

EGI Technical Forum 2012

Report of Contributions

Contribution ID: 0

Type: **Presentation**

Status of the StratusLab Sustainability Efforts

Thursday, September 20, 2012 2:25 PM (20 minutes)

Description of the work

StratusLab was a short, two-year European project that succeeded in developing an complete, open-source cloud distribution. Near the end of the project, the consortium developed a sustainability strategy that described the plans for continued support and development of the StratusLab software as well as the development tools and resources. Key parts of the strategy were 1) continued use of the StratusLab software on academic infrastructure, 2) incorporation of the software in commercial products, and 3) expanding use of the StratusLab cloud services in software engineering, particularly in DevOps settings. The presentation will describe the StratusLab sustainability strategy and its current status four months after the formal end of the project.

Wider impact of this work

The presentation will be of interest to running projects that are planning their own sustainability strategies. It will offer insight into what has and has not worked for StratusLab.

Printable Summary

StratusLab was a short, two-year European project that succeeded in developing an complete, open-source cloud distribution. The presentation provides a description of the StratusLab sustainability strategy and the status of current efforts.

Primary author: Dr LOOMIS, Charles (CNRS/LAL)

Co-authors: KATHYRN, Cassidy (Trinity College Dublin); DAVID, O'Callaghan (Trinity College Dublin)

Presenter: Dr LOOMIS, Charles (CNRS/LAL)

Session Classification: Sustainability of Technology Providers

Track Classification: Community and Co-ordination (Sergio Andreozzi: track leader)

Contribution ID: 1

Type: **Presentation**

FermiCloud - A Cloud Computing Infrastructure for Science at Fermilab

Tuesday, September 18, 2012 11:00 AM (30 minutes)

Description of the work

FermiCloud is a Cloud Computing Infrastructure as a Service (IaaS) deployment at Fermilab using the OpenNebula framework. The specific goals of the FermiCloud project are direct and indirect support for science at Fermilab together with collaboration and support for compatible cloud computing infrastructures.

The FermiCloud project has:

- Developed x.509 authentication plugins and contributed them to the OpenNebula project. These x.509 plugins are generally available in V3.2 of OpenNebula.
- Developed extensions to the Gratia accounting framework to support Cloud accounting.
- Demonstrated MPI performance using virtualised resources that are equivalent to “bare metal” MPI performance.

The FermiCloud project is in the process of:

- Deploying a robust, multi-user, distributed and replicated SAN to assure fault tolerant operations.
- Developing the necessary monitoring frameworks to assure continued operations.

Link for further information

<http://www-fermicloud.fnal.gov>

Wider impact of this work

The FermiCloud project has demonstrated operations at the forefront of delivering cloud computing capabilities to support physics research:

- By starting small, developing a list of requirements, building on existing Grid knowledge and infrastructure to address those requirements, FermiCloud has managed to deliver an Infrastructure as a Service cloud computing capability that supports science at Fermilab.
- The Open Science Grid software team is using FermiCloud resources to support their RPM “refactoring”.

Printable Summary

FermiCloud is the project to deploy an Infrastructure as a Service (IaaS) Cloud Computing capability in support of the Fermilab Scientific Program. FermiGrid currently totals 23 systems that can host as many as 384 virtual machine units. The work to improve the applicability and robustness of FermiCloud, together with specific accomplishments with respect to direct and indirect support of science achieved by the collaboration of personnel from Fermilab and the Korean Institute of Science and Technology Information (KISTI) using open source cloud computing frameworks will be presented.

Primary author: CHADWICK, Keith (Fermilab)

Presenter: CHADWICK, Keith (Fermilab)

Session Classification: Providing cloud services

Track Classification: Virtualised Resources: challenges and opportunities (Michel Drescher: track leader)

Contribution ID: 3

Type: **Presentation**

Ibercloud: orchestrating services to provide virtualized access to IberGrid

Tuesday, September 18, 2012 2:00 PM (20 minutes)

Description of the work

The objective of the IberCloud project is the deployment of a federated cloud IaaS testbed for scientific computing within the Ibergrid collaboration. In the interest of being as generic as possible, we pursue the integration of several popular cloud back-ends, namely resources behind OpenNebula and OpenStack deployments.

We have also analyzed the different APIs currently available in the community in order to identify the pros and cons of each of them (OCCI, EC2, libcloud and deltacloud).

As for user access, we provide a unique portal as interface to register, authenticate and accessing the services provided at each site contributing resources to Ibercloud.

We have also investigated the requirements of several communities of scientific users of the cloud in the Iberian Peninsula. In particular this applies to usage of generic software on single ore machines and parallel computing using MPI on Infiniband switches.

We describe as well several use cases in the area of Environment applicable to the ESFRI Lifewatch.

Link for further information

<http://cloud.ibergrid.eu>

Wider impact of this work

The main impact we expect is the broadening of the users base in Ibergrid. Currently we have a clear portfolio of users which are regularly using the Grid infrastructure in classic Grid. However there are a large number of users for which the Grid infrastructure is simply too much for their needs. The extra time needed to acquire expertise in using the Grid in a classic way does not compensate because their needs are much more focussed: deployment of databases on-the-fly, interactive access to resources for a few hours, just two name two examples.

We believe there is a large number of such customers in the scientific institutions of Spain and Portugal, which can be served if an appropriate on-demand cloud is setup as part of our federated infrastructure.

Printable Summary

Following users feedback, IberGrid is deploying a set of services to provide computing capabilities as configurable virtualized machines, available to the users on demand. In order to do that we have implemented together several OpenSource solutions available to the community. We have also investigated several use cases, among them interactive single-core access, as well as support

to parallel computing in low-latency clusters. We present the architecture as a cloud-like service, as well as our view in current issues, potential solutions, and future work.

Primary author: FERNANDEZ DEL CASTILLO, Enol (CSIC)

Co-authors: Mr LOPEZ GARCIA, Alvaro (CSIC); SIMON, Alvaro (FCTSG); ALFONSO, Carlos (UPV); FERNANDEZ, Carlos (FCTSG); MOLTO, German (UPV); BORGES, Goncalo (LIP); Dr BLANQUER, Ignacio (UPVLC); Dr CAMPOS, Isabel (CSIC); MARCO DE LUCAS, Jesus (CSIC); GOMES, Jorge (LIP); DAVID, Mario (LIP Lisbon); NUNEZ VEGA, Miguel Angel (CSIC); CABALLER, Miguel (UPV)

Presenter: FERNANDEZ DEL CASTILLO, Enol (CSIC)

Session Classification: Providing cloud services

Track Classification: Virtualised Resources: challenges and opportunities (Michel Drescher: track leader)

Contribution ID: 4

Type: **Workshop**

CHAIN Interoperability Workshop - Introduction to project

Wednesday, September 19, 2012 2:00 PM (30 minutes)

Printable Summary

The workshop aims to assess the progress made on interoperability with the demo of the Science Gateway and a feed-back from the VRCs.

A preliminary version of the Roadmap will also be presented. CHAIN aims to coordinate and leverage the experience of several e-Infrastructures Initiatives addressing various regions of the World and specifically those emerging in Asia and Africa, but also Latin America and Mediterranean.

The project is now focused on two milestones related to:

- Develop a Roadmap for the first model for world-wide collaboration in e-Infrastructures and Grids in particular;
- Demonstrate interoperability by means of a Science Gateway with applications from VRCs that can be considered of intercontinental span and/or willing to aggregate people and resources (computing, data repositories, etc.) from different regions of the world, to validate the above-mentioned model.

Description of the work

CHAIN aims to further coordinate and leverage the experience of several e-Infrastructures Initiatives addressing various regions of the World and specifically those emerging in Asia and Africa, but also Latin America and Mediterranean.

CHAIN focuses on the harmonised and optimised interaction model for e-Infrastructure and specifically Grid interfaces between Europe and the rest of the world.

The interoperations among regional infrastructures both at the operational and organizational levels will be one of the major goals.

The project, started on 1st December 2010, is now focused on two milestones related to:

- Develop a Roadmap for the first model for world-wide collaboration in e-Infrastructures and Grids in particular;
- Demonstrate interoperability by means of a Science Gateway with applications from VRCs that can be considered of intercontinental span and/or willing to aggregate people and resources (computing, data repositories, etc.) from different regions of the world, to validate the above-mentioned model.

Link for further information

<http://agenda.ct.infn.it/conferenceDisplay.py?confId=816>

www.chain-project.eu

www.gisela-grid.eu

Wider impact of this work

The workshop is expected to analyse the progress made since the previous workshops in the matching of User Communities' requirements and Resource Providers' plans demonstrating the usefulness of the Science Gateway approach towards interoperability. Hopefully it will help to pinpoint the remaining critical issues and the necessary technical developments, if any, that still need to be addressed in order to fulfil the requirements of the VRCs.

Primary authors: RUGGIERI, Federico (INFN - Roma Tre); MATYSKA, Ludek (CESNET); BARBERA, Roberto (University of Catania and INFN)

Co-authors: Prof. MARECHAL, Bernard (CETA - CIEMAT); PRNJAT, Ognjen; Dr GAVILLET, Philippe (CIEMAT)

Presenters: RUGGIERI, Federico (INFN - Roma Tre); MATYSKA, Ludek (CESNET); BARBERA, Roberto (University of Catania and INFN)

Session Classification: CHAIN Interoperability Workshop

Track Classification: Community and Co-ordination (Sergio Androozzi: track leader)

Contribution ID: 5

Type: **Internal Project Meeting**

EGI CSIRT (closed) meeting - agenda at <https://indico.egi.eu/indico/conferenceDisplay.py?confId=1160>

Monday, September 17, 2012 11:00 AM (1h 30m)

Description of the work

This is EGI CSIRT internal meeting

Link for further information

<https://indico.egi.eu/indico/conferenceDisplay.py?confId=1160>

Wider impact of this work

A closed meeting for EGI CSIRT members

Printable Summary

EGI CSIRT members will review past 6-months milestones and activities. The plan for the coming months will also be discussed and agreed.

Primary authors: KELSEY, David (STFC); Dr GABRIEL, Sven (Nikhef)

Presenter: Dr GABRIEL, Sven (Nikhef)

Session Classification: EGI CSIRT (Closed)

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 6

Type: **Training**

EGI-CSIRT Security Training Hands-on Wrap-Up

Friday, September 21, 2012 1:30 PM (1h 30m)

Description of the work

The outcomes of the tutorial hands-on will be discussed.

This should result in some best practices on how to reduce the attack surface of your site.

Wider impact of this work

Participants to the training hands-on are invited to discuss their findings. Other interested people are also welcome to this session

Printable Summary

This session is dedicated to the discussion and wrap-up of the training hands-on part I (tuesday) and II (wednesday) executed during the week.

Presenters: BRUNETTI, Riccardo (INFN); Dr GABRIEL, Sven (Nikhef)

Session Classification: EGI-CSIRT Security Training Hands-on Wrap-Up

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 7

Type: **Internal Project Meeting**

The EGI Software Vulnerability Group (SVG) - Introduction

Friday, September 21, 2012 11:00 AM (5 minutes)

Description of the work

Software Vulnerabilities may be reported by anyone to the EGISVG by e-mail to report-vulnerability@egi.eu. Vulnerabilities reported are investigated jointly by the developers and the SVG Risk Assessment Team (RAT). If they are valid and applicable to EGI the RAT then carries out a risk assessment, and sets a target date for resolution according to the risk.

Vulnerabilities are also detected by the pro-active examination of code known as 'vulnerability assessment'.

Various pieces of Grid Middleware are assessed using first principles vulnerability assessment techniques jointly developed by the University of Wisconsin and the Universitat Autònoma de Barcelona. Various pieces of Grid Middleware are or have been assessed using these techniques and any vulnerabilities found are addressed.

Vulnerability prevention is also carried out through developer education (mainly in EMI) and through considering what new software is allowed onto the EGI infrastructure.

Link for further information

<https://wiki.egi.eu/wiki/SVG>

Wider impact of this work

The purpose of the EGI Software Vulnerability Group is to eliminate existing vulnerabilities from the deployed infrastructure, primarily from the grid middleware, prevent the introduction of new ones and prevent security incidents.

As far as we are aware, and at the time of writing, no incidents have occurred due to vulnerabilities in Grid Middleware which tends to indicate that our procedures for handling vulnerabilities are effective.

Printable Summary

This session will report on the work and progress of the EGI Software Vulnerability Group (SVG).

A summary of the process for handling software vulnerabilities reported (which may be reported by anyone) and the current status will be included.

A report on the current status of 'Vulnerability assessment', which is the pro-active examination of code for vulnerabilities will also be presented.

Primary authors: HEYMANN, Elisa (UAB); CORNWALL, Linda (STFC)

Presenters: HEYMANN, Elisa (UAB); CORNWALL, Linda (STFC)

Session Classification: Operations

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 8

Type: **Session**

Security Policy Group and EGI Security Threat Risk Assessment public reports - Introduction

Tuesday, September 18, 2012 11:00 AM (5 minutes)

Description of the work

The Security Policy Group (SPG) is responsible for developing the policy needed to provide NGIs with a secure, trustworthy distributed computing infrastructure. The SPG output defines the behaviour expected from NGIs, Sites, Users and other participants to maintain a beneficial and effective working environment. At the time of writing there are 15 policy documents that have been developed by members of the SPG and approved by EGI Executive Board.

The EGI security threat risk assessment was carried out in several states. Firstly, a team was established for carrying out this work. Then the threats were selected and the current situation and mitigation for each threat was established as far as effort allowed. Then the risk was computed by asking everyone to give their ratings on the likelihood and impact. Guidelines for likelihood and impact were given. The findings were then reported.

Link for further information

<https://wiki.egi.eu/wiki/SPG>

Wider impact of this work

Security policies are available which have been widely adopted both within European Grid projects and in the wider Grid community.

For the EGI security threat risk assessment a comprehensive review has taken place of security in the EGI infrastructure. A number of risks have been identified where priority is needed for their mitigation.

Printable Summary

The Chair of the EGI Security Policy Group (SPG) will report on the current status of security policies and plans for future work.

A report will be presented on the EGI Security Threat Risk Assessment, which was carried out during the first half of 2012. Members of the EGI Security Assessment team identified 75 threats to the EGI infrastructure and considered the risk associated with each of those threats. The method used to compute the risk and the security threats identified as having the highest risk will be presented.

This session is open to anyone and will be of particular interest to those interested in security and what the security threats are to the EGI infrastructure. It will also be of interest to those who wish to know what security policies are in place or planned and the behaviour expected in the EGI environment.

Primary authors: KELSEY, David (STFC); CORNWALL, Linda (STFC)

Presenters: KELSEY, David (STFC); CORNWALL, Linda (STFC)

Session Classification: EGI InSPIRE

Track Classification: Community and Co-ordination (Sergio Andreozzi: track leader)

Contribution ID: 9

Type: **Workshop**

Authentication and Authorization Infrastructure workshop

Wider impact of this work

This workshop is meant to be a meeting point between user communities, infrastructure providers and technology providers. Being hosted within the EGI Technical Forum programme, the workshop is a good opportunity to steer the discussion towards the point of view of the European Grid Infrastructure and the users who are accessing its resources.

Description of the work

This workshop is meant to be a meeting point between user communities, infrastructure providers and technology providers.

High level overview of the proposed sessions:

First part

- Presentations by NGIs and EGI.eu about the current AA solutions used within the existing middleware and emerging federated cloud platforms of EGI.
- Presentations by representatives scientific communities about the use cases and requirements for the integration of diverse AA solutions to the EGI infrastructure.
- Presentations by technology providers and representatives of community support teams about emerging AA solutions and software tools that already exist or could be integrated with the EGI infrastructure platform. (for example Identity federations; OpenID; science gateway frameworks; online certificate storages, robot certificate providers)

Second part:

Open discussion focused on identifying commonalities in requirements, in solutions, analysing any gaps between use cases and state of the art; defining an action plan for the more harmonised adoption of emerging AA solutions within the production infrastructure.

The presentation and discussion sessions of the workshop will present state of the art AA solutions from the community, will refine requirements for AAIs and will define next steps for the community towards a more harmonised adoption of emerging AA solutions within EGI.

Printable Summary

Resources of the European Grid Infrastructure provide services based on the main grid middleware platforms. These services use proxy certificates for user authentication and authorisation (AA) purposes.

A growing number of the existing and potential new user communities of EGI consider personal certificate based access as a main barrier of the infrastructure uptake. Some of these communities –together with their supporting NGIs, NRENs and scientific projects –developed ad-hoc solutions to simplify or hide the certificate based AA infrastructure (AAI) for the users

Worldwide coverage, scalability, robustness, integrate-ability with current and emerging EGI platforms and simple usage are the main criteria for a AAI to be adopted within the European Grid Infrastructure. The workshop will bring together representatives from the three main actors of the EGI ecosystem: existing and potential user communities and their support teams, platform technology providers and resource providers.

Primary authors: Dr SIPOS, Gergely (EGLEU); SOLAGNA, Peter (EGLEU)

Presenters: Dr SIPOS, Gergely (EGLEU); SOLAGNA, Peter (EGLEU)

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 11

Type: **Session**

Marketing and Communications Strategies

Tuesday, September 18, 2012 11:00 AM (10 minutes)

Printable Summary

Marketing and communicating your work effectively is a key aim for EC-funded projects. With Horizon 2020 coming onstream in 2014, funders are increasingly focused on getting the research they support noticed. Marketing strategies need to focus on establishing engagement with a range of audiences, including scientists, research infrastructures and policy makers. This interactive session focuses on building outreach strategies that will open up two way channels with different audiences and encourage dialogue. This is particularly important when addressing new audiences, such as new users of e-infrastructures like the ESFRI projects, in order to ensure that their requirements can be gathered and to promote new tools and services available to them.

Description of the work

This session will include practical advice on communicating the results from research carried out on a national, international or global basis. The main focus will be on communicating with new users, and the work carried out to date in reaching out the ESFRI communities. In particular, the session will focus on work with the ENVRI project, to gather a network of contacts within the environmental sciences research infrastructures that will enable outreach to their diverse user communities.

The session will also cover the use of Virtual Teams to enhance marketing and communications activities. Previous VTs have included a team gathering NGI-related content for the EGI website, and a team that developed a plan for EGI's attendance at the European Geophysics Union. The session will propose new ideas for communications-focused VTs in the third year of EGI-InSPIRE and outline plans for implementation.

Wider impact of this work

The workshop is aimed at researchers and members of the community who need to communicate their work to targeted audiences. It will be of particular interest to those who are new to communicating research in these areas, or who have particular problems in getting their message out that they would like to share with other communications specialists. We hope to see a good attendance by the NGI International Liaisons and communications teams from collaborating projects, in particular ENVRI and the other ESFRI projects.

Primary authors: GATER, Catherine (EGLEU); ONEILL, Neasan (EGLEU)

Presenters: GATER, Catherine (EGLEU); ONEILL, Neasan (EGLEU)

Session Classification: Marketing and Communications Strategies

Track Classification: Community and Co-ordination (Sergio Andreozzi: track leader)

Contribution ID: 12

Type: **Presentation**

DIRAC: breaking the frontiers of the Grid

Thursday, September 20, 2012 4:00 PM (20 minutes)

Description of the work

The DIRAC is an open source software project currently supported by CPPM, UB and CERN, that provides an unique framework for distributing computing, able to integrate in a single system all the activities of large international scientific collaborations, and providing transparent access and interoperability among all their distributed resources.

