test cases using a simple syntax in a human readable format manner. Additionally, the framework capabilities can be extended by creating test libraries from the APIs of different test tools.

ACKNOWLEDGMENT
The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7) Capacities Programme under grant agreement no. RI-283481.

Presenter: FORGÁCS, István (4D Soft Kft.)

Session Classification: Lightning Talks Industry Session

Contribution ID: **35** Type: **not specified**

SCI-BUS and the SimBusPro Portal

Wednesday, 5 June 2013 11:45 (15 minutes)

I. INTRODUCTION

This abstract presents a portal solution for the optimization of business processes by using simulation. The objective is to utilize the SCI-BUS [1] framework in order to simulate business processes that have been modeled with Bonita Open Solution which is BPMN2 editor and has extended simulation capability. As of now portal only supports Bonita Open Solution [2]. Technically, the simulation feature via the SimBusPro portal can be an aspiring idea for other BPMN user communities.

II. TERMINOLOGY

BPMN2 (Business Process Model and Notation) is a modeling notation that allows users to define their own complex business processes as charts and gives clear understanding for all technical developers, business analysts and managers. BPMN2 is mostly graphical representation standard only for specifying business processes with business process diagrams [3]. BPMN2 doesn't have a built-in simulation feature yet.

III. ABSTRACT

SimBusPro Portal was developed to simulate "BPMN2 based Business Process Models" and there are two main benefits of the portal. The first one is to simulate a heavy load of business process models with massive computer power, which are generated from base BPMN2 models. The second one is to provide visual reports for the comparison between all business processes simulations of a model group. [4] In addition to functionalities of portal, this project gains value by means of the user-friendly interface that hides the DCI complexity. Portal currently uses Local Desktop Grid in company network as DCI.

Via SimBusPro Portal, users can import their pre-modeled Business Process Models in BPMN2 standard and easily produce multiple alternatives based on imported models by adjusting model parameters and then executes simulation. Simulation interpretation is also possible by using the generated reports that includes graphical-statistical charts. Users can store and keep track of those simulation reports and their own process models at personalized portal account which is connected to portal's database. Using this account, users can improve and work on their models anytime. Detailed usage of portal can be found in SimBusPro User Guide [5].

Auto-generated large number of model's simulation is time consuming and costly business. At this point, DCI and gUSE technologies help to reduce the simulation times and to get results quicker. This feature will help to reduce costs and increase efficiency.

Simsoft will sustain this project by making a deal with wide range customer portfolio, such as hospitals, factories, etc. is covered. Currently our sales team is visiting different types of costumers to introduce our project with presentations. In this scope, some of our projects are; modeling and optimizing processes of cement factory, supporting university project group who models and optimizes service routes of factory workers of a defense industry firm. Finally, we are planning to simulate departmental processes that include income and expenses department of a university hospital.

IV. ACKNOWLEDGMENT

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7) Capacities Programme under grant agreement no. RI-283481.

V. REFERENCES

[1] SCI-BUS, https://www.sci-bus.eu/

[2] Bonita Open Solution, www.bonitasoft.com/
[3] Object Management Group BPMN Standard, http://www.omg.org/spec/BPMN/2.0/PDF/
[4] SimBusPro Portal, http://simbuspro.com/liferay-portal-6.1.0/
[5] SimBusPro User guide, http://simbuspro.com/liferay-portal-6.1.0/documents/
Ferhat Kaya, Mert Aktaş, Sevilay İmre, İsmail Bıkmaz, Gökçe Yıldırım, Aydın Okutanoğlu and Veysi İşler, SimSoft, Turkey

Presenter: İŞLER, Veysi (Simsoft)

Session Classification: Lightning Talks Industry Session

Contribution ID: **36** Type: **not specified**

Overview about the SCI-BUS Project

Wednesday, 5 June 2013 13:30 (30 minutes)

Presenter: KACSUK, Peter (MTA SZTAKI)

Session Classification: Science Gateway Sustainability and Industry Utilization Workshop

Contribution ID: **37** Type: **not specified**

World Leading Storage Cloud at ETH Zurich

Wednesday, 5 June 2013 14:00 (45 minutes)

Presenter: STEIGER, Tilo (ETH Zurich)

Session Classification: Science Gateway Sustainability and Industry Utilization Workshop

Contribution ID: **38** Type: **not specified**

Introduction to Science Gateway Sustainability

Wednesday, 5 June 2013 14:45 (15 minutes)

Presenter: SUDHOLT, Wibke (CloudBroker GmbH)

Session Classification: Science Gateway Sustainability and Industry Utilization Workshop

Contribution ID: **39** Type: **not specified**

Podium Discussions of Speakers and Audience

Wednesday, 5 June 2013 15:00 (30 minutes)

Session Classification: Science Gateway Sustainability and Industry Utilization Workshop

Contribution ID: **40** Type: **not specified**

Discussions in Small Groups

Wednesday, 5 June 2013 16:00 (1 hour)

Session Classification: Interactive Session

Contribution ID: **41** Type: **not specified**

Summary of Discussions

Wednesday, 5 June 2013 17:00 (1 hour)

Session Classification: Interactive Session