Starting from the LHCb use case, where it was initially developed, DIRAC has proven to integrate general purpose EGEE grid resources with commercial Amazon cloud resources and private HTC solutions in order to optimize the overall cost of large scale projects like Belle II.

This work has been continued and DIRAC can connect now to OCCI/OpenNebula and RESTful/CloudStack (OpenStack is under development at the moment), providing much wider integration with private, public and hybrid cloud solutions. In combination with CernVM and cvmfs solutions, it provides a very flexible platform for general purpose usage that will be validated within the EGI FedCloud TaskForce effort.

At the same time, an integrated web portal interface, a versatile command-line and programatic python API, and current new developments allow now to connect the DIRAC WMS and DMS engines, to existing third party generic application portal solutions.

Effort is now devoted to handle the use case of small user groups accessing a common shared infrastructure like the EGI NGIs, providing a single and easy to use entry point to all user of general purpose VOs. This includes the possibility for larger groups or advanced users to easily extend the basic DIRAC functionality reusing the extensions provided by larger collaborations already making use of DIRAC.

This integration effort, in all senses, and the opportunity to transfer to a much wider audience the accumulated experience from the Particle Physics community is an unique opportunity to return to the society part of the effort that in the last years has been dedicated to LHC experiments.

Link for further information

<http://diracgrid.org>

<https://indico.cern.ch/search.py?p=DIRAC&confId=149557&collections=Contributions>

Wider impact of this work

Using DIRAC to transfer the knowledge and expertise that Particle Physics communities have accumulated over the last year of extensive use of distributed computing, by providing an easy to use complete solution to end users and smaller communities is expected to boost the usage of these expensive computing resources towards a much wider base of users, currently discouraged in many cases by the existing spartan middleware tools.

The unique capability to integrate resources, users and tools make DIRAC a new “interware” solution that should allow to break through the frontier that currently separates non-expert users

(or users without close connections to experts) and the usage of high throughput distributed computing solutions that would allow these researches to boost their scientific work and its impact.

Printable Summary

The DIRAC project provides a framework for building ready to use distributed computing systems. It has been proven to be a useful tool for large international scientific collaborations like LHCb, ILC, Belle II, and others, integrating in a single system all their computing activities and distributed computing resources: Grids, Clouds and HTP clusters.

At the same time DIRAC offers an unique solution for NGIs that need to provide simple and transparent access a wide variety of users and users communities with different levels of ICT expertise. This is currently being explored as pilot case in FRANCE and IBERGRID.

Additionally, DIRAC provides unique tools like its File Catalog that is at the same time replica, meta-data and provenance catalog, integrated in a single service and closely connected to the DIRAC WMS and DMS. Or a newly developed RESTful API that it is being tested against existing Application specific portal frameworks like gUSE/WS-pGRADE or InSilicoLab.

Primary author: Dr GRACIANI DIAZ, Ricardo (University of Barcelona)

Co-authors: CASAJUS, Adria (CERN); Dr TSAREGORODTSEV, Andrei (Centre de Physique des Particules de Marseille, CPPM); Dr MENDEZ, Victor (Port d'Informació Científica, PIC)

Presenter: Dr GRACIANI DIAZ, Ricardo (University of Barcelona)

Session Classification: VREs - Community Contributions

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 14

Type: **Session**

EMI Security for Grids and Clouds - aims of double session

Thursday, September 20, 2012 2:00 PM (5 minutes)

Description of the work

The work in the Security Area that is aimed towards the future includes:

- The Security Token Service (STS) where a UNICORE client is being developed. We are hoping to show either a preview or a demo of this client with the STS service.
- The Common AuthN libraries (CANL) have been developed and are being integrated into EMI services.
- The Hydra key storage system, used by Biomedical researchers, will have undergone changes on both the client and service-side. These changes incorporate newer technologies and experience from other Grid services. A distributed key storage system can be considered essential for data storage on “unknown” Cloud resources.
- The Argus Environment Execution Service (EES) provides connection between the security of the Grid world and the flexibility of the Virtual Machine (or Cloud) environment. The EES may exist with the Argus Authorization service and ensures that an appropriate site-specific execution environment for a virtual machine is procured.

Wider impact of this work

The EMI Security services will need to be understood by users (system administrators and researchers) if they are to continue in the more virtualized computing environment after traditional Grid projects finish.

Printable Summary

Within the the EMI project the Security Area work has started in May 2010 in order to achieve the project goals.

These goals include easing the usage of Grid and Cloud resources by researchers and unifying the EMI software stack.

This session will present some of the work performed on the way to achieving these goals. These presentations are to be user-oriented and include the STS UNICORE Client, Hydra key storage service, Argus Environment Execution Service and Common Authentication libraries. This session is aimed at middleware users and software developers who will gain an understanding of how this work can be used.

Primary author: WHITE, John (University of Helsinki, Finland)

Presenter: WHITE, John (University of Helsinki, Finland)

Session Classification: Resource Infrastructure Services

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 15

Type: **Poster**

An example of GSG tools integration: Starlight Spectral Synthesis Code.

Printable Summary

Starlight Spectral Synthesis Code is widely used by several astrophysics research groups specialized in extragalactic astronomy to estimate physical properties of galaxies. Due to this situation, Grid Computation and Extragalactic Astronomy Departments of the IAA have collaborated to develop a toolbox that helps scientists use Grid Computing intensively and to optimize its use. In this article we present all the migration details of the Starlight Spectral Synthesis Code to Grid infrastructure. Furthermore we present the study of the optimizations made to use this code in Grid Computing, showing improvements, problems and anecdotes found while using this infrastructure.

Description of the work

Starlight Spectral Synthesis Code is used intensively, so we are looking for solutions to decrease the computation time in every spectra analysis. We have estimated around a thousand spectra for each galaxy, and 30 minutes for each spectrum analysis, so to analyze a complete galaxy we will spend a total computation time of 500 hours. It is pretended to study thousands of galaxies so it is not possible to run the analysis in a sequential computation environment. This is the reason to use a distributed computation platform. There are several alternatives to run the Starlight Spectral Synthesis Code in a distributed infrastructure, as HPC, Volunteer Computation, Cluster Computing, or Grid, because it is very easy to chop it into small pieces of code. However, Grid Infrastructure is the best option due to it comprises several nodes, which contain several available slots where these code pieces can be run. Starlight Spectral Synthesis Code is not a completely stable software due to the investigation requirements: there are new items to investigate as time goes by so the software has to be modified or extended, so new versions appear constantly. Moreover, this code was modified and parallelized to be used under Grid computing. Due its unstability each new version has to be ported to Grid Computing. IAA Grid Computing Department can do it immediately because of their proximity to IAA Extragalactic Astronomy Department, so the Starlight Spectral Synthesis Code is contained in the IAA Grid node. But the portability to the rest of Grid nodes is more complicated. To do so we have studied another solutions.

Link for further information

<https://grid.iaa.csic.es>

Wider impact of this work

Is it worth running Starlight Code in Grid environment?

Considering the runtime: The total running time improves if compared with the sequential way, including implicit times

of overhead in Grid Infrastructure.

Considering the reliability: GSC tool utility to detect fail or aborted jobs makes more robust Starlight portability avoiding losses in the analysis results.

Considering the ease of use: Due to StarLight integration through GSG tools, an easy management has been achieved using directories, so the final user can find the data easily.

Considering the development of Starlight Code in Grid Infrastructure: The Starlight Code has been integrated with the modular GSG tool. This modularity makes easier to add new functionalities and to modify the existing ones, leaving the door opened to further development of StarLight Code in Grid.

But nothing is perfect: On the other hand there are potential improvements especially when we look for the total utilization of Grid Infrastructure.

Primary author: RAMON RODON ORTIZ, Jose (CSIC)

Co-authors: Mr PÉREZ, Enrique (IAA-CSIC); Mr RUEDAS, José (IAA-CSIC); Mr ORTEGA, Julio (UGR); Mr GARCÍA, Rubén (IAA-CSIC); Mrs PASSAS, Varo (IAA-CSIC)

Presenter: RAMON RODON ORTIZ, Jose (CSIC)

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 17

Type: **Presentation**

EGI IPv6 activities: status and planning

Tuesday, September 18, 2012 2:00 PM (20 minutes)

Description of the work

EGI set up a distributed IPv6 testbed to directly assess the level of IPv6 compliance of the UMD family middleware products.

In collaboration with the HEPiX IPv6 Working Group, direct testing of the behavior of the UMD middleware has been assessed.

Installation, Configuration and Testing of basic functionality of ARC and gLite has been assessed using the IPv6 protocol.

Installing and configuring the services using IPv6 required some specific steps enabling the proper execution of all required steps.

Also, the verification of the proper response by the system using Dual Stack has been carried out.

Starting from Workload Management and InfoSysy related components, then moving to other functional domains, an analysis of the overall level of IPv6 compliance in the middleware has been carried out.

Test reports have been written, listing required changed and identified issues.

Link for further information

<http://net.egi.eu>

Wider impact of this work

IPv6 compliance at large seems at hand, not far from being fully achievable.

IPv6 is supported by essentially all Operating Systems, it has been enabled by major internet service providers, it is supported by essentially all European NRENs and the European Backbone GEANT. The situation in USA is similar.

The UMD middleware family products are already with some tech inside well likely to be full IPv6 compliant. What is missing is a final grip to certify the middleware and have it tested in IPv6.

Work has to be done also at the operational tools level to test and verify the proper functioning of the tools using IPv6.

Coordination at all levels with collaborating projects and initiative is vital and fundamental.

An hoped impact for this work is to trigger a general discussion between EGI, W-LCG, EMI, IGE and other tech providers to permanently include IPv6 in the loop of certification and testing&validation activities.

Printable Summary

This will be a joint presentation by the IPv6 team of EGI, open to contributions and inputs by collaborating projects, on the overall status of the IPv6 activities, the identified issues, the level of compliance observed so far in the middleware and the required involved external dependencies. It will also highlight the current testbed infrastructure, and wrap up the outcome of the testing activities.

Primary author: Dr REALE, Mario (GARR)

Presenter: Dr REALE, Mario (GARR)

Session Classification: Network Support

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 18

Type: **Presentation**

GWpilot: a personal pilot system

Thursday, September 20, 2012 11:30 AM (30 minutes)

Description of the work

By using pilot-job systems, Grid execution performance can be improved. This is so because it is possible to somehow reserve the available resources by means of submitting regular Grid jobs (the pilots) that lately allocate the application tasks for doing the real calculation. Nevertheless, the current implemented systems usually lack some features such as user-sharing and easy-to-installing capabilities or standard interfaces that could prevent their deployment.

GWpilot is a new developed general purpose framework system that overcomes the limitations or drawbacks derived from the aforementioned approaches. In addition to the common functionalities already implemented in other systems to overcome remote queues, correctly fit tasks to pilots or discard bad resources, it can also abandon a passive role in order to effectively coordinate the pilot-task matchmaking with advanced scheduling techniques such as pre-allocation, reservation, data-allocation awareness, etc. This is done by a more flexible, extendable and tuneable methodology based on embedding the pilot system in the GridWay metascheduler.

As consequence, the framework is suitable in both user and VO levels since it is compatible with standards such as DRMAA, OGSA-BES or JSDL and it permits a fine troubleshooting, this is, any user belonging to any scientific discipline or association can use it. Therefore, GWpilot can be used in any kind of HTC calculation since legacy applications can be easily adapted to these interfaces. Moreover, it enables a multi-level scheduling ecosystem that allows application-oriented systems (workflow managers) or self-schedulers to profit from the customized characterisation of resources supplied by pilots.

As use cases, several official EGI applications are employed to demonstrate the suitability of GWpilot as well as the performance gain obtained.

Link for further information

<http://www.ciemat.es/portal.do?IDR=343&TR=C>

<http://www.gridway.org/>

Wider impact of this work

GWpilot system is a framework specially devoted to single users and institutions who are interested in performing massive High Throughput Computing calculations. It can be used in principle with any Grid application and/or VO, so it is ideal for users who in principle do not count on ad-hoc systems. In this way, it is expected a high impact in the scientific community due to it offers standardised interfaces that allow a straightforward execution of previously gridified applications. It also provides an easy and standalone installation that facilitates its quickly deployment and use.

Last, in addition to the common advanced scheduling features already present in other pilot systems, GWpilot accomplishes the performance requirements of users, institutions and even VO levels.

Printable Summary

Grid overheads can be overcome by pilot-job systems. Huge collaborations rely on ad-hoc ones that fit their need, but a researcher working on his own cannot rely on a general purpose application. This work is devoted to this kind of standalone users, either inexpert or highly skilled ones. Some frameworks offer pilot-job advantages to conventional users, but they usually lack some features such as user-sharing, ease of installation or they suffer from standard Grid interfaces that could prevent their deployment. GWpilot is a newly developed general purpose framework based on GridWay that offers functionalities already implemented in other pilot systems to overcome remote queues, correctly fit tasks to pilots or discard bad resources, but also coordinates the pilot-task matchmaking with advanced scheduling techniques. The system offers CLI tools and Grid standard interfaces that make it suitable for running legacy applications or being coupled to other systems such as workflow managers

Primary author: Mr RUBIO-MONTERO, Antonio J. (CIEMAT)

Co-authors: HUEDO, Eduardo (UCM); MAYO-GARCIA, Rafael (CIEMAT)

Presenters: Mr RUBIO-MONTERO, Antonio J. (CIEMAT); MAYO-GARCIA, Rafael (CIEMAT)

Session Classification: Resource Infrastructure Services

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 19

Type: **Presentation**

Future Proof storage with DPM

Tuesday, September 18, 2012 11:00 AM (20 minutes)

Wider impact of this work

Data access on the grid has traditionally meant custom protocols and custom software blocks. The work presented here tries to break with this tradition, by enabling access using standard protocols and opening DPM for integration with standard building blocks both for frontends and storage backends.

The goal is to reduce the dependencies on custom software and reuse as much as possible the existing open source tools fulfilling the same requirements, such as Apache2, Memcache, Hadoop, S3 based cloud providers, among many others.

Description of the work

HTTP is the most widely used web access protocol. With DAV, it provides a standards based mean of accessing data on the grid, a valid alternative to traditional grid protocols like GridFTP.

NFS is a popular remote access protocol. With version 4.1 it introduces sessions, strong authentication, bulk operations, among many other features. With the addition of pNFS it also separates metadata from physical data access, while enabled parallel reading and writing of the data. This puts it as a very strong solution for POSIX access to data on the grid, offering additional benefits like free clients and kernel level caching.

DPM now provides frontends exposing both of these protocols, and we have been working on providing excellent performance and validating our software against popular scientific frameworks like ROOT or experiment specific software.

At the same time, and triggered by some performance limitations found while evaluating the analysis use cases against DPM, we have done a significant refactoring of the software. The result is a new plugin based library called 'dmlite', which exposes interfaces for namespace management, pool management, and access to data.

We present the global architecture of the new libraries, the plugins developed and the ones in development. These include memcache for scalability and excellent performance of the nameserver, hadoop as an alternative solution for pool backends, S3 for extending DPM storage to the cloud, among others.

We also present extensive performance results of each of these parts of DPM, including comparisons against previous setups where appropriate.

Printable Summary

The Disk Pool Manager (DPM) is a lightweight solution for grid enabled disk storage management. Operated at more than 240 sites it has the widest distribution of all grid storage solutions in the WLCG infrastructure. It provides an easy way to manage and configure disk pools, and exposes multiple interfaces for data access (rfio, xroot, nfs, gridftp and http/dav) and control (srm). During the last year we have been working on providing stable, high performant data access to our

storage system using standard protocols, while extending the storage management functionality and adapting both configuration and deployment procedures to reuse commonly used building blocks. In this contribution we present extensive performance results of the work recently accomplished, we summarize bottlenecks discovered during this analysis and the solutions followed and we present new components opening DPM to integration with other popular storage systems.

Primary author: ROCHA, Ricardo (CERN)

Co-authors: Mr ALVAREZ AYLLON, Alejandro (CERN); FURANO, Fabrizio (CERN); KEEBLE, Oliver (CERN)

Presenter: KEEBLE, Oliver (CERN)

Session Classification: New middleware products

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 20

Type: **Presentation**

Prototyping a Competence Center for the Computational Life Science Communities - Experiences from the ScalaLife Project

Tuesday, September 18, 2012 2:00 PM (20 minutes)

Description of the work

The Life Sciences have rapidly become one of the major beneficiaries of the European e- Infrastructures, placing a growing demand on the capabilities of simulation software and on the support services. The ScalaLife project has set to address some of the specific problems associated with this growth, acting along two distinct and complementary directions.

On the one hand, the project is implementing new techniques for efficient small-system parallelization, establishing open software standards for data storage and exchange, and developing new hierarchical approaches. The latter are explicitly based on ensemble and high-throughput computing for new multi-core and streaming/GPU architectures. Automated hierarchical parallelization will be of particular importance for Grid applications with running on many-core and accelerator systems.

On the other hand, the project is committed to the long- term support of the Life Science users and communities, providing both training and expert advice. ScalaLife is documenting and developing training material for the new techniques and data storage formats implemented by the project. The project has created a pilot for a cross-disciplinary Competence Centre, which enables the Life Science community to effectively exploit the key applications developed as part of the project on the existing European e-Infrastructures, including Grid resources provided by EGI.

Link for further information

www.scalalife.eu

Wider impact of this work

The experiences gathered in the development of the ScalaLife Competence Center provide valuable foundations for future efforts to establish similar centers in other areas of computational science. In addition, the center fosters the interactions and exchange of expertise between the HPC and EGI communities.

Printable Summary

Life Science research has become to a large extent dependent on computing resources for simulation and data analysis. Despite the advances in software development, which are pushing Life Science codes towards exascale capabilities, efficient utilization of those resources is often not achieved. Some of the reasons include a lack of HPC awareness and expertise among regular users, a steep learning curve for powerful software features, a lack of documentation of best practices as well as a disconnect between code developers and users.

The ScalaLife project (scalalife.eu) has engaged in prototyping a cross-disciplinary Competence Center as a long-term structure for:

- Maintenance and optimization of key Life Science software
- Provision of training and support infrastructure
- Development of an adequate framework and associated policies to foster collaboration among software developers and users

Primary author: APOSTOLOV, Rossen

Presenter: APOSTOLOV, Rossen

Session Classification: Experience in community building

Track Classification: Community and Co-ordination (Sergio Andreozzi: track leader)

Contribution ID: 21

Type: **Presentation**

VENUS-C Interoperable Toolkit v2.0

Wednesday, September 19, 2012 4:30 PM (30 minutes)

Description of the work

The objective of the OVF4ONE tool is to render the Opennebula cloud inter-operable and open and to enable horizontal scaling and migration of infrastructure services across heterogeneous cloud environments.

OVF4ONE tool provides an OCCI standard interface on top of Opennebula cloud with the OCCI rendering defined in OVF service description format.

In detail, OVF4ONE converts the OCCI Restful calls to Opennebula OCA(OpenNebula Cloud API) calls and the OVF service descriptions to the Opennebula Virtual machine descriptions. The work involves mapping of the Opennebula template parameters to the OVF service Descriptor and thus allowing the Opennebula VM template to be defined in OVF packaging format. OVF is extensible and thus the OpenNebula template values that could not be mapped to OVF are placed in the "Product properties" section of the OVF file.

OVF4ONE enables infrastructure services migration and scaling between Opennebula and other cloud providers.

This project has been realised as part of Venus-C EU project focused on easing migration across heterogeneous target platforms.

Link for further information

<http://opennebula.org/software/ecosystem/ovf4one>

Wider impact of this work

Provides a standard interface for accessing infrastructure as a service offerings for the Opennebula cloud and thus facilitates scaling, migration of services across Opennebula and other cloud infrastructures. Integrates cloud standards OCCI and OVF thus enhancing the advantages of OCCI interface by providing OVF service descriptions for packaging, configuration and distribution of virtual machines, that in turn helps in rendering the virtual resources portable across various cloud infrastructures.

Printable Summary

With the rapid adoption of virtualization in the arena of cloud a need for an open and inter-operable cloud and a standard way to package and distribute virtual machines is imminent. Open Cloud Computing Interface (OCCI) provides an API that allows for the runtime management of infrastructure as service offerings. Open Virtualization Format (OVF) was created in order to provide a platform independent, efficient, extensible, and open packaging and distribution format for virtual machines.

The idea behind OVF4ONE is to render the Opennebula cloud interoperable with other cloud offerings by providing an OCCI interface that allows OVF service descriptions.

Primary author: Mr UBERTO LAURI, Gian (Researcher)

Co-author: Ms RAMASAMY BALRAJ, Kanchanna (Junior Researcher)

Presenter: Ms RAMASAMY BALRAJ, Kanchanna (Junior Researcher)

Session Classification: Cloud Management Solutions

Track Classification: Virtualised Resources: challenges and opportunities (Michel Drescher: track leader)

Contribution ID: 22

Type: **Workshop**

Tutorial on NGI Service Portfolio Design

Thursday, September 20, 2012 4:00 PM (1h 30m)

Description of the work

Since the beginning of the EGI community, there has been an overall move from a project based structure to a more sustainable, self supporting approach that can to compete with commercial alternatives. This change has included the formation of National Grid Infrastructures, which are expected to be the primary entities that service user needs through signing agreements to provide services to Virtual Organisations and Virtual Research Communities.

NGIs have evolved out of the national groups involved in Grid computing over the last decades, and their variability represents the different size, approach, funding situation and user communities seen in the Grid community of different European countries. A mix of their different states, organisational structures and goals means that while all active NGIs work hard to provide value to a series of customers, it is often difficult to define either precisely who the customer segments they serve are or what set of services they consume. In many cases an NGI will provide dozens of technical services to a range of customers but they will not be relationally grouped.

This session will help NGIs to define their services and customers in a formal manner. This will not involve changing the technical service they provide, simply to use techniques from commercial IT Service Management (ITSM) to describe services and customers in a way that supports effective provision and management of services.

The session will be run by experts fro the gSLM and FedSM projects as well as EGI.eu staff, to demonstrate a process for defining a service portfolio. This will be illustrated with examples and will encourage participants to raise issues and concerns they have from their own NGI experience.

At the conclusion of the session, NGI participants should have an understanding of how to begin work on a service portfolio, as well as knowing whether they can seek support and expertise needed to complete this task.

Link for further information

www.gslm.eu

Wider impact of this work

This session will compliment the session on sustainability of National Grid Infrastructures, and continue the collaboration of EGI with the gSLM and FedSM projects on improving management and sustainability of NGIs and the EGI ecosystem. It will provide practical assistance for NGIs, to make the challenge of designing a service portfolio an approachable task. Through working on service portfolios, many fo the questions relation to business models and service level management identified as issues for EGI should also be raised, so this session should act as a first step in encouraging NGIs to tackle these issues in a formal manner.

Printable Summary

National Grid Infrastructures are currently in a transition phase from startup states to sustainable organisational structures. A key part of this is precisely defining what services they offer and who will make use of them, This is necessary as not only are NGIs variable in goals, structure and resources, but also because they have often grown out of other organisations. This means that they have often evolved to provide value to many groups without clearly defining either the groups or services.

This session will use knowledge and approaches from commercial IT Service Management and experts from the gSLM & FedSM projects to help NGIs understand who to clearly define service portfolios. This will support later development of business models and structures for efficient and cost effective management of NGIs. The session will work through the process of designing a service portfolio and provide a schema to support NGIs is carrying out the same process themselves.

Primary authors: APPLETON, Owen (Emergence Tech Limited); ANDREOZZI, Sergio (EGI.EU); SCHAAF, Thomas (Ludwig-Maximilians-Universität München)

Co-author: Mr SZEPIENIEC, T. (CYFRONET)

Presenters: APPLETON, Owen (Emergence Tech Limited); ANDREOZZI, Sergio (EGI.EU); Mr SZEPIENIEC, T. (CYFRONET); SCHAAF, Thomas (Ludwig-Maximilians-Universität München)

Session Classification: Training on Service Portfolio design for NGIs

Track Classification: Community and Co-ordination (Sergio Andreozzi: track leader)

Contribution ID: 23

Type: **Presentation**

perfSONAR MDM for the European grid and beyond: status and news.

Tuesday, September 18, 2012 4:00 PM (30 minutes)

Wider impact of this work

This presentation focuses on the monitoring infrastructures and activities that would affect the grid community: support to service availability, help to evaluate transfer cost, reliability assessment

Printable Summary

Grid computing in a geographically distributed environment is based on efficient, reliable networking.

The grid initiatives EGI supports and coordinates span multiple network domains, requiring then a monitoring system which is able to cross network borders.

perfSONAR MDM is an ideal network monitoring solution since it would allow a seamless coverage over multiple domains, is able to deal with different infrastructures and has a flexible and intuitive UI.

PerfSONAR MDM is currently used in 15 NRENs in Europe and an a liaison activity is currently ongoing talking to grid initiatives in Latin America (GISELA) and research projects worldwide.

Description of the work

perfSONAR MDM (Multi-Domain-Monitoring) is the multi-domain monitoring service for the GÉANT Service Area (GSA). It enables NREN NOCs and PERTs to collaborate in providing seamless network performance, working together to identify, prevent and solve performance issues for network users.

perfSONAR MDM provides easy, transparent end-to-end monitoring, giving its users - primarily NOC and PERT engineers - access to network measurement data from multiple network domains. Monitoring data is collated from all those domains where the perfSONAR service is deployed, in order to visualise network characteristics, present the information in a standardised format and enable troubleshooting of related issues.

Link for further information

perfsnar.geant.net

Primary author: VICINANZA, Domenico (DANTE)

Presenter: VICINANZA, Domenico (DANTE)

Session Classification: Network Support

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 24

Type: **Presentation**

Talk 1

Track Classification: Community and Co-ordination (Sergio Andreozzi: track leader)

Contribution ID: 25

Type: **Workshop**

Integrating publicly owned e-Infrastructures with commercial clouds: Business Models

Description of the work

EGI has developed a strategic framework through 3 key documents: EGI Strategic Plan, Evolving Business Models and Sustainability Plan defining roles, responsibilities, long-term directions and a business model framework for the EGI community. The aim of this session is to discuss specifically how EGI interacts with commercial cloud providers by defining potential business models to support it. The focus will be on exploring ideas for public-private partnerships, understanding the financial implications within the research community, exploring resource allocation options and defining methods of compensation, legal and contractual constraints, compliance, accounting/billing models and service level agreements.

Link for further information

<https://documents.egi.eu/document/1040>; <https://documents.egi.eu/document/960>; <https://documents.egi.eu/document/114>

Wider impact of this work

This session is key to kick-off a strategic discussion in the wider context of integrating e-infrastructures with commercial cloud providers. Exploring business models ideas is an important pillar together with the policy issues and technical aspects being considered in other sessions.

Printable Summary

The aim of this session is to kick-off the discussion on potential business models suitable for enabling commercial cloud providers to operate within the open EGI ecosystem. The focus will be on exploring ideas for public-private partnerships, understanding the financial implications within the research community, exploring resource allocation options and defining methods of compensation, legal and contractual constraints, compliance, accounting/billing models and service level agreements. The information gathered will be part of a written summary following the workshop outlining potential business models and specific areas where further work is required. The topics to be discussed is of interest for policy-makers, funding agencies, commercial organisations and technology and infrastructure providers and operators.

Primary authors: ANDREOZZI, Sergio (EGLEU); NEWHOUSE, Steven (EGLEU)

Co-authors: MARINOVIC, Damir (EGLEU); HOLSINGER, Sy (EGLEU)

Presenter: NEWHOUSE, Steven (EGLEU)

Track Classification: Community and Co-ordination (Sergio Andreozzi: track leader)

Contribution ID: 26

Type: **Workshop**

Sustainability of National Infrastructures

Description of the work

In order for EGI to become sustainable, it is essential that also the national grid infrastructures (NGIs) achieve sustainability. The aim of this session is to continue the discussion on this matter by addressing key sustainability issues such as costs, cost models, service portfolio, business model and service management structures. The session will open with a presentation on a checklist for sustainable NGIs that will help to stimulate discussion and sharing of best practices or approaches across NGIs. Further presentations will develop some specific aspect and through a final panel discussions, key results will be identified and agreed together with future actions. Presentations and discussions build on the three key EGI-InSPIRE deliverables (EGI2020 Strategy, Evolving the EGI Business Model, and EGI Sustainability Plan) and on the results and support from collaborating projects (e-FISCAL, gSLM, FedSM).

Link for further information

<https://documents.egi.eu/document/1040>; <https://documents.egi.eu/document/960>; <https://documents.egi.eu/document/114>

Wider impact of this work

The final goal for the community is that by the end of EGI-InSPIRE, the implementation of the initial phases of the EGI2020 strategy will have established an open ecosystem where individual actors supported through public and/or private funding have been able to define their own added value and business models to support their activities sustainably. This foundation will enable EGI to continue to play a key role in bringing the digital European Research Area online.

Printable Summary

This session aims at guiding NGIs to analyse the key issues that need to be faced to become sustainable: costs, cost models, service portfolio, business models and service management structures. The evolution of the European Grid from a project-based structure to a network of organisations with a common strategy to provide federated services is a major step forward for long-term sustainability of the EGI community. In this context, each national infrastructure needs to develop or consolidate its own business model and service management structures to achieve sustainability. This dedicated session is designed to better understand the development and progress of the sustainability plans and strategies coming from the national infrastructures, therefore participation from these representatives is essential.

Primary author: ANDREOZZI, Sergio (EGLEU)

Co-authors: MARINOVIC, Damir (EGLEU); APPLETON, Owen (Emergence Tech Limited); HOLSINGER, Sy (EGLEU); SCHAAF, Thomas (Ludwig-Maximilians-Universität München)

Presenter: ANDREOZZI, Sergio (EGLEU)

Track Classification: Community and Co-ordination (Sergio Andreozzi: track leader)

Contribution ID: 27

Type: **Workshop**

Sustainability of Technology Providers

Description of the work

The technology area within the EGI ecosystem is built upon open-source or commercial software coming from community and generic technology providers that is put together by platform integrators to meet the needs of particular research communities. For instance, the EMI project integrates a platform for high-throughput computing from software that is developed within the project primarily for the EGI community (i.e. community technology providers such as the product teams within EMI) with software developed outside the EGI community (i.e. generic technology providers such as Apache) to meet particular use cases coming from their target research community (e.g. WLCG).

The aim of this session is to share vision and strategies from the perspectives of technology providers to guarantee sustainable software and support for the EGI community.

Presentations and discussions will be focused on the sustainability plans of current technology providers offering software and services to the EGI community as well as an opportunity for a third party perspective in where they see the value of being involved with EGI.

Link for further information

<https://documents.egi.eu/document/104>;

<https://documents.egi.eu/document/970>;

<https://documents.egi.eu/document/960>;

Wider impact of this work

The final goal for the community is that by the end of EGI-InSPIRE, the implementation of the initial phases of the EGI2020 strategy will have established an open ecosystem. As EGI aims play a key role in bringing the digital European Research Area online, the technology needed to implement its long-term strategy will rely on external technology providers ensuring the required software and services are delivered and supported. Therefore this session is dedicated to better understanding the development and progress of the sustainability plans and strategies coming from the technology providers and potential opportunities moving into the future.

Printable Summary

The technology area within the EGI ecosystem is built upon open-source or commercial software coming from community and generic technology providers that is put together by platform integrators to meet the needs of particular research communities. For instance, the EMI project integrates a platform for high-throughput computing from software that is developed within the project primarily for the EGI community (i.e. community technology providers such as the product teams within EMI) with software developed outside the EGI community (i.e. generic technology providers such as Apache) to meet particular use cases coming from their target research commu-

nity (e.g. WLCG). The aim of this session is to share vision and strategies from the perspectives of technology providers to guarantee sustainable software and support for the EGI community.

Primary author: ANDREOZZI, Sergio (EGLEU)

Co-authors: MARINOVIC, Damir (EGLEU); HOLSINGER, Sy (EGLEU)

Presenter: ANDREOZZI, Sergio (EGLEU)

Track Classification: Community and Co-ordination (Sergio Andreozzi: track leader)

Contribution ID: 28

Type: **Presentation**

Dynamic federations: storage aggregation using open tools and protocols

Tuesday, September 18, 2012 11:20 AM (20 minutes)

Description of the work

The Dynamic Federations project considers HTTP/WebDAV and NFS4.1-based storage elements and makes them able to cooperate through an architecture that properly feeds the redirection mechanisms that they are based upon, thus giving the functionalities of a “loosely coupled” storage federation.

One of the key features is to use standard clients (provided by OS'es or open source distributions, e.g. Web browsers) to access an already aggregated system; this approach is quite different from aggregating the repositories at the client side through some wrapper API, like for instance GFAL, or by developing new custom clients.

The challenge, here undertaken by the providers of dCache and DPM, but pragmatically open to other Grid and Cloud storage solutions, is to build such a dynamic system while being able to accommodate name translations from existing catalogues (e.g. LFCs), experiment-based metadata catalogues, or stateless algorithmic name translations, also known as “trivial file catalogues”.

Other technical challenges are latency and scalability, and the ability to create worldwide storage federations that are able to redirect clients to repositories that they can efficiently access, for instance trying to choose the endpoints that are closer or applying other criteria.

The project is in an advanced demoable status, and the inclusion in the EMI distribution will be one of the next short term steps.

Wider impact of this work

The features of a loosely coupled, high-performance federation of open-protocols-based storage elements will open many possibilities of evolving the current computing models without disrupting them, and, at the same time, will be able to operate with the existing infrastructures, follow their evolution path and add storage centers that can be acquired as a third-party service.

Printable Summary

A number of storage elements and catalogues now offer standard protocol interfaces like NFS 4.1/pNFS and WebDAV, for access to their data repositories, in line with the standardization effort of the European Middleware Initiative (EMI). Here we report on work that exploits the federation potential of these protocols, making it possible to build a system that offers a unique view of the storage and metadata ensemble and the possibility of integration of other compatible resources such as those from cloud providers.

Such so-called storage federations of standard protocols-based storage elements give a high performance unique view of their content that does not need to index the content of all the storage, thus promoting simplicity in accessing the data they contain and offering new possibilities for resilience and data placement strategies.

Primary authors: Mr DEVRESSE, Adrien (CERN); Mr ALVAREZ AYLLON, Alejandro (CERN); FURANO, Fabrizio (CERN); KEEBLE, Oliver (CERN); FUHRMANN, Patrick (DESY); ROCHA, Ricardo (CERN)

Presenter: FURANO, Fabrizio (CERN)

Session Classification: New middleware products

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 29

Type: **Presentation**

A workflow manager for NMR: objectives, design and implementation

Friday, September 21, 2012 1:30 PM (22 minutes)

Description of the work

Nuclear Magnetic Resonance (NMR) is an important method for the analysis of macromolecular structure and dynamics. There are a host of different computational processes involved in the interpretation of NMR data. These traditionally make use of different programs, which store data in a wide range of formats. Inter-conversion between these formats is not trivial - the underlying information is extremely complex - and this has limited the ability to chain these processes together in an automated way. Our objective is to develop an application that will provide integrated access to the NMR processing programs presented by the WeNMR portals, allowing them to be combined into workflows in a simple and streamlined manner.

The CcpNmr Workflow Management System (WMS) uses the CCPN data model and API for data storage. Seamless conversion to and from program-specific formats is effected by separate Python modules, using the known input data to disambiguate the program output. For the actual calculations WeNMR web portals that present the relevant NMR analysis programs, are exposed as WSDL web services. These are accessed via a common web-based GUI, with individual protocol interfaces specified by templates. The basic architecture is a GWT client linked to a Java/Hibernate server deployed under Tomcat. The workflow management is carried out using Taverna.

Importantly, WMS supports the systematic tracking and management of the data itself. NMR data analysis involves many different processes with subtly different input data combined either sequentially or as alternatives. WMS keeps a record of all the various data versions and helps scientists keep track of their work - something which a survey of potential users at the start of the project identified as a significant unmet need.

As a case study, we present a workflow that takes a single set of input data, sends it to several different WeNMR structure calculation portals, and integrates the results for subsequent analysis.

Wider impact of this work

WMS is expected to; (i) improve accessibility of the WeNMR services, and (ii) provide a framework for the development of novel workflows to address scientific problems. As structural biology becomes more established, there will be ever increasing demand from non-specialist users for simple, automated tools that provide reliable results. The establishment of standardized user interfaces and seamless data transfer should be very helpful in this regard.

In addition, the WMS should be extremely useful for managing the large amounts of data generated in NMR spectroscopy. We expect spectroscopists to be relatively open to the use of LIMS tools like WMS, given that NMR data processing is already heavily computerized. The direct integration with the programs that are already in common use in the field should make WMS compatible with established work practices without the need for additional data capture steps, thus minimizing the demands on the user and ultimately increasing take-up.

Printable Summary

As part of the WeNMR project, we are developing a workflow and data management system targeted at scientists using nuclear magnetic resonance spectroscopy (NMR). This “end-user platform” will support seamless task execution and data flow across the WeNMR portals for NMR-related calculations, as well as our CcpNmr analysis suite. NMR is a key method for investigating the structure of biological macromolecules, involving both manual and computational steps. For the computational steps there is a growing desire to use distributed computing infrastructures, as the popularity of the WeNMR VO testifies. However, combining grid-based computational tools into more complex workflows presents significant technical challenges, particularly in terms of format inter-conversion. The prize at the end is a framework for the development of novel, increasingly automated workflows for NMR data analysis. In addition, it will provide improved laboratory data management and simpler access to WeNMR services.

Primary author: Dr FOGH, Rasmus (University of Cambridge)

Co-authors: Prof. LAUE, Ernest (University of Cambridge); Dr IONIDES, Jonn (University of Cambridge); Dr LUNDBORG, Magnus (University of Cambridge); Dr STEVENS, Timothy (University of Cambridge); Dr BOUCHER, Wayne (University of Cambridge)

Presenter: Dr FOGH, Rasmus (University of Cambridge)

Session Classification: Workflow community workshop

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 30

Type: **Poster**

Using Adaptation Strategies to Improve Grid Operations

Description of the work

An adaptation strategy in ACTRESS is based on a feedback control loop that is represented in a model following an intuitive sense/compute/control decomposition. A target system context (WMS host, L&B service, etc.) is modelled by sensors and effectors representing observable context information (system load, job state, queue size, etc.) together with available actions (service restart, job resubmission, etc.). These touch points are usually defined using interfaces provided by an operating system and middleware services. The actual adaptation strategy is expressed in controllers. Based on the sensors inputs, a user describes what effectors should be triggered in order to realize the envisioned adaptation.

There is a wide range of possible strategies including a job resubmission upon failure with automatic black listing of the computing element where the execution has failed, a detection and a suppression of a malfunctioning resource, purging saturated job queues, or replicating data in several storage elements by choosing the closest or the highest capacity server.

To aid the development of these adaptive strategies, the toolkit provides both a design time and a runtime support. The design part currently contains (1) a language to express both a structure and a behavior of the different elements and is compiled into the adaptation model, (2) a validator for the adaptation model that checks its well-formedness, and (3) a verifier that performs an explicit verification of assumptions about the model expressed in the linear temporal logic using the SPIN model checker. The runtime support consists of a code generator that translates the adaptation model concepts into a Scala source code together with deployment scripts.

Link for further information

<https://salty.unice.fr>

Wider impact of this work

The presented toolkit enables grid administrators and users to develop solutions for some of the recurring problems they face during grid operation that would otherwise require their immediate attention. While the actual development of the adaptation strategies, especially of the more advanced ones is still a challenging problem, the ACTRESS toolkit aims at making the overall process easier by providing the tools and well defined building blocks for rapid prototyping.

The work is being evaluated in the context of the NeuGRID project that operates the gLite infrastructure, however, it is a general framework for assembling an external adaptive behavior on the top of existing systems. Among others, our ongoing long term goal is also to create a library of reusable elements and adaptation patterns that encapsulates the common services in the contemporary distributed computing infrastructures.

This work has been partly funded by the ANR SALTY project (ANR-09-SEGI-012).

Printable Summary

Due to the complexity of grid infrastructures human administration costs of grid operations are high and end-users are not completely shielded from the system heterogeneity leaving them to deal explicitly with the reliability issues.

Instead of trying to achieve complete reliability within the middleware itself, we propose new operation modes where grid administrators describe their goals and knowledge about the system configuration, optimization and failure recovery into executable adaptation strategies. These strategies then run autonomously within the infrastructure freeing administrators from some common repetitive operational issues and allowing them to focus on a higher level supervision of the system.

This poster presents the ACTRESS toolkit that can be used by the grid administrators and users to rapidly prototype such adaptation strategies while abstracting them from the painful low-level implementation details.

Primary author: Mr KŘIKAVA, Filip (I3S - CNRS UMR 7271)

Co-authors: Dr ROJAS BALDERRAMA, Javier (I3S - CNRS UMR 7271); Dr MONTAGNAT, Johan (CNRS); Dr COLLET, Philippe (I3S - CNRS UMR 7271)

Presenters: Mr KŘIKAVA, Filip (I3S - CNRS UMR 7271); Dr ROJAS BALDERRAMA, Javier (I3S - CNRS UMR 7271)

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 31

Type: **Demonstration**

A General Purpose Grid Portal for simplified access to Distributed Computing Infrastructures

Description of the work

Since one of the major obstacles that very often prevent new users from using the grid infrastructure is the AA complexity, the portal is based on an external authentication layer providing strong user identification by means of an accredited identity federation.

The portal is interfaced to an online CA, which provides users with digital certificates in a transparent way; a long term proxy is then created and saved into a MyProxy server in an encrypted form.

Through the portal users can then submit either simple jobs or complex workflows to the grid by simply providing a passphrase to retrieve and decrypt the proxy without the need of directly managing the X509 certificate.

Users belonging to a specific community can have a customized view where they can submit their applications to several different middleware stacks.

The portal also provides a simple but powerful data-management tool that does not require custom software on the client side and allows users to upload data to the Grid, move data among Grid sites and download the output of the grid jobs. Users are notified by mail when the upload/download operations are completed.

The demo will show how users register to the portal, how the portal will get a digital certificate from the online CA on behalf of the user, how users register to a VO, how the proxy is created and stored in the repository, how simply jobs can be submitted to the grid and customized view for SPES experiment community.

Link for further information

<https://wiki.italiangrid.org/twiki/bin/view/IGIPortal/WebHome>

Wider impact of this work

This solution can have a huge impact on the user communities that currently are not using the grid because of the too complex access and it can then greatly improve the number of grid users.

Printable Summary

The Italian National Grid Infrastructure (IGI) has developed a web portal which provides a powerful and easy to use access gateway to distributed computing and storage resources.

To hide the complexity of X509 certificates the portal is interfaced to an online CA that provides certificates to users belonging to an accredited identity federation.

Primary author: BENCIVENNI, Marco (INFN)

Co-authors: CECCANTI, Andrea (INFN); CESINI, Daniele (INFN); Dr MICHELOTTO, Diego (INFN-CNAF); GIORGIO, Emidio (INFN); MISURELLI, Giuseppe (INFN); GAIDO, Luciano (INFN); VERONESI, Paolo (INFN); BRUNETTI, Riccardo (INFN); ZAPPI, Riccardo (INFN); Dr CECCHINI, Roberto (INFN--Firenze); Dr VENTURI, Valerio (INFN-CNAF)

Presenters: Dr MICHELOTTO, Diego (INFN-CNAF); BENCIVENNI, Marco (INFN)

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 32

Type: **Presentation**

IPv6 Testbed for ARC-CE

Tuesday, September 18, 2012 2:20 PM (20 minutes)

Description of the work

The idea is to use IPv6-only in all aspects of testing: installation, configuration and testing of Arc client and server. For doing that a lot of other services have to be IPv6-ready, such as IPv6-enabled repositories for installation, IPv6-ready resource manager, scheduler and VOMS for submission etc. We encounter problems on different levels, beginning with a resource manager (Slurm, Torque and SGE are currently not working with IPv6). After establishing a fully functional ARC-based cluster using v6 protocol, we should also test its security, stability and provide monitoring for it. But this will happen in the subsequent steps.

Wider impact of this work

IPv6 testing is necessary if we want to deploy IPv6 clusters in the future.

Printable Summary

IPv4 exhaustion is becoming a reality, it will soon be impossible to deploy or even enlarge big clusters, clouds or other services which are based on IPv4. Transition to IPv6 is (or it will be) mandatory.

At Arnes we have set up an IPv6 testbed in 2012, used for ARC server and client testing. Services are tested using dual-stack and IPv6-only, though our goal is to prepare a cluster that is completely functional using IPv6-only. Beside the cluster's functionality, it is also necessary to think about its stability and security. In this presentation we will discuss the setup, testing and experiences that we have gained so far.

Primary author: Mrs KRASOVEC, Barbara (Arnes)

Presenter: Mrs KRASOVEC, Barbara (Arnes)

Session Classification: Network Support

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 33

Type: **Presentation**

Running OpenFOAM in parallel on the grid: CFD in the study of cardiovascular disease

Thursday, September 20, 2012 4:20 PM (20 minutes)

Description of the work

The NGI_IT computing infrastructure has been successfully configured [1] for supporting researchers from the “Mario Negri” Institute to securely access grid resources and to use OpenFOAM (v.2.0.1. and openmpi v.1.5.3), the open-source CFD toolbox library [2]. Preliminarily, the Italian infrastructure has been validated with parallel jobs of very big cylindrical meshes (1,824,200 cells) simulating a transient case of 100 time steps. From a technical point of view, a parallel CFD run is performed in three different steps. Initially, the master node prepares the case by using OpenFOAM’s domain decomposition `decomposePar` utility. After this operation, `processorNN` directories for each CPU is created with decomposed mesh, fields, solution controls, model choice and discretisation parameters. Distribution of input data to each slave node is subsequently guaranteed by the grid infrastructure. The CFD case is then solved with `icoFoam` solver on a per-processor basis, where each CPU node uses its local disk space to store results data in time directories. At the end of the run, results are collected from the slave nodes to the master node, and the resulting tar-ball is saved on a grid Storage Element using user-defined post hooks. Finally, the numerical and graphical post-processing is done locally after using the `reconstructPar` utility of OpenFOAM to recreate the case to a single CPU.

In the present work, patient-specific, image-based [3] transient CFD of the carotid bifurcation, with physiological flow wave form as boundary conditions have been processed using the `icoFoam` solver in parallel. CFD results were then post-processed to calculate haemodynamic wall parameters that predict disturbed flow, athero-prone, sites on the vascular wall like the time averaged wall shear stress (TAWSS), the oscillatory shear index (OSI) and the relative residence time (RRT). Disturbed flow is localized by the wall surface area exposed to low and oscillating wall shear stress having $RRT > 1$.

Link for further information

<https://wiki.italiangrid.it/twiki/bin/view/UserSupport/OpenFoam>

Wider impact of this work

NGI_IT infrastructure has been successfully used to efficiently run parallel CFD simulations in large arteries, but we are only at the beginning of our research roadmap towards patient-specific pilot clinical studies. The performance we have obtained in this first phase are also quite encouraging: compared to 1 CPU, for runs on 4 and 8 nodes we have measured a speed-up of large CFD simulations of 3.0 and 3.9, respectively [4]. Future improvements of our numerical simulations will include setting of fluid-structure interaction (FSI) for blood vessels, a new challenge for the cardiovascular biomechanics research.

Moreover, the extensive range of partially derivate equations (PDE) solvers OpenFOAM offers,

from chemical reactions, turbulence, heat transfer, to solid dynamics and electro-magnetism may attract other researchers to use OpenFOAM on the grid.

Printable Summary

Haemodynamic forces play a fundamental role in regulating the vascular structure. If different from the physiological range, namely in “disturbed flow” conditions, these factors are implicated in the aetiology of vascular wall disease. Such evidences have led to the need of in vivo characterization of haemodynamics at patient-specific level. However, the introduction of computational fluid dynamics (CFD) in clinical research is confined by the limited availability of systems able to run transient simulations for large domains and for many subjects. Towards this end, we are setting up the OpenFOAM framework for large CFD cases to be run in parallel on the IGI grid infrastructure. OpenFOAM has efficient algorithms to solve the Navier-Stokes equations for compressible and non-Newtonian fluids as blood. Furthermore, it has adaptive mesh refinement and parallelization features required to perform patient-specific big CFD simulations aimed at characterizing flow patterns in large arteries.

Primary authors: Dr ENE-IORDACHE, Bogdan (Mario Negri Institute for Pharmacological Research); Dr LA ROCCA, Giuseppe (INFN)

Co-authors: Dr CESINI, Daniele (INFN); Dr GIORGIO, Emidio (INFN); Dr GAIDO, Luciano (INFN); Dr RIZZI, Massimo (Mario Negri Institute for Pharmacological Research)

Presenters: Dr ENE-IORDACHE, Bogdan (Mario Negri Institute for Pharmacological Research); Dr LA ROCCA, Giuseppe (INFN)

Session Classification: VREs - Community Contributions

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 34

Type: **Presentation**

ERINA+: Welcome and Overview

Wednesday, September 19, 2012 11:00 AM (10 minutes)

Description of the work

e-Infrastructures receives considerable support from the EC, in order to sustain the evolution of e-Science, by improving technology helping researchers, along with the circulation of their survey results across Europe. The EC financed 91 e-Infrastructures projects since 2009 and, in a context of limited resources, the impacts associated to this financing requires to be carefully assessed to be sure that the concrete goals set by the EU for the e-infrastructure domain are achieved, if and how the financed projects contribute to these goals. The ERINA+ study goes exactly in that direction. Pivotal of the project is the development of a methodology to assess both the impact of e-Infrastructures on the general economy.(and society) and the impact produced by e-Infrastructures projects to the e-Infrastructures, therefore evaluating also the value they add to the general economy and social welfare. The methodology will produce an assessment in qualitative and quantitative terms and means. Its main components include: a mapping of e-Infrastructures and e-Infrastructures projects, the analysis of their socio-economic impacts, an valuation of e-Infrastructures performance also from the point of view of their users and other stakeholders and the analysis of the impact on European Research Area development. A web tool developed by the consortium will enable projects to perform a self-assessment that will help them in understanding their impacts and in improving their sustainability plans.

The presentation will present:

- (1) The ERINA+ Overview
- (2) The ERINA+ Methodology
- (3) The ERINA+ Data Gathering Tools
 - a. Project Self-Assessment Webtool
 - b. Project Users' Questionnaire
 - c. Key Stakeholder Questionnaire
- (4) The preliminary data collection results.

Wider impact of this work

The impact will be the presentation ERINA+ Project Self-Assessment Webtool, the opportunity for the attendees to better understand the methodology, the opportunity to further transfer the ERINA+ methodology and for spreading a culture of socio-economic impact assessment and to have an open discussion.

Printable Summary

The ERINA+ consortium will be presenting the final shape of the assessment methodology, as a comprehensive means to enable a socio-economic impact analysis of e-Infrastructures. Part of the discussion will be dedicated as well to the main operational tools which, above all, helped the Consortium in the gathering of information and, therefore, allowed the projects to enter the second phase of data analysis, coming to a first overall assessment of the e-Infrastructures domain

through:

- The ERINA+ Project Self-Assessment Webtool, released in July 2012;
- Online questionnaires for retrieving information and feedback from e-Infrastructures from two perspectives: (a) e-Infrastructure Key Stakeholders of and (b) Users of the projects' services.

Primary author: MANIERI, Andrea (Engineering)

Co-authors: BELLINI, Francesco (Eurokleis); BENEDIKT, Josef (ZSI); Dr PASSANI, antonella (T6 ecosystems); NARDI, nadia (Engineering Ingegneria Informatica spa)

Presenter: NARDI, nadia (Engineering Ingegneria Informatica spa)

Session Classification: e-Infrastructure Impact Assessment Methodologies

Track Classification: Community and Co-ordination (Sergio Andreozzi: track leader)

Contribution ID: 35

Type: **Presentation**

New features on Accounting Portal Release Electra

Wednesday, September 19, 2012 11:20 AM (20 minutes)

Description of the work

The new Electra (v4.2) release of the Accounting Portal was made available on <http://accounting.egi.eu>. This is a release, with many internal changes, including many internal code changes and reorganization, a move from frames to full CSS design and addressability, mobile support (including QR), better graphs, XML datafeeds on tree views, full user data on site and VO admin views, by country usage views, and several code optimizations and changes.

The document will detail and flesh out each of these new features and their impact on both performance and user interaction.

Link for further information

<http://accounting.egi.eu>

Wider impact of this work

This document details the changes in the new version of the Accounting Portal and the new functionalities and performance offered to users with this new release.

It is of interest of operations staff, specially site admins, vo admins and any organism using the Portal data to gauge and track the grid infrastructure usage, possibly mediated by a large number of variables.

Printable Summary

The new Electra (v4.2) release of the Accounting Portal was made available on <http://accounting.egi.eu>. This is a release, with many internal changes, including many internal code changes and reorganization, a move from frames to full CSS design and addressability, mobile support (including QR), better graphs, XML datafeeds on tree views, full user data on site and VO admin views, by country usage views, and several code optimizations and changes.

This document details these changes and the new functionalities and performance offered to users with this new release.

Primary author: DIAZ ALVAREZ, Ivan (FCTSG)

Co-authors: SIMON, Alvaro (FCTSG); FERNANDEZ, Carlos (FCTSG); FREIRE GARCIA, Esteban (FCTSG)

Presenters: SIMON, Alvaro (FCTSG); DIAZ ALVAREZ, Ivan (FCTSG)

Session Classification: Operations Workshops

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 36

Type: **Presentation**

TopHat on the Grid: an automatic workflow for sequence alignment exploiting EGI/IGI grid infrastructure

Friday, September 21, 2012 1:52 PM (22 minutes)

Description of the work

The contribution presents an execution of a complex workflow based on a reverse engineering of TopHat over the EGI/IGI grid infrastructure. It aligns RNA-Seq reads to human genomes using the ultra high-throughput short read aligner Bowtie. The workflow then analyzes the mapping results to identify splice junctions between exons. The job submission is executed by means of an already developed service called JST (<https://indico.egi.eu/indico/contributionDisplay.py?contribId=127&sessionId=18&confId=207>). It handles the execution requests and deals with the real grid job submission, monitoring and re-submission. In order to provide a reliable workflow the exit status of each step is checked and each calculation could be re-executed in case of failure. Data transfers are executed using a grid Storage Element as temporary buffer. This tool could be used to exploit both standard grid resources and WNoDeS (<http://web.infn.it/wnodes/index.php/wnodes>) enabled cloud resources. A particular attention will be devoted to explain how we address the problem of transferring input and output data that usually exceed the 3GB size for each job. This implementation provides an improvement of algorithm of TopHat making it parallelizable. Three main blocks have been identified in TopHat, each of which is composed by several segments that can be executed independently on the Grid. Each segment is analyzed by Bowtie. Its aim is to map each short read segment onto the human genome reference. The main advantage of using the cloud solution based on WNoDeS is the possibility of deploying all the computational steps of the workflow in the cloud environment and not only the most CPU intensive ones as in the grid environment. Indeed for few steps of the workflow we need a dedicated environment that is hard to replicate over a standard grid infrastructure.

Link for further information

www.computer.org/portal/web/csdl/doi/10.1109/ISMS.2012.76
bioinformatics.oxfordjournals.org/content/25/9/1105.full.pdf+html

Wider impact of this work

The described activities are the first prototype implementation of a mixed workflow that makes use of local dedicated machines, grid worker nodes and cloud resources and provide a good example of flexible and dynamic resource allocation. This work is useful to all the researchers that could not easily deploy their analysis on the EGI/IGI grid infrastructure only. Our framework allows to dispatch over the grid infrastructures only steps that are really CPU consuming with a minimal impact and modification on the already working workflow. Biotechnological laboratories are producing more and more sequencing data that open new horizons and new challenges to both

computing infrastructures and software engineering. Our approach allows to reduce the elaboration time in the selected blocks; for example it is easy to obtain a big speedup factor in specific steps of the workflow.

Printable Summary

We will present the activity related to the use of TopHat, a fast splice junction mapper for RNA-Seq reads over a grid distributed infrastructure. TopHat manages data for Next Generation Sequencing technology that allows a more accurate analysis as: detection of new isoforms, differential gene expression analysis and detection of aberrant mutations. This technology allows to obtain from the molecules of DNA/RNA smaller fragments, called read, which can be sequenced in parallel. The workflow is executed partially over a dedicated infrastructure hosted in the Istituto Superiore Mario Boella, while the steps that are very CPU consuming are executed dynamically on the grid. The processing framework is able to submit jobs to the grid infrastructure by means of Job Submission Tool. The jobs could be executed both on a standard grid infrastructure or on a cloud infrastructure based on WNoDeS like the IGI one.

Primary author: Dr DONVITO, Giacinto (INFN)

Co-authors: ACQUAVIVA, Andrea (Politecnico di Torino); CESINI, Daniele (INFN); Dr MOSSUCCA, Lorenzo (Istituto Superiore Mario Boella); GAIDO, Luciano (INFN); TERZO, Olivier (Istituto Superiore Mario Boella); RUIU, Pietro (Istituto Superiore Mario Boella)

Presenter: Dr DONVITO, Giacinto (INFN)

Session Classification: Workflow community workshop

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 37

Type: **Presentation**

Testing SLURM batch system for a grid farm: functionalities, scalability, performance and how it works with Cream-CE

Thursday, September 20, 2012 11:00 AM (30 minutes)

Description of the work

We will show all the work done in order to install and configure the batch system itself together with the security configuration needed.

In this presentation we will show the results of the deep testing that we have done on SLURM, in order to be sure that it will cover all the needed functionalities like: priorities, fairshare, limits, QoS, failover capabilities and others.

We will report also on the possibility of exploiting this batch system within a complex mixed farm environment where grid job, local job and interactive activities are managed exploiting the same batch system.

From a point of view of the scalability we will show how the SLURM batch system is able to deal with the increasing number of node, CPU and jobs served.

We will also show the performance achieved with several client accessing the same batch server. We also will make some comparison with other available open source batch system both in terms of performance and functionalities.

We will also provide feedback on mixed configuration with SLURM and MAUI as job scheduling. We will also describe the work done in order to support SLURM on Cream-CE. Indeed SLURM users community has also expressed a lot of interest in having SLURM supported by the CREAM CE. The integration effort already started at the BLAHP (Batch Local ASCII Helper Protocol) layer, and a basic, prototype-level support for standard jobs is ready. Still, due to the wide range of customisation and different deployment models allowed by SLURM, the task of offering an homogeneous interface for all the deployment scenarios will not be a simple remapping of what has been done for other batch systems.

Wider impact of this work

This activity will give to the computing farm administrator some information required to choose a new open source batch system in order to provide better scalability, without the need of buying costly proprietary batch system. This work will provide information that could help to understand the capabilities and performance of SLURM batch system and provide few feedback on the possibility to use it in a large and complex computing farm infrastructure.

Moreover in this work it will be evident the work needed to support a new batch system in the Cream-CE.

Finally, this work will give the opportunity to use the SLURM as batch system in a computing farm belonging to the EGI/IGI grid infrastructure.

Printable Summary

As the grid computing farm are increasing in size in terms of nodes but even more in terms of CPU slots available, it become of great interest to have a scheduler solution that could scale up to tens of thousands of CPU slots and hundreds of nodes.

In order to try to keep the Total Cost of Ownership as low as possible it will be preferred to have an easy to use and open source solution.

SLURM is able to fulfil all those requirements and it looks promising also in terms of community that is supporting it, as it is used in several of the TOP500 supercomputing.

For this reason we deeply tested the SLURM batch system in order to prove if it could be a suitable solution.

In the work we will present the result of all the test executed on SLURM batch system and the results of the development activity carried on in order to provide the possibility to use SLURM as back-end of a Cream-CE.

Primary author: Dr DONVITO, Giacinto (INFN)

Co-authors: DORIGO, Alvisè (INFN); REBATTO, David (INFN); FRIZZIERO, Eric (INFN); PRELZ, Francesco (INFN); ZANGRANDO, Luigi (INFN); MEZZADRI, Massimo (INFN); SGARAVATTO, Massimo (INFN); ANDREETTO, Paolo (INFN); BERTOCCO, Sara (INFN)

Presenter: Dr DONVITO, Giacinto (INFN)

Session Classification: Resource Infrastructure Services

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 38

Type: **Presentation**

Use Case: Running Monte Carlo LHCb simulations using DIRAC with EGI Federated Cloud

Thursday, September 20, 2012 11:00 AM (20 minutes)

Description of the work

DIRAC is integrated with private, public and hybrid clouds. In combination with CernVM, it provides a very flexible platform for general purpose usage and for this reason it has been chosen for this use case. Minor developments have been required in order to allow multiple cloud providers to be managed from a single DIRAC instance.

The resulted extension to the DIRAC Virtual Engine needs to be tested before being used in production. After performing some small-scale functionality tests, the goal is to execute a large-scale test aiming to consume 2000 core hours at each site on the EGI Federated Cloud infrastructure. Small Virtual Machines (VM) of 1 Core and 2 GB of memory are required. The deployed images are the CernVM with about 14GB of filesystem size on local disk. The scaling test will be run in multiple clouds at the same time.

After the integration and scaling testing, the cloud resources will be connected to the LHCb DIRAC production instance and configure the VMs to match and execute MC simulation jobs. Site Managers will establish a maximum number of virtual machine images running at the same time. The expected contribution from each site, integrated over the duration of this last phase, should be of the order of several thousands of core days. The LHCb Production Managers will take care of the production from LHCb-DIRAC portal and are not supposed to see any difference beyond the additional computing resources available.

The possibility to connect commercial public CPU providers with a contribution at the same level will also be explored.

Link for further information

<http://diracgrid.org>

Wider impact of this work

The use case will prove the viability of a Monte Carlo production using Grid and Cloud resources through DIRAC. This is one of the DIRAC team efforts to provide wider integration of resources, users and tools making DIRAC “the interware” solution. At the same time this use case is contributing to the EGI Federation Cloud aim of a production level.

The expertise of this use case will be useful for other LHCb activities. Furthermore, the experience can be directly transferred to other communities already using DIRAC interware solution (ILC, Belle II, CTA, BES III), and to the NGI infrastructures already providing DIRAC interware as Service for their national user communities (France-Grilles and IberGrid).

Printable Summary

The DIRAC project provides a framework for building ready to use distributed computing systems.

It has been proven to be a useful tool for large international scientific collaborations integrating in a single system, their computing activities and distributed computing resources: Grids, Clouds and HTC clusters. In the case of Cloud resources, DIRAC is currently integrated with Amazon EC2, OCCI/OpenNebula and CloudStack. Some Monte Carlo (MC) simulation campaigns were realized at the large scale project Belle II, providing over 10.000 thousand CPU days from Amazon. Until now, all cases have made use of a single cloud at a time.

The work aims to integrate the resources provided by the multiple private clouds of the EGI Federated Cloud and additional WLCG resources, providing high-level scientific services on top of them by using the DIRAC framework. New design and development are discussed. Initial integration and scaling tests are necessary to anticipate possible problems.

Primary author: Dr MENDEZ, Victor (Port d'Informació Científica, PIC)

Co-authors: Mr ADRIÀ, Casajús (Universidad de Barcelona); Dr MERINO, Gonzalo (Port d'Informació Científica, PIC); Dr DIAZ, Ricardo (Universidad de Barcelona)

Presenter: Dr MENDEZ, Victor (Port d'Informació Científica, PIC)

Session Classification: Virtualised Resources

Track Classification: Virtualised Resources: challenges and opportunities (Michel Drescher: track leader)

Contribution ID: 39

Type: **Presentation**

Latest improvements in the GridWay metascheduler

Thursday, September 20, 2012 5:00 PM (20 minutes)

Description of the work

The Initiative for Globus in Europe (IGE) is a project funded by FP7 to coordinate European Globus activities. This work shows the most significant developments in the GridWay metascheduler supported by IGE.

The most important developments have been the inclusion of two new execution drivers for CREAM and BES execution services. The new execution drivers for GridWay provide an abstraction layer that enables users to submit to resources managed by CREAM or BES. CREAM accepts job submission requests, described using JDL, as well as other job control and monitoring requests. On the other hand, BES defines activities using the JSDL specification. Therefore, GridWay translates its job templates and operations to the above job description languages and the operations provided by the services, respectively. GridWay also has a new remote BES-compliant interface, based on GridSAM, providing access to GridWay's metascheduling and interoperation capabilities through a standard interface.

A new installation procedure is now available, in which the GridWay core and drivers are configured, built and installed independently. This feature avoids recompile GridWay in order to add new drivers. It is also possible to build and install GridWay according to the Filesystem Hierarchy Standard (FHS), while a new configuration option to write logs according to the syslog format has been included to this end.

Other developments in the GridWay core implies the adaptation of the scheduler to submit jobs to resources based only on their rank, and not on their free slots, and the choice of setting resource requirements in job templates, like the maximum time or the memory needed for a single execution of the executable. Finally, other new features in GridWay include a randomized job state polling, and the possibility of disposing jobs when they have finished the execution.

Link for further information

<http://gridway.org/>

<http://www.ige-project.eu/>

Wider impact of this work

The new features of GridWay provide a higher level of interoperation with grid infrastructures, responding to the need to leverage existing infrastructures based on different grid technologies. The CREAM and BES drivers provide end-users the ability to submit, control and monitor the execution of jobs to sites managed by these services.

The improvements in the GridWay core include features demanded by the user community, such as the possibility of writing logs according using syslog that allows sysadmins to store the logs in a central point. Other features provide a more flexible installation procedure, the submission of jobs to resources based only on their rank, or the choice of setting resource requirements to be

forwarded to LRMS. These improvements contribute to provide a more powerful metascheduling tool.

Printable Summary

The GridWay metascheduler performs unattended, reliable, and efficient execution of jobs on heterogeneous and dynamic grids over different grid middlewares. The development of GridWay is being currently supported by the IGE project.

This work presents the latest developments of GridWay, which include new execution drivers that enable the interoperation with other grid infrastructures although they are not based on Globus, the possibility of writing logs with syslog, the forwarding of requirements to LRMS, or a new installation procedure.

Primary authors: HUEDO, Eduardo (UCM); M. LLORENTE, Ignacio (Universidad Complutense de Madrid); Dr MARÍN CARRIÓN, Ismael (Complutense University of Madrid)

Presenter: Dr MARÍN CARRIÓN, Ismael (Complutense University of Madrid)

Session Classification: VREs - Community Contributions

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 40

Type: **Presentation**

A Grid execution model for the ANSYS engineering simulation software

Thursday, September 20, 2012 2:40 PM (20 minutes)

Printable Summary

The increasing availability of computer power on Grid platforms has prompted the implementation of complex computational codes on distributed systems and, at the same time, the development of appropriate visual interfaces and tools able to minimize the skills requested to the final user to carry out massive Grid calculations. The work performed to implement a complex engineering simulation software usecase on distributed systems making use of the IGI web portal is here presented and discussed. The work has been carried out within a collaboration with the User Support Unit of the Italian Grid Initiative (IGI) and the INFN-Legnaro National Laboratories (INFN-LNL).

Wider impact of this work

The porting of legacy applications onto the Grid infrastructure, together with the development of the related workflows and gateway, is being carried out as part of a more general effort to build a solid platform, offered to users as a service, for assembling accurate multi scale realistic simulations. This is the case of the INFN-LNL which activity, in the context of the SPES project R&D [M. Manzolaro, G. Meneghetti, A. Andrichetto, Nucl.Instrum. Methods Phys. Res., Sect. A 623 (2010) 1061-1069], is strongly focused on the electro-thermal design of high temperature devices for the production of Radioactive Ion Beams. The related numerical analyses are strongly non-linear, mainly because of the radiative heat transfer computation, and require a relevant computational power to obtain a solution. The implemented case study demonstrates the validity of this approach and makes available a reusable example for other groups interested in porting their applications to production Grid systems.

Description of the work

The application chosen requested to be ported to the Grid environment, making use of the IGI resources available under the GRIDIT VO, is the commercial suite called ANSYS Inc. ANSYS is an engineering simulation software (computer-aided engineering, or CAE) that offers a comprehensive range of engineering simulation solution sets providing access to virtually any field of engineering simulation that a design process requires. In the present work the suite ANSYS has been installed and configured in some Worker Nodes that belongs to the IGI domain. A web interface to run the simulation exploiting the production Grid services was provided through a dedicated portlet of the IGI Portal which is a powerful and easy to use gateway to distributed computing and storage resources. The web GUI was built in strong partnership with the user community interested in porting the application to gather their requirements. In a typical usecase the user provides the initial input files and configuration parameters and waits for the results until the calculation is terminated. This process may take several hours, even days using the resources available to run this application. As the granted amount of cpu time is limited on Grid sites, special care

was taken in handling the checkpointing of the calculation. Another crucial aspect of such long simulations is the evolution's audit of the calculations at runtime, so a facility that exploits Grid Storage Elements was created to make temporary files and partial output available for inspection at runtime. The present strategy enable the user to check the consistency of the output at runtime, evaluating possible strategies aimed at saving time, computing resources and at avoiding waste of license usage.

Primary author: COSTANTINI, Alessandro (INFN)

Co-authors: ANDRIGHETTO, Alberto (INFN-LNL); MONETTI, Alberto (INFN-LNL); CESINI, Daniele (INFN); MICHELOTTO, Diego (INFN/IGI); GIORGIO, Emidio (INFN); GAIDO, Luciano (INFN); BENCIVENNI, Marco (INFN); MANZOLARO, Mattia (INFN-LNL)

Presenter: COSTANTINI, Alessandro (INFN)

Session Classification: Science Gateways

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 41

Type: **Presentation**

G-NEMO: a Grid empowered version of the state-of-the-art European modelling framework for oceanographic research

Tuesday, September 18, 2012 11:00 AM (22 minutes)

Description of the work

The application chosen to be implemented in the Grid environment, making use of the IGI resources available under the GRIDIT VO, is the simulation framework called NEMO, one of the most important simulation frameworks chosen to be ported in the Grid environment to demonstrate the capabilities of the new IGI platform for parallel applications. NEMO is an ocean modelling framework which is composed of “engines” nested in an “environment”. The “engines” provide numerical solutions of ocean, sea-ice, tracers and biochemistry equations and their related physics. The “environment” consists of the pre- and post-processing tools, the interface to the other components of the Earth System, the user interface, the computer dependent functions and the documentation of the system. In the present work, the NEMO configuration used in MFS was ported to the gLite based infrastructure of the IGI domain. As an added value, a web interface to run the simulation was provided through a dedicated portlet integrated in the IGI Portal which is a powerful and easy to use gateway to distributed computing and storage resources. The web GUI was built in strong partnership with the user community interested in porting the application to gather their requirements. A typical simulation could last hours and even days of concurrent calculation using the parallel resources available in the Italian Grid Infrastructure, so special care was taken in handling the safe interruption and restart of the simulation and a facility to inspect files at runtime was also created.

Wider impact of this work

The implemented case study demonstrates the validity of this approach as part of a more general effort to build, as a service, a solid platform for enabling the access to computing and storage resources for those communities interested in porting their applications to the Grid. The porting of legacy applications onto the Grid infrastructure, together with the development of the related workflows and gateway, is being carried out as part of a more general effort to build a solid platform for assembling accurate multi scale realistic simulations. This is the case of the Research and Development unit at INGV which is focused in development new physical processes in the ocean general circulation model NEMO, in order to produce more accurate forecast and analysis of the Mediterranean Sea, and aims to increase the chance to have new modelled processes in the operational data production service which runs daily at INGV. Otherwise these developments may remain an academic practice.

Printable Summary

The availability of computer power on Grid platforms has prompted the implementation of complex codes on distributed systems and, consequently, the development of appropriate visual inter-

faces and tools able to minimize the skills requested to the final user to carry out massive Grid calculations. In this work the Nucleus for European Modelling of the Ocean (NEMO) package has been implemented on the Italian Grid Initiative (IGI) infrastructure making use of the IGI web portal. NEMO is developed by a European Consortium established between CNRS (F), Mercator-Ocean (F), NERC (UK), UKMO (UK) and since 2011 CMCC (I) and INGV (I). The work has been carried out within a collaboration with the User Support Unit of IGI and the National Group for Operational Oceanography (GNOO) at National Institute of Geophysics and Vulcanology (INGV) and benefits of the activities of a more general project developed inside IGI aimed at deploying a scalable, reliable and easy-to-use HPC distributed platform.

Primary author: COSTANTINI, Alessandro (INFN)

Co-authors: CESINI, Daniele (INFN); MICHELOTTO, Diego (INFN-IGI); GIORGIO, Emidio (INFN); GAIDO, Luciano (INFN); BENCIVENNI, Marco (INFN); ADANI, Mario (INGV); DRUDI, Massimiliano (INGV); ODDO, Paolo (INGV); ALFIERI, Roberto (INFN); BOCCIA, Vania (INFN)

Presenter: COSTANTINI, Alessandro (INFN)

Session Classification: Research Infrastructures

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 42

Type: **Presentation**

EGI delves into the sea depths: the collaboration with EMSO begins

Tuesday, September 18, 2012 2:44 PM (22 minutes)

Description of the work

EMSO is going to exploit the power of EGI to create a data infrastructure to serve the wide communities of scientists studying marine mammals' acoustics, Oceanography, Geophysics, Astroparticle Physics, and Ecology. The capabilities offered by the EGI computing infrastructure will be explored through a pilot activity that has been recently agreed with the Italian Grid Initiative as part of the NGI_IT User Support strategy. This activity aims at deploying a Grid-based solution for the data repository and for the basic tools needed to manage data acquired at the Western Ionian Sea node near Catania. The computing model and data distribution infrastructure designed for EMSO will be based on layered structure similar to the one already adopted for the LHC community. Tier0s will be created as close as possible to the experimental sites and will host raw data, Tier1s will host replicas of the data needed for the analysis and will run analysis jobs. In the future the need for a Tier2 layer will be evaluated. Furthermore, when the infrastructure will be set up at the European scale, it will be considered the creation of an higher level service brokering and optimizing the usage of resources. . For the pilot activity, the scientific data will be generated by two offshore experimental sites, one is located offshore Catania, the other offshore in Porto Palo, about 45Km southwest of Syracuse. From these sites, data will reach harbors through fiber cables. The biggest part of data will come from hydrophones at the rate of 6Mb/s. Data collected from the seafloor observatories are firstly saved on a server located in the harbors and then transferred to LNS (INFN South Laboratories) in Catania where a grid Storage Element will be configured to export the RAW data to the whole EMSO collaboration. The pilot activity includes also porting to Grid of some analysis software applications.

Link for further information

<http://www.emso-eu.org/>

Wider impact of this work

The successful gridification of the real time data and software applications, the development of a web-based interface will ease the data management and the usage of the ported applications. The outcome of the pilot will also be used to involve other partners and other NGIs in supporting EMSO Grid activities. If other NGIs are interested, the pilot activity could evolve towards an EGI Virtual Team and create a computing infrastructure serving scientists and other stakeholders in Europe and outside Europe for long-term deep water observation and investigation. Moreover, it will promote the catalytic process and synergic effort at EC and national levels, coordinating and harmonizing all available resources.

Printable Summary

Understanding processes in the marine environment are the key-enable to address complex present-day challenges, such as potential impacts of climate change, prevention of marine ecosystems and mitigation of natural hazards. EMSO, the European Multidisciplinary Seafloor Observatory, is a distributed network of platforms and several deep-seafloor observatories deployed on specific sites around European waters, reaching from the Arctic to the Black Sea passing through the Mediterranean Sea, with the scientific objective of long-term monitoring of environmental processes related to the interaction between the geosphere, biosphere, and hydrosphere, including natural hazards. The reliable and scientific-based information available through these seafloor observatories will help decision makers to study multiple, interrelated processes over time scales ranging from seconds to decades and enable the long time-series collection of multiple variables at fixed locations.

Primary author: Dr LA ROCCA, Giuseppe (INFN)

Co-authors: CESINI, Daniele (INFN); Dr EMBRIACO, Davide (INGV); GIORGIO, Emidio (INFN); Dr ZANOLIN, Francesco (INGV); Dr GIOVANETTI, Gabriele (INGV); Dr RICCOBENE, Giorgio (INFN); Dr MARINARO, Giuditta (INGV); GAIDO, Luciano (INFN); Dr BADIALI, Lucio (INGV); Dr FAVALI, Paolo (INGV)

Presenter: Dr LA ROCCA, Giuseppe (INFN)

Session Classification: Research Infrastructures

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 43

Type: **Workshop**

Workshop: Future advancements of tools and regionalization

Wednesday, September 19, 2012 2:00 PM (10 minutes)

Description of the work

EGI-InSPIRE, within the work package 7, JRA1, develops and maintains the operational tools needed to properly run the EGI infrastructure and to support users and operators in their grid activities. The list of developed tools includes: the Grid Configuration Database (GOCDB), the EGI Helpdesk (GGUS), the Service Availability Monitor (SAM), the Accounting Repository, the Accounting Portal, the Metrics Portal, the Operations Portal.

In this workshop the tools product teams will present the status and development plans for most of the tools focusing on those that are intended to be regionalized and deployed at the NGI level. The product teams presentations will be the basis for an open discussion with all the involved actors that can bring feedback and requirements. The status of tools regionalization after two years of developments will be analysed and a coherent view of the regionalization activities until the end of the project should emerge during the workshop.

Wider impact of this work

Several actor access daily the operational tools developed by EGI, end users, operators on duty and shifters, site managers, grid managers, VO managers. All of them could be interested in this workshop and are encouraged to attend and provide feedback.

Printable Summary

Operational Tools are fundamental in the day by day operation of the GRID. They are accessed by several actors of the grid ecosystem, such as: end users, operators on duty and shifters, site managers, grid managers, VO managers.

EGI develops most of the needed operational tools: the Grid Configuration Database (GOCDB), the EGI Helpdesk (GGUS), the Service Availability Monitor (SAM), the Accounting Repository, the Accounting Portal, the Metrics Portal, the Operations Portal.

In this workshop we will analyse and discuss the foreseen roadmap for the tools, in particular the evolution of those that are (or will be) regionalized to be deployed at the NGI level.

Primary author: CESINI, Daniele (INFN)

Co-authors: Ms PACKER, Alison (STFC); SIMON, Alvaro (FCTSG); TRIANTAFYLLIDIS, Christos (GRNET); LORPHELIN, Cyril (CNRS); Mr IMAMAGIC, Emir (SRCE); GREIN, Guenter (KIT - Karlsruhe Institute of Technology); DIAZ ALVAREZ, Ivan (FCTSG); CASSON, John; Dr GORDON, John

(STFC); BABIK, Marian (CERN); SOLAGNA, Peter (EGLEU); Dr FERRARI, Tiziana (EGLEU); Dr ANTONI, Torsten (KIT-G); Mr ROGERS, William (STFC); MEREDITH, david (STFC)

Presenter: CESINI, Daniele (INFN)

Session Classification: Operations Workshops

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 44

Type: **Internal Project Meeting**

OTAG Meeting (Closed)

Friday, September 21, 2012 1:30 PM (1h 30m)

Description of the work

OATG-13 will analyse and prioritise the main open requirements for the operational tools developed by EGI. The outcome of the Workshop on future advancements of tools and regionalization will also be presented and discussed.

Wider impact of this work

Closed meeting - only for the members of the EGI OTAG group.

Printable Summary

EGI OTAG-13 - 13th Operational Tools Advisory Group Meeting (closed)

Primary author: CESINI, Daniele (INFN)

Presenter: CESINI, Daniele (INFN)

Session Classification: OTAG Meeting

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 45

Type: **Presentation**

From Bare-Metal to Cloud

Tuesday, September 18, 2012 11:30 AM (30 minutes)

Description of the work

ZHAW and GWDG had a shared, common problem, namely how to deploy infrastructural service technology (e.g. OpenStack, CloudStack) with the least amount of user interaction (i.e. automated) during the deployment process, across a large number of servers. The solution allows for the easy deployment of operating systems on to bare-metal (physical servers) and the deployment and management of specified software packages upon those provisioned bare-metal systems. To accomplish the combination of Foreman and puppet was chosen. For the work, it was assumed that the network architecture, partitioning etc. is already determined.

Wider impact of this work

Many organisations face the challenges of deploying cloud technology. This presentation and demonstration focuses on the latter part with a particular concentration on infrastructural service offerings. The approach is not only applicable to physical but also virtual infrastructure and it can be also used to automate the provisioning of PaaS-style services (e.g. Hadoop or CloudFoundry). The challenge of deploying cloud infrastructure technologies is one that has many solutions but in totality are fragmented and require sourcing from different places. This work collates these different sources. The knowledge gained by participants from this presentation and demonstration will bring a full end-to-end picture and means to executing automated deployments to each. On parting from the session, participants will have a good understanding of an end-to-end automation stack and will have the knowledge to implement this either as simulated infrastructure or on actual physical infrastructure.

Printable Summary

This presentation will detail what measures have been taken to automate the provisioning of OpenStack clusters at two research labs, ICCLab and GWDG. The presentation will describe the technology stack, discuss the individual technologies used and share the information with others. It will conclude with a demonstration of provisioning a multi-cluster OpenStack deployment upon virgin bare metal servers.

Primary authors: EDMONDS, Andy (Zurich University of Applied Sciences (ZHAW)); KASPRZAK, Piotr (Gesellschaft für wissenschaftliche Datenverarbeitung mbH (GWDG))

Presenters: EDMONDS, Andy (Zurich University of Applied Sciences (ZHAW)); KASPRZAK, Piotr (Gesellschaft für wissenschaftliche Datenverarbeitung mbH (GWDG))

Session Classification: Providing cloud services

Track Classification: Virtualised Resources: challenges and opportunities (Michel Drescher: track leader)

Contribution ID: 46

Type: **Presentation**

The HEP transition to IPv6

Tuesday, September 18, 2012 2:40 PM (20 minutes)

Description of the work

The requirements for a successful transition to IPv6 on the WLCG production infrastructure are simple to express: namely that the functionality, performance and security of all services must be at least as good as they are today under IPv4.

The HEPiX IPv6 Working Group has been investigating the many issues which feed into the decision on the timetable for a transition to the use of IPv6 in HEP Computing. The activities include the analysis and testing of the readiness for IPv6 and performance of the many different components essential for HEP computing, taking a full systems view of infrastructure and applications. We have been working closely with the HEP experiment collaborations and related IPv6 activities in EGI. The working group is also considering other operational issues such as the implications for security arising from a move to IPv6.

A survey of all important applications and operational tools is well underway. This is restricted to an analysis of the IPv6 readiness of the important WLCG outward-facing services and essential applications and operational tools. In addition to this survey we have already deployed a distributed IPv6 testbed connected to the national research networks that have been providing production IPv6 routing for a few years. We have been testing various Grid services to determine their behaviour in a dual stack environment. During the rest of 2012 we will need to perform larger-scale tests of the functionality, performance, and security of many different components of WLCG.

Wider impact of this work

The decision as to when to support dual stack, IPv6 and IPv4, services needs to take the needs of many stakeholders into account. This will have a major impact on operations, including networking support and the security team, and careful testing and planning is required. For the WLCG community we will also need to involve other Infrastructures such as Open Science Grid in the USA.

Printable Summary

Moving distributed applications to IPv6 involves much more than the routing, naming and addressing solutions provided by the campuses and national networks. Application communities need to perform a full analysis of their applications, middleware and operational tools to determine whether they are IPv6-compliant and establish how much development work is required to update those which are non-compliant. The HEPiX IPv6 Working Group has been investigating the many issues which feed into the decision on the timetable for the transition to IPv6 and has deployed a distributed IPv6 testbed. This paper presents the work being done by the group in collaboration with WLCG and IPv6 activities in EGI.

Primary author: KELSEY, David (STFC)

Presenter: KELSEY, David (STFC)

Session Classification: Network Support

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 47

Type: **Presentation**

Grid Network Accounting for Incident Response

Description of the work

It would be useful to better log the network flows and connections through our firewall for, amongst other reasons, to be able to respond more confidently to CSIRT advisories, that often ask to check for connections to/from a given IP address.

Grid service administrators may not have administrative access to switch or router equipment in their hosting environment that can record such information (using netflow or sflow).

A solution should be installable by grid site administrators on the hosts / vms / firewalls they control.

For logging / probe solutions that run on firewalls it is important that the monitoring does not significantly degrade the throughput or latency of the system.

A key requirement is that logs should be relatively compact. For instance, full iptables plain-text logging of all packets would require significant storage.

In general, grid sites may not have funding to purchase commercial network traffic logging and analysis hardware or software, and as mentioned above the grid site administrators may not have administrative access. If it is available at a site it can of course be used, but we assume that a site wishes to implement a low-cost software-based solution.

Probes, repeaters, and tunnels should be configured securely to avoid tampering.

There are a variety of options

- * iptables rules with rate limiting
- * traffic analysis software such as ntop
- * software based probes such as fprobe, nprobe, softflow, argus.

The approach is to install network probes on firewalls or vm hosts controlled by grid site administrators and to feed the output to collectors which store the flow data. Various tools can be used to visualise and search the stored data.

For NGI-wide deployment, probes need to send flow data through secure tunnels to a central collector.

Issues include simplifying deployment of probes and collectors and evaluation of the performance impact.

Wider impact of this work

The aim of this work is to describe a usable method of network logging to better equip grid sites and NGIs of all sizes (and sites in similar distributed infrastructures) to respond to network security incidents.

Printable Summary

Grid infrastructure providers and research communities run distributed Computer Security Incident Response Team functions, which coordinate incident response on a global scale. Unfortunately, international research collaborations can provide a path for security incidents to propagate easily.

To make use of network information from investigations grid sites must record network activity in sufficient detail. Firewall logs provide partial records of network activity. Network administrators can make use of logging facilities provided by routers. However, grid sites may not have direct access to such facilities and so need a solution that can record traffic efficiently in terms of network performance, storage space, and search.

We describe the use cases for network logging in the grid CSIRT context; we explore several approaches and related work in this area; we describe a deployment of our chosen solution at NGI scale; we evaluate the solution in practice; and provide recommendations.

Primary author: O'CALLAGHAN, David (TCD)

Co-author: KENNY, Stuart (Trinity College Dublin)

Presenter: O'CALLAGHAN, David (TCD)

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 48

Type: **Poster**

NGI-DE General Operations Policy

Description of the work

Grid Computing in Germany started in 1997. Projects like D-Grid, EGEE and EGI-InSPIRE enabled the development of a heterogeneous and productive e-infrastructure. To ensure a sustainable and seamless operation of the existing e-Infrastructure a common understanding of policies and procedures is needed.

The NGI-DE General Operations Policy (GOP) will be based on the EGI procedures and policies, the "D-Grid-Betriebskonzept"(German Grid Operations Concept) and German law.

The idea of the GOP is to provide a framework, where European and national procedure could be integrated.

The GOP will be a high level document without technical details. It will be a permanent agreement signed by all partners. The GOP will be written by NGI-DE and include Scope, Definitions and the General Policies.

Details will be put in annexes and appendixes. Examples for that are documents about "NGI-DE user requirements", "NGI-DE security and privacy", "NGI-DE infrastructure services"(including e.g. "NGI-DE help desk"and "NGI-DE accounting") and "NGI-DE resource centre requirements" (including e.g. "NGI-DE resource centre certification for gLite", "NGI-DE resource centre certification for UNICORE", "NGI-DE resource centre certification for Globus", "NGI-DE resource centre certification for dCache").

The working language of all technical documentation is English to ensure an easy exchange of documents with other NGIs.

Link for further information

<http://www.ngi-de.eu/> - no public information at the moment

Wider impact of this work

NGI-DE presents the actual work in progress to discuss the ideas with other NGI operation groups.

This framework and its annexes and appendixes could be of interest for any NGI as a blue print.

The GOP is a building block for a sustainable e-infrastructure.

Printable Summary

To ensure a sustainable and seamless operation of the existing e-infrastructure a common understanding of policies and procedures on a national level is needed.

The NGI-DE General Operations Policy (GOP) will be based on the EGI procedures and policies, the "D-Grid-Betriebskonzept"(German Grid Operations concept) and German law.

The idea of the GOP is to provide a framework, where European and national procedures could be integrated. NGI-DE presents the actual work in progress to discuss the ideas with other NGI

operation groups.

Primary authors: STREIT, Achim (KIT-G); NILSEN, Dmitry (KIT-G); WEBER, Pavel; Dr ANTONI, Torsten (KIT-G); BUEHLER, Wilhelm (Karlsruhe Institute of Technology)

Presenters: NILSEN, Dmitry (KIT-G); BUEHLER, Wilhelm (Karlsruhe Institute of Technology)

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 49

Type: **Demonstration**

GGUS report generator

Description of the work

With the start of EGI Inspire various requirements on reporting came up which had not been covered by the previous version of the GGUS report generator. After collecting all requirements on a workshop held at KIT in October 2011 the GGUS team started implementing a new report generator.

The first version was brought online in time for the EGI Inspire review at the end of June 2012.

Since then further enhancements have been implemented adding all the requested functionality.

The metrics calculations are documented in the egi wiki https://wiki.egi.eu/wiki/FAQ_Report_Generator_%28GGUS%29

Wider impact of this work

The GGUS report generator allows operations staff, support teams and user communities to calculate metrics reports at any time according to their needs without relying on the assistance of the GGUS team.

Printable Summary

The GGUS report generator is offering individual and pre-defined metrics reports on the performance of GGUS and the integrated support units. The relevant key parameters for the reports can be modified for adapting the reports to the user's needs.

Primary authors: Mr ROGGE, Carsten (KIT); GREIN, Guenter (KIT - Karlsruhe Institute of Technology)

Co-authors: DRES, Helmut (KIT-G); Dr ANTONI, Torsten (KIT-G)

Presenter: GREIN, Guenter (KIT - Karlsruhe Institute of Technology)

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 50

Type: **Presentation**

The SARA/BigGrid HPC Cloud, From Beta to Production

Tuesday, September 18, 2012 12:00 PM (25 minutes)

Description of the work

We have been involved in High Performance Computing and later on Grid computing for a long time. Nevertheless, there are a number of users or user communities that for whom Grid or HPC infrastructures are not helping them to do their work. Those infrastructures may have the computational and I/O performance that they need but they lack flexibility. Clouds have that flexibility and a HPC Cloud can cater for the users performance needs. This has motivated us some two years ago to investigate the possibilities of a HPC Cloud infrastructure. We have setup a proof-of-concept type of environment and allowed a number of users to try it out. Their enthusiasm has encouraged us to replace the test environment with a full-fledged production HPC Cloud environment. In this presentation we will give an overview of our HPC Cloud infrastructure and discuss the experiences that we have gained from running our new production infrastructure. Extra features created by SARA for sharing of data and security of the Virtual machines within the user environment are also explained.

Link for further information

<https://www.sara.nl/systems/hpc-cloud>

Wider impact of this work

With this presentation we hope to get users and sites interested in HPC Clouds and we hope to share and exchange experiences with them.

Printable Summary

For over two years we have experience in running a HPC cloud at SARA. Starting out with a Proof-of-Concept type of environment we now have a production HPC Cloud infrastructure that we offer to our users. In this new infrastructure using Open Source solutions, users can startup their own HPC compute clusters having their own private high-band width interconnect between their VMs and a high-bandwidth connectivity to the outside as well as fast access to storage. In this presentation we will discuss the setup of our infrastructure, our experiences that we have gained so far and explain what was modified to the used open source tools.

Primary authors: Mr MASSCHELEIN, Jhon (SARA); TROMPERT, Ron (SARA)

Presenter: Mr MASSCHELEIN, Jhon (SARA)

Session Classification: Providing cloud services

Track Classification: Virtualised Resources: challenges and opportunities (Michel Drescher: track leader)

Contribution ID: 51

Type: **Presentation**

Messaging Service and Client Software

Tuesday, September 18, 2012 2:50 PM (20 minutes)

Description of the work

See summary.

Wider impact of this work

See summary.

Printable Summary

Messaging is an appealing technology which simplifies the integration between applications at different levels. It is used in production within EGI and being evolved in the EMI project.

Messaging is not just an API, it is a service like databases or web servers and it requires to be deployed on dedicated servers. In order to get the best in terms of reliability and scalability the EMI Messaging Product Team recommends to use dedicated messaging services made of independent brokers.

It is not trivial to write robust applications which make usage of messaging. In order to prevent code duplication and extra maintenance the EMI Messaging Product Team is providing a set of tools and libraries which can be used to build easily flexible and robust applications.

The proposed architecture for a messaging service and the set of tools and libraries will be presented along with their benefits.

Primary authors: CONS, Lionel (CERN); PALADIN, Massimo (CERN)

Presenters: CONS, Lionel (CERN); PALADIN, Massimo (CERN)

Session Classification: New middleware products

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 52

Type: **Demonstration**

Globus Accounting using Grid-SAFE

Wednesday, September 19, 2012 12:15 PM (15 minutes)

Description of the work

Grid-SAFE allows usage data from geographically distributed Globus GRAM5 computational sites to be stored, monitored and forwarded to centralised accounting resources. The aim of the session is to present a conceptual and practical overview of the capabilities of Grid-SAFE, from initial deployment to use in a Globus GRAM5 grid environment for resource operators.

Agenda:

1. Introduction to Grid-SAFE (~10 minutes): this will present an overview of Grid-SAFE itself, including its architecture and capabilities, as well as work achieved within IGE towards improvement, it's current technical and release status within IGE, and future plans and directions for the software.

1. Deployment and configuration of Grid-SAFE on a Globus GRAM5 grid (~10 minutes): Grid-SAFE is designed to have a straightforward process of installation. This demonstration will show the deployment of Grid-SAFE from initial installation to configuration as an accounting portal within a site context, as well as configuring it as a contributor to EGI centralised accounting resources.
2. Using Grid-SAFE to monitor grid job usage (~10 minutes): this will demonstrate the practical use of Grid-SAFE's functions to monitor job usage.

Wider impact of this work

IGE has built on the developments of Grid-SAFE by improving it as a rapidly deployable accounting solution tailored to Globus, enabling such Globus-oriented resources to conveniently contribute accounting usage data to the wider EGI accounting ecosystem. IGE is continuing to achieve this by:

- Greatly improving its ease of installation via IGE packaging
- Increasing its effectiveness as a monitor of Globus GRAM5 resources
- Progressing it as an external accounting client to centralised EGI accounting infrastructure

Printable Summary

Grid-SAFE allows usage data from geographically distributed Globus GRAM5 computational sites to be stored, monitored and forwarded to centralised accounting resources. The aim of the session is to present a conceptual and practical overview of the capabilities of Grid-SAFE, from initial deployment to use in a Globus GRAM5 grid environment for resource operators.

Agenda:

1. Introduction to Grid-SAFE (~10 minutes)
2. Deployment and configuration of Grid-SAFE on a Globus GRAM5 grid (~10 minutes)
3. Using Grid-SAFE to monitor grid job usage (~10 minutes)

Primary author: CROUCH, Stephen (University of Southampton)

Co-author: J.S.ROBINSON, j.s.robinson

Presenter: CROUCH, Stephen (University of Southampton)

Session Classification: Operations Workshops

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 53

Type: **Workshop**

Workshop: IGE Appliances for Grid Administrators and Users

Wednesday, September 19, 2012 2:00 PM (1h 30m)

Wider impact of this work

This contribution is of relevance to both grid administrators and end-users. This workshop will give practical insight on what is released by IGE as appliances. These appliances can be further used to easily and rapidly setup a grid environment based on GT 5.x, either for training purposes or for production use.

Printable Summary

The workshop is focused on the presentation of the appliances (virtual machines) provided by IGE to the European Grid ecosystem. It targets both, service providers and end-users, and it will cover the topics:

- An introduction to the components packaged by IGE (both software adaptations and GT5.x)
- IGE appliances available on the StratusLab Marketplace
- Short demo on how to setup a GridFTP, a GRAM5, or a GridWay etc. machine based on our appliances

The demo will also show the smooth integration of the IGE packaged software with the various operating system distributions. For that, the installation and configuration of an IGE component will be performed from scratch.

Description of the work

The workshop is focused on the presentation of the appliances (virtual machines) provided by IGE to the European Grid ecosystem. It targets both, service providers and end-users, and it will cover the topics:

- * An introduction to the components packaged by IGE (both Globus Toolkit 5.x and software adaptations):

Major functionality covered by the IGE packaged software will be briefly introduced: file transfer (GridFTP), job management (GRAM5), meta-scheduling (GridWay) etc.

- * IGE appliances available on the StratusLab Marketplace:

IGE releases appliances with pre-installed GT 5.x software. They will be introduced in this workshop.

- * Short demo on how to setup a GridFTP, a GRAM5, or a GridWay etc. machine based on our appliances

The demo will also show the smooth integration of the IGE packaged software with the various operating system distributions. For that, the installation and configuration of an IGE component will be performed from scratch.

Link for further information

<http://www.ige-project.eu>

Primary author: Dr LUCIAN MUNTEAN, Ioan (UTC)

Co-author: Dr COLESA, Adrian (Technical Univesity of Cluj-Napoca)

Presenter: Dr LUCIAN MUNTEAN, Ioan (UTC)

Session Classification: Cloud Management Solutions

Track Classification: Virtualised Resources: challenges and opportunities (Michel Drescher: track leader)

Contribution ID: 54

Type: **Training**

WNoDeS training

Tuesday, September 18, 2012 4:00 PM (1 hour)

Description of the work

The training session will be focused on two cases: the former deals with the installation and configuration of WNoDeS for batch jobs and the latter with the installation and configuration of WNoDeS for grid jobs. The pre-requisites and the technicalities for both cases will be illustrated and widely explained. The WNoDeS mixed mode feature and its configuration will also be described during the training.

Torque/Maui will be the reference batch system and Scientific Linux (either SL6 or SL5) the operating system for the tutorial.

The installation and configuration will then be validated by submitting both batch and grid jobs. At the end of the training session participants will be able to install, configure and run WNoDeS either on their grid site or on a standalone farm.

Link for further information

<http://web.infn.it/wnodes/>

Wider impact of this work

Large computing centers supporting a range of applications by communities need to leverage appropriate configurations or tools for the management of the computing resources in order to increase efficiency without decreasing the security level. Communities also need to have access to resources by a number of different interfaces, authentication methods and exploiting the existing local and distributed infrastructures. WNoDeS provides administrators and end users with several benefits for resource provisioning. Administrators can use the same set of services to manage the resources that are allocated dynamically to satisfy user requests. Resource centers can provide their users with a cloud infrastructure that is built upon non-invasive changes in their production computing infrastructure. End users can exploit grid and cloud interfaces based on their applications and requirements, potentially enjoying an increased number of available resources through easy and intuitive user interfaces.

Printable Summary

WNoDeS is a software framework to integrate grid and cloud provisioning through virtualization. Based on the needs of the INFN CNAF computing center (10,000 computing cores, 9 PB of disk space, 10 PB of tape space and about 80,000 jobs per day), WNoDeS has proved to be scalable and reliable; developed in a joint effort between INFN and the Italian Grid Infrastructure (IGI), in December 2011 has been accepted in the EMI software stack and has recently been released within the EMI-2 release. WNoDeS features a seamless integration with the EMI middleware that allows to transparently run Grid Jobs both on real and virtual Worker Nodes, instantiating virtual

machines from an existing VM repository.

The proposed tutorial is intended for site administrators and aims at providing an understanding of how to dynamically extend the resources at a site by installing, configuring and testing the WNoDeS service.

Primary authors: RONCHIERI, Elisabeth (INFN); GIORGIO, Emidio (INFN)

Co-authors: Mr ITALIANO, Alessandro (INFN CNAF); SALOMONI, Davide (INFN CNAF); DONVITO, Giacinto (INFN); DALLA TORRE, Gianni (INFN CNAF); GAIDO, Luciano (INFN)

Presenters: RONCHIERI, Elisabeth (INFN); GIORGIO, Emidio (INFN)

Session Classification: New middleware products

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 55

Type: **Presentation**

Monitoring Virtual Machine Status with L&B

Wednesday, September 19, 2012 4:00 PM (30 minutes)

Description of the work

The number of “job” types L&B recognizes has been extended with two additional types: one covering the states of an individual virtual machine (VM) and one corresponding to a collection of virtual machines, representative of a cluster. All the other mechanisms such as event delivery, storage, querying API, notifications or statistics are already in place in L&B. A single instance of an L&B server in its standard setup can be used to monitor compute jobs and VMs at the same time. The basic VM state diagram is a simplified version of the state diagram defined by the Open Nebula toolkit. A prototype has been setup with Open Nebula as the virtualization stack. It provides call-out hooks on major events, which makes it easy to instrument L&B event delivery. Events can be delivered either over L&B’s legacy event delivery chain, or through messaging (STOMP/OpewWire), which leaves potential adopters with a choice of mechanism. The L&B team is looking further at other messaging protocols applicable to event delivery. L&B events can be sent not only from the virtualization stack, but also from Dom0 or even from within the virtual machine itself. This makes some of them partly redundant, but in fact this is building up on one of L&B’s other advantages since redundant events make for a more reliable state determination in case of localized failures, and distinguishing between similar events received from various sources provides for fine-grained status monitoring.

By abstraction a VM state diagram is also applicable to physical machines. Since L&B can record relationships between individual processes it monitors, such as compute jobs and sandbox transfers, or compute jobs and virtual machines, it can also report on the relationship between a virtual and physical machine. This information is often found useful by infrastructure administrators.

Link for further information

<http://egee.cesnet.cz/en/JRA1/LB/>

Wider impact of this work

Although the solution has been prototyped with Open Nebula, it is intended to combine with other widely used virtualization stacks. There are multiple options for instrumenting event delivery. L&B client library provides C or Java bindings and command-line tools are also available for use in scripting languages. Some virtualization solutions, OpenStack for instance, already use messaging at the moment. The possibility of tapping into that source of information and using it “as is” to monitor its processes is another opportunity for further investigation.

Using a common monitoring tool is advantageous for heterogeneous infrastructures relying on multiple virtualization solutions at once. L&B can become a common point of reference, or a source of unified-format notifications to be used by other elements for further processing.

Printable Summary

gLite's Logging and Bookkeeping is a well-established tool for monitoring processes in the grid, typically compute jobs. It collects event information from various grid elements and sums it up to determine the current status of any such process at the given moment. Aside of various types of compute jobs L&B has also been used to monitor the states of sandbox transfers. With a great portion of the grid infrastructure becoming virtualized, a virtual machine is becoming a "process" just like a normal job, with a distinguished lifetime, a state diagram with well defined state transitions, and a similar requirement for status monitoring. Once the likeness becomes apparent, it is obvious that monitoring can be done by the same tools. Another advantage of using the same tool for monitoring virtual machines and compute jobs, which run on them, is that it makes it easy to keep track of the relationship at the same time, which simplifies problem solving.

Primary authors: KOURIL, Daniel (CESNET); Mr DVORAK, Frantisek (CESNET); Mr FILIPOVIC, Jiri (CESNET); SITERA, Jiri (CESNET); MATYSKA, Ludek (CESNET); Mr POUL, Marcel (CESNET); VOCU, Michal; SUSTR, Zdenek (CESNET)

Presenters: Mr FILIPOVIC, Jiri (CESNET); SUSTR, Zdenek (CESNET)

Session Classification: Cloud Management Solutions

Track Classification: Virtualised Resources: challenges and opportunities (Michel Drescher: track leader)

Contribution ID: 56

Type: **Poster**

Seamless Integration of BOINC Applications into MetaCentrum

Description of the work

The Laboratory of Security and Applied Cryptography (LaBAK), Faculty of Informatics, Masaryk University, Brno uses its simulation platform to run a wide set of computations generated by its current research focus - genetic programming of security protocols and optimization of intrusion detection systems for WSNs. Its demands exceed the capacity of the available 280-CPU computer lab. The research group has a strong potential to achieve more results if short-term peaks in demand of computing power are satisfied. The MetaCentrum virtual cluster service is based on the idea that providing a virtual machine is just another form of a computing job. Our scheduler controls resource assignment regardless of job type (job/VM) and we want to demonstrate that even Windows-based infrastructure controlled by BOINC workload manager can be seamlessly provided by our resources. The application is an in-house Windows-based development of the research group with port to the MetaCentrum (Debian Linux, batch manager) possible, but costly. From previous experiments, we had MetaCentrum Windows golden image ready. Licenses represented an issue but turned out to be relatively cheap as long as research prices (MSAA) remain applicable. The networking-related part is important as well - all the virtual machines appear to be locally connected to the BOINC server. It is achieved by a dedicated VLAN which is by design part of the virtual cluster service. In our case the VLAN was connected directly to the BOINC server machine by a permanent OpenVPN tunnel. The new L2 segment is completely private, not routed to the Internet. It is permanently seen by BOINC, dynamically created machines are connected to this VLAN by MetaCentrum's virtual cluster management software. The BOINC-based workload manager proved itself well scalable simplifying the Windows image complexity (discovery, no need for other services) and the overall architecture (robustness, dynamic nature of VMs – an automatic resubmit).

Link for further information

<http://www.metacentrum.cz/>

Wider impact of this work

This work is important as a demonstration of Czech NGI's strategy to be flexible and accessible for established or emerging research groups who have specific requirements or already operate their own infrastructure. One of the key roles of an NGI consists in covering the peaks in demand of established groups otherwise satisfied by their own resources, admins and infrastructure tools. We also want to show that this can be done under one resource broker umbrella – without a resource pool dedicated solely for the virtualized services (cloud). From the research group's point of view we demonstrate the possibility to expand their current infrastructure and enter MetaCentrum with minimal changes in SW and computational model. We have already demonstrated a 500-CPU virtual cluster but we are heading for a bigger one.

Printable Summary

For MetaCentrum, the Czech NGI, the strategic aim is to narrow the gap an established user group must cross to use the NGI's infrastructure. We have developed "virtual cluster" virtualization technology to implement that strategy. This poster gives an example of one application group relying on stable and established infrastructure based on BOINC running on university computer lab hardware. The infrastructure runs security simulation experiments for wireless sensor networks (WSNs) with Sensor Security Simulator (S3), a software developed by this research group. Thanks to the virtual cluster service we have been able to provide MetaCentrum's resources as a dynamically available extension to the group's current infrastructure seamlessly, without any significant investment on the user group's side. Additional resources appear as new BOINC worker nodes running original Windows environment available over the local network segment (VLAN).

Primary author: SITERA, Jiri (CESNET)

Co-authors: MULAČ, Miloš (CESNET); Mr RUDA, Miroslav (Cesnet); VONDRUŠKA, Pavel (CESNET); ŠVENDA, Petr (Masaryk university, Brno, Czech Republic); STIBOR, Tomáš (CESNET); SUSTR, Zdenek (CESNET); TÓTH, Šimon (CESNET)

Presenter: SITERA, Jiri (CESNET)

Track Classification: Virtualised Resources: challenges and opportunities (Michel Drescher: track leader)

Contribution ID: 57

Type: **Presentation**

Enhancing the CREAM-CE with the High Availability Cluster

Tuesday, September 18, 2012 3:10 PM (20 minutes)

Description of the work

This work is focusing on enhancing CREAM service with the High Availability. It will guarantee the ability of the CREAM service to be continuously available for serving the user requests even during planned and unplanned outages. The whole system complexity will be hidden to the user whereas the overall service will benefit of an increasing of performance, scalability, availability and fault tolerance which are relevant features in the production environment. We will show the relevant criteria adopted for modeling the envisaged CREAM architecture based on clustering paradigm. We will discuss about its implementation by highlighting the entailed issues: the merging and/or logical centralization of information coming from different sources in the cluster, the potential bottlenecks and SPOFs (Single Point Of Failures) which affect several services used by CREAM-CE, such as databases, sandboxes, BLAH and how those components should be carefully replicated. Moreover we will present the status of the work.

Wider impact of this work

This work will provide the user the guarantee to take advantage of a CREAM service continuously available and enhanced with high performance for serving her requests even during planned and unplanned outages. The whole system will gain of features desirable and required in all production environments.

Printable Summary

One of the new functionalities foreseen to be implemented after the second EMI (European Middleware Initiative) major release in the CREAM-CE is the ability to be continuously available for serving user requests even in case of planned and unplanned outages. Like several popular Internet services, the idea is to achieve this goal relying on a cluster of commodity computers seen as a single service. This won't allow only to implement the High Availability (HA) capability, but also to improve the overall scalability and performance. The whole system complexity must of course be hidden to the user, who is not interested in distinguishing a single CREAM service from the clustered one. We present the envisaged architecture and the status of the developments.

Primary authors: Dr DORIGO, Alvise (INFN); Dr AIFTIMIEI, Cristina (INFN); Dr DONGIOVANNI, Danilo (INFN); Dr REBATTO, David (INFN); Dr FRIZZIERO, Eric (INFN); Dr CAPANNINI, Fabio (INFN); Dr PRELZ, Francesco (INFN); Dr ZANGRANDO, Lisa (INFN); Dr MEZZADRI, Massimo (INFN); Dr SGARAVATTO, Massimo (INFN); Dr ANDREETTO, Paolo (INFN); Dr BERTOCCO, Sara (INFN); Dr TRALDI, Sergio (INFN)

Presenters: Dr ZANGRANDO, Lisa (INFN); Dr ANDREETTO, Paolo (INFN)

Session Classification: New middleware products

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 58

Type: **Presentation**

Using R, self scaling Matlab clusters and Galaxy clusters on a HPC cloud infrastructure

Wednesday, September 19, 2012 5:00 PM (30 minutes)

Description of the work

BiG Grid, the Dutch NGI started with a production IAAS HPC cloud infrastructure based on OpenNebula and KVM. With a self-service portal individuals and projects can have access to a virtualised infrastructure they can configure at will.

Offering users and projects such a development opportunity created a lot of creative projects and approaches.

Offering services such as Matlab, R and Galaxy workflows as a hosted infrastructure that scales on demand emerged as the way to go.

In the past 10 months we've been co-developing and implementing these type of compute services.

In this talk these experiences are presented, as well as issues and challenges.

Link for further information

www.biggrid.nl

www.cloud.sara.nl

Wider impact of this work

We think the community can learn from our development approach and can be inspired by the possibilities of having a IAAS / SAAS infrastructure that can be used as a piloting test-bed to quickly develop such dynamically scaling clusters.

We also welcome input for collaboration and discussion.

Printable Summary

BiG Grid, the Dutch NGI started with a production IAAS HPC cloud infrastructure based on OpenNebula and KVM. With a self-service portal individuals and projects can have access to a virtualised infrastructure they can configure at will.

In this talk experiences with R, Matlab(tm) and Galaxy (cloudman) are presented, where a clear shift from IAAS to SAAS is shown.

Primary author: VISSER, Tom (SARA)

Presenter: VISSER, Tom (SARA)

Session Classification: Cloud Management Solutions

Track Classification: Virtualised Resources: challenges and opportunities (Michel Drescher: track leader)

Contribution ID: 59

Type: **Presentation**

GridWay interoperability through BES

Thursday, September 20, 2012 12:00 PM (30 minutes)

Description of the work

The interaction between grid infrastructures and middlewares is still a challenging problem. Therefore, interoperability is a desired capability of grid infrastructures to allow VOs access the resources provided by the existing infrastructures although based on different middlewares. Otherwise, interoperation techniques are needed.

GridWay provides support for some of the few established standards for interoperability, and it also provides components for interoperation. Its modular architecture is based on drivers, acting as adapters for different grid services providing resource discovery and monitoring, job execution and management, and file transfer.

GridWay now provides support for the OGF OGSA Basic Execution Service standard both as a client and as a server, by means of a new execution driver and a new interface, respectively. BES addresses interoperability between job management services and is pursuant toward job submission standardization. The OGSA BES specification defines WS interfaces for creating, monitoring, and controlling computational entities such as processes, Web Services or parallel programs, called activities, within a defined environment.

Given the importance of interoperating with BES-enabled endpoints, a BES driver for GridWay has been developed. The new BES driver provides an abstraction layer that enables users to submit jobs through GridWay to BES interfaces, and control and monitor their execution.

A BES service-level interface for GridWay is now also supported, which provides access to GridWay's job metascheduling capabilities using the BES implementation of GridSAM. Communication between both systems is based on the Java DRMAA API. It is worth noting that GridSAM provides a job submission interface for submitting computational jobs to many LRMS, but it does not provide metascheduling capabilities.

Both the BES driver and interface were demonstrated in March 2012, in the context of the Grid Interoperability Now OGF group.

Link for further information

<http://gridway.org/doku.php?id=ecosystem:gridwaybes>

<http://www.omii.ac.uk/wiki/GridSAM>

<http://www.ige-project.eu>

Wider impact of this work

This work is motivated by the need of interoperation and interoperability on grid environments, addressing the interoperation at the metascheduler level. The main contribution of this work is the support of a new standard for interoperability in GridWay, and presents new components based on BES to enable the interoperation of different grid technologies. It provides end-users a metascheduling tool able of interoperating with the most used grid technologies.

The BES driver for GridWay allows to users to submit, control and monitor their execution jobs to sites managed by a standard service. The BES interface for GridWay allows users to access metascheduling capabilities using an open-standards based interface for job submission.

Printable Summary

The GridWay metascheduler enables large-scale, reliable and efficient sharing of computing resources over different grid middlewares, providing a single point of access for them. The current development of GridWay is being supported by the IGE project.

This work addresses the problem of interoperability at the metascheduler level. The Basic Execution Service (BES) standard address interoperability between job management services and pursues for job submission standardization. GridWay provides support for BES both as a client and as a server, providing to end-users more possibilities of interoperability. Thus, a BES driver for GridWay that enables the interoperability with BES-enabled endpoints, and a BES-compliant interface that enables the remote access to GridWay's metascheduling capabilities through a standard interface, have been developed.

Primary authors: HUEDO, Eduardo (UCM); M. LLORENTE, Ignacio (Universidad Complutense de Madrid); Dr MARÍN CARRION, Ismael (Complutense University of Madrid); CROUCH, Stephen (University of Southampton)

Presenter: Dr MARÍN CARRION, Ismael (Complutense University of Madrid)

Session Classification: Resource Infrastructure Services

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 60

Type: **Demonstration**

Monitoring National Infrastructure with L&B

Description of the work

Besides its contribution to EGI, the Czech NGI operates its own distributed environment – the METACentrum. In addition to traditional computing clusters METACentrum has been providing virtualization services, either relying on its own virtualization infrastructure that allows creation of virtual clusters or extension of physical clusters with virtual nodes, or providing pure Open Nebula-based cloud services. Naturally the batch system can be used to schedule jobs for physical and virtual nodes alike. There are multiple instances of the batch system capable of forwarding jobs among themselves to improve load balancing or find alternatives in match making. L&B has been selected as a monitoring tool for the infrastructure for its ability to monitor processes whose events arrive from different nodes in a grid, to provide reliable message delivery and up-to-date status information regardless of occasional grid component or communication failures.

To achieve the goal L&B has been extended with state diagrams for torque jobs and virtual machines. Instrumentation of L&B logging calls from Torque has been implemented at the source code level, while virtualization stacks (both Open Nebula and home-brewed Magrathea) make use of existing callback hooks to log events through command line executables. Similar logging calls are made from Dom0 and also from within the virtual image, so that events triggering high-level state changes are redundant and distinguishing between similar events received from various sources provides for fine-grained status monitoring.

Another reason to choose a single monitoring tool for different processes is its ability to keep track of relationships. ID of a VM used to run a given computing job can be stored with the job. Similarly, with just a simple abstraction, VM state diagram can also be used for physical machines, and then a similar relationship can be established between a VM and the physical resources used to execute it.

Link for further information

<http://egee.cesnet.cz/en/JRA1/LB/>

Wider impact of this work

The work described here is intended not only as a production-level solution for METACentrum, but also as a prototype of a general-purpose cloud monitoring solution. Since L&B simply connects up to the infrastructure through messaging (supporting not only its own messaging protocol but also STOMP or OpenWire), it can be used to monitor any virtualized infrastructure capable of sending event messages over any of these channels. L&B also brings in other advanced features such as notifications or user tags that can be stored with other status information. It supports multiple authentication mechanisms, namely PKI (used across the EGI grid) and Kerberos (used in METACentrum).

The presented solution is also a prototype of integrating L&B's HTTPs interface into a Web portal to bring user-related information directly to the users' personal pages.

Printable Summary

gLite's Logging and Bookkeeping (L&B) is a monitoring tool equipped for monitoring the states of all kinds of processes related to grid and cloud computing. Besides traditional gLite WMS jobs and logical groupings thereof such as DAGs or collections it also supports input/output sandbox transfers, native CREAM jobs, Torque jobs and, as a recent addition, Virtual Machine states. With that, L&B has been deployed over the Czech NGI's infrastructure to provide users with a uniform view of all their processes, be it traditional jobs submitted to Torque-managed clusters, virtual machines managed through Czech NGI's own virtualization solution, or Open Nebula Cloud Machines. Where applicable, mutual links between jobs, virtual machines and physical machines are also recorded and made available as a part of the status information.

Primary authors: KOURIL, Daniel (CESNET); Mr DVORAK, Frantisek (CESNET); Mrs KRENKOVA, Ivana (CESNET); CHUDOBA, Jiri (CESNET); Mr FILIPOVIC, Jiri (CESNET); SITERA, Jiri (CESNET); MATYSKA, Ludek (CESNET); Mr POUL, Marcel (CESNET); VOCU, Michal; Mr RUDA, Miroslav (Cesnet); Mr TOTH, Simon (CESNET); SUSTR, Zdenek (CESNET)

Presenter: SUSTR, Zdenek (CESNET)

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 62

Type: **Presentation**

Cloud infrastructure for the on demand provisioning of Worker Nodes

Thursday, September 20, 2012 2:00 PM (30 minutes)

Wider impact of this work

This work, in a way, introduces elasticity to the Grid. Using if not the same but also similar implementations administrators of Grid resources can modify on demand the number computing resources offered based on certain thresholds (i.e. the total number of submitted jobs and the underlying number of physical job slots). We feel that the user experience of the Grid will be enhanced through such implementations, especially for users that rely heavily on parametric job types or large parallel jobs (or both).

Printable Summary

We showcase the development and usage of a Quattor based Openstack reference cloud which is used for the on demand provisioning of additional Worker Nodes under our Grid based infrastructure. The need of being able to add computing resources on-the-fly has gradually emerged over the years as a way to leverage the large number of jobs that may occasionally be routed towards a Grid site. After considering several PaaS based alternatives that could be used to harness such a need we decided to deploy and operate an Openstack based cloud.

Description of the work

Using an Openstack reference cloud we introduce the concept of elasticity to our Grid site by adding and removing on demand computing resources in the form of virtualized Worker Nodes (vWNs). The vWNs are added under the central batch job queueing system whenever a large number of jobs is directed towards the site (i.e. whenever the number of queued jobs exceeds the total number of physical job slots offered). Such situations are not uncommon in our experience as several users rely heavily on the usage of parametric jobs. Large batches of parametric jobs may be directed by the WMS to a single site as at the time of submission the WMS does not calculate dynamically the impact of the total number of jobs on the Grid site but rather treats them independently using the information supplied at some earlier given point in time by the information system. By bringing elasticity to the Grid we manage to leverage such abrupt demands on computing resources by deploying vWNs to handle the large number of jobs. Distinguishing among serial and parallel jobs we do so by provisioning two types of instances one for single CPU jobs and one for multi-CPU jobs consisting in the later case of 8 CPUs per instance. Once the number of queued jobs decreases to a number lower than the number of available job slots (physical not virtualized) no additional vWNs are created and the system is left to 'cool off'.

Primary authors: KANELLOPOULOS, Christos (GRNET); TRIANTAFYLLIDIS, Christos (GR-

NET); KOROSOGLOU, Paschalis (GRNET)

Presenter: TRIANTAFYLLIDIS, Christos (GRNET)

Session Classification: Virtualised Resources

Track Classification: Virtualised Resources: challenges and opportunities (Michel Drescher: track leader)

Contribution ID: 63

Type: **Presentation**

HINTS network trouble shooting - Latest news

Tuesday, September 18, 2012 4:30 PM (30 minutes)

Printable Summary

HINTS ("Hints for Instantaneous Network Troubleshooting") is a lightweight quick network troubleshooting tool, delivered to the GRID end user as a WEB page.

With "probes" installed in Grid sites, one can run on-demand basic bidirectional tests (Ping, Traceroute, Nmap, DNS resolving, Iperf) between two of this probes.

On the server side, the registered users can involve only probes they are allowed to use in their tests.

By using HINTS, a user does not require any specific network privilege to run basic tests between two GRID sites, thus preventing the involvement of the network administrators at the first diagnostic actions.

This presentation should be attended by Network Support NGI representatives and anyone from the Network Support Community.

Description of the work

Originally developed within EGEE SA2, HINTS has been endorsed as the suggested tool for Network Troubleshooting on demand by the EGI Network Support.

In this presentation the overall structure and basic principles of HINTS will be described, together with the deployment carried out so far and the provided functionality.

Work has been carried out on the packaging to provide easily deployable probes for site administrators.

The recent work on HINTS consists mainly in:

- Providing SL5 binaries (today, only SL6 binaries are available)
- Thinking of HINTS as a service more than a server and its probes. The idea would be to provide a "community" HINTS server, preventing every NGI from installing and running their own HINTS server. By delivering such a service, installing HINTS would be much easier and effortless to run.
- Though HINTS' probes are heavily based on PerfSONAR MDM "Measurement Points", someone using both systems today has to install both binaries, and moreover, on two separated servers.

In a joint initiative with the PerfSONAR MDM team, we are working on letting HINTS server use already installed PS Measurement Points. Again, the purpose is to make HINTS easier to install, and particularly this point would greatly increase the number of usable probes. And of course, an NGI would no more have to run two different servers, one for PF's Measurement Point, and one for HINTS' probe.

Link for further information

<https://aresu.dsi.cnrs.fr/spip.php?rubrique27>

Wider impact of this work

Provide the EGI GRID user community with a lightweight, easy tool to perform on-demand network troubleshooting.

Primary author: LENORMAND, Olivier (CNRS)

Co-author: Dr REALE, Mario (GARR)

Presenter: LENORMAND, Olivier (CNRS)

Session Classification: Network Support

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 64

Type: **Presentation**

D4Science Infrastructure: a novel open approach to distributed software management based on Maven

Thursday, September 20, 2012 11:40 AM (20 minutes)

Description of the work

gCube, the D4Science infrastructure powering technology, is a JAVA software framework featuring the declarative and interactive creation of transient Virtual Research Environments by aggregating and deploying on-demand content resources and application services. It has been designed to exploit the peculiarities of modern application servers (JAVA WS-core, Tomcat, Jboss), which ensure Web-services/web-applications “hot” deployment, and to bridge existing private and public cloud systems. Since its inception, the focus of the framework has been on the optimal resource allocation and management of software packaged following the gCube policies. This was indeed a weak point of the system imposing custom packaging rules and limiting de-facto the software potentially deployable on the infrastructure. To overcome it, the gCube enabling layer recently took an extended approach for dealing with software available as artifact in an approved Maven Repository. This also had the great advance of introducing Maven as official distribution repository for gCube software. The approach firstly comprised the exploitation of Maven within the build and integration process. Then, it required the development of a new service, the Software Gateway, which acts as a gateway over a cluster of Maven Repositories granting access to the stored information and software for deployment purposes. And finally, all the rest of the gCube services managing software were adapted to exploit the new deployment model, not based on the previous policies. We will report on this new open approach along with the extensions applied to the gCube resource model to cope with maven artifacts. To give a complete picture of the impact on the development, we will cover major aspects of the gCube build process, the “mavenization” of the legacy gCube artefacts and the activities performed at integration side to deal with Maven. Lastly, a look ahead to future activities to support deployment of non-JAVA based artifacts.

Link for further information

<http://www.gcube-system.org/>

https://gcube.wiki.gcube-system.org/gcube/index.php/Data_e-Infrastructure_Management_Facilities

Wider impact of this work

Simplified management and deployment of software coming from third-party repositories is one of the main goals in a growing and open environment as the D4Science infrastructure. Communities coming from different EU projects, with different requirements in term of software management are today exploiting the infrastructure. The introduction of a de-facto standard as Maven orthogonal to many aspects of software management has been a major step forward towards interoperability and sustainability. From one side we enable users’ applications to be deployed on the infrastructure at nearly “zero” cost. From the other side, integrating Maven grants access to gCube

software to other Maven-systems and opens the possibility to adopt Maven at different levels. This is a crucial point to evaluate as the end of EMI project's end is approaching and the future of its building facilities, greatly exploited by gCube since many years, is yet unsettled.

Printable Summary

The D4Science infrastructure is a Hybrid Data Infrastructure (HDI) deployed and maintained throughout three EU projects (DILIGENT, D4Science I and II) and actually supporting two EU projects (iMarine and EUBrazilOpenBio). A Hybrid Data Infrastructure is an innovative approach based on the integration of several technologies, including Grid and private/public Clouds, to provide an elastic access to and usage of data and data-management capabilities. Equipped with services supporting the creation of Virtual Research Environments, it creates dynamic and distributed applications tailored to serve a specific need whose constituents are acquired by the HDI. We report on a major extension performed on the infrastructure: the deployment and activation of Maven artifacts. This is a key aspect for the sustainability of the infrastructure promoting the transparent exploitation of third-party applications in a Virtual Research Environment.

Primary authors: MANZI, Andrea (CERN); Dr SIMEONI, Fabio (FAO); Mr GIAMMATTEO, Gabriele (ENG); Mr SIMI, Manuele (ISTI-CNR)

Co-author: PAGANO, Pasquale (CNR - ISTI)

Presenter: MANZI, Andrea (CERN)

Session Classification: Virtualised Resources

Track Classification: Virtualised Resources: challenges and opportunities (Michel Drescher: track leader)

Contribution ID: 65

Type: **Presentation**

Log and reporting architecture for achieving compliance in distributed and multitenant infrastructures

Description of the work

Although the distributed computing models and virtualization technologies have introduced substantial benefits, the fact that both physical and software resources can be geographically distributed and shared by different users, even from different administrative domains, has heightened the common security and regulatory issues of traditional IT infrastructure. Most important initiatives on security compliance, including CSA (Cloud Security Alliance) and ENISA, emphasize the importance to certify the resource offering to the common security standards, including ISO27002, PCI-DS and audit framework such as SAS-70 II and the need to have specific audit event, log and report management mechanisms providing the evidence of adherence to the reference security and/or regulatory framework. In this presentation we will describe an architecture of a SIEM (System Information and Event Management) based, scalable and flexible system which can be deployed in a variety of distributed and virtualized infrastructure and which provides the following capabilities: collection in a secure way of audit messages from different nodes (and in different format) of the distributed infrastructure, normalization of them in a standard and common format independent from the source, application of security policies over the normalized messages and (in case of events that may require corrective actions or other types of responses) generation of alert, summary of data in reports in conformity with ISO 27k, PCI-DSS and HIPAA. The proposed solution allows defining security controls and generate alerts and report on a perimeter or a per-tenant basis. These controls are mainly related to: user authentication, access to resources, attacks from unknown or untrusted sources, attacks/infections at network/host level, operations on services (such as install, invoke, remove, terminate), system corruption and hardware failure, access to sensitive data, geolocalization of data, data retention/transfer/deletion.

Wider impact of this work

The presentation delves into what has been traditionally an IT problem and which has been exacerbated in the area of distributed and virtualized infrastructure in a real-life business, that is the compliance to regulations and security certification standards.

In this context, the proposed and standard based solution provides key capabilities for the security management in this type of infrastructure, while accomplishing the most important requirements of scalability, failover and high-performance.

Adopting such a solution may help to overcome one of the most difficult barriers for organizations considering moving to a distributed and virtualized infrastructure.

Printable Summary

Addressing compliance to security certification standards such as ISO27002, PCI-DSS and to EU directives on data privacy is more and more revealing to be a key factor in the adoption of distributed

and virtualized computing model in the real-life business. In this context, logging, monitoring, auditing and reporting practices, while transcending the compliance regimes, represent the primary instrument of assurance for security manager and auditors that compliance objectives are being met or, if not fully met, then progressively improved.

The aim of this presentation is to provide an overview of impelling requirements coming from most common security certification standards and regulation, and then to present an architecture of a logging and reporting component that aim at supporting infrastructure providers to achieve compliance objectives.

Primary author: Mrs BONELLI, Lucia (Engineering Ingegneria Informatica spa)

Co-authors: Mr MANIERI, Andrea (Engineering Ingegneria Informatica spa); Mr IMMEDIATA, Angelo (Engineering Ingegneria Informatica spa); Mr LUZZI, Antonio (Engineering Ingegneria Informatica spa); Mrs GIUDICIANNI, Luisa (Engineering Ingegneria Informatica spa)

Presenter: Mrs BONELLI, Lucia (Engineering Ingegneria Informatica spa)

Track Classification: Virtualised Resources: challenges and opportunities (Michel Drescher: track leader)

Contribution ID: 66

Type: **Presentation**

Remote analysis of human voice - environment for voice training and ORL medicine

Thursday, September 20, 2012 12:00 PM (20 minutes)

Link for further information

<http://medicus.cesnet.cz/doku.php/publikace>

Description of the work

The remote desktop protocol (RDP) transfers events from mouse and keyboard from user's client application to the remote system where it is interpreted and graphical changes are transferred back to client's application which visualizes it. RDP protocol also introduces some support in sound recording redirection over RDP, however, due to our measurements, we decided to introduce our own customization of RDP protocol to redirect sound recording over RDP without loss of information and control of quality of recorded data.

The system is being used in the research of voice training of students of pedagogy and singing and by some otorhinolaryngologists.

The test deployment on the cloud provided by the METACENTRUM were tested with the promising results and some challenges. The cloud environment OpenNebula allows user friendly interface to deploy virtual machines almost without intervention of some professionals or support. The intervention were needed if we installed a custom virtual machine with MS Windows operating system which is prerequisite of the analytical application. The license agreement of the operating system needs to be fulfilled also on the virtual machine in the cloud. This may cause some troubles to providers who doesn't have appropriate license schema and currently METACENTRUM doesn't provide some multilicense choice for MS Windows operating systems.

Wider impact of this work

The RDP protocol is mostly used for administration purposes. Another web technologies are used to provide application on the web - HTTP, HTML, Javascript and others. However the introduced work shows an application which can be accessed remotely via RDP without complicated reimplementation and provide an application support to establish and cooperate via virtual research environment. Deployment of the application to the cloud removes the maintenance task of hardware from provider of the application to the provider of the cloud. However, the maintenance should be moved also on the level of operating system and generic scientific software. This and other expectations from scientific cloud providers are now being discussed within the research community.

Printable Summary

The aim of the project FONIATR described within this presentation is to built and maintain a system to support remote access to analytical tools and cooperation of specialists in the field of

human voice pathologies (othorinolaryngologist) and in the field of speech or voice training.

The analytical application can analyze input signal - human voice and provide a graphical output in terms of voice range profile supporting decision and cooperation of specialists. On top of that the anonymized analysis of voices would provide a database for further statistical research.

The application is deployed on the server and can be accessed by remote desktop protocol.

The METACENTRUM activity of CESNET provides access to scientific cloud. The deployment on the test cloud provided by the METACENTRUM were tested with the promising results and challenges of license of underlying software, cost of services etc.

Primary author: Mr KULHÁNEK, Tomáš (CESNET)

Co-author: Dr FRIČ, Marek (Academy of Performing Arts in Prague)

Presenter: Mr KULHÁNEK, Tomáš (CESNET)

Session Classification: Virtualised Resources

Track Classification: Virtualised Resources: challenges and opportunities (Michel Drescher: track leader)

Contribution ID: 67

Type: **Training**

EMI 2 Matterhorn - tutorial for system administrators

Wednesday, September 19, 2012 11:00 AM (10 minutes)

Description of the work

This system administrator oriented tutorial is based on the second release of the EMI middleware, EMI-2 “Matterhorn”.

It will provide an overview of the EMI services for Computing and Data Services, providing then a practical tutorial about the installation and configuration of the most used services, giving concrete examples and highlighting best practices and most common mistakes.

Link for further information

<http://www.eu-emi.eu>

Wider impact of this work

EMI products are increasingly being adopted from NGI's. This training is directed mainly toward new grid system administrators, although also experienced ones can find it beneficial by refreshing their knowledge.

Printable Summary

The European Middleware Initiative (EMI) delivers a consolidated set of middleware components for deployment in Distributed Computing Infrastructures, extending the interoperability between grids and other computing infrastructures, strengthening the reliability of the services, and establishing a sustainable model to maintain and evolve the middleware that will fulfil the requirements of user communities. EMI middleware is not built from ground up, but rather, it delivers a consolidated and streamlined set of services and components from existing middleware projects ARC, gLite, UNICORE and dCache by re-factoring existing components. It also defines and implements standards and phases out duplicate or obsolete components from the original middleware stacks. The middleware components are divided into four areas (Compute, Data, Security, Infrastructure). This tutorial will show how to install and configure the most commonly used services for job and data management.

Primary author: GIORGIO, Emidio (INFN)

Co-author: CASSIDY, Kathryn (TCD)

Presenters: GIORGIO, Emidio (INFN); CASSIDY, Kathryn (TCD)

Session Classification: EMI 2 Matterhorn - tutorial for system administrators

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 68

Type: **Presentation**

Towards Deployment of EMIR in Federated Grid Infrastructures

Tuesday, September 18, 2012 2:10 PM (20 minutes)

Description of the work

The National Grid Infrastructures (NGI) in EGI are autonomous in nature and based on heterogeneous Grid middlewares. The lack of common decentralized service discovery would enforce NGI or EGI operators to manage multiple middleware specific registries. Alternatively a centralized indexing registry has to be setup, which however could easily become a bottleneck and susceptible to single point of failure. The EMI's EMIR offers a distributed service registry which has been developed using the concept of independent NGIs, whereby each EMIR server instance can be deployed on each participating site and can be connected to top level NGI, thus forming a hierarchical network of EMIR registries. Furthermore all the participating middleware services can be indexed in a controlled fashion while deciding which services should (or should not) be available at the ancestor registries. While the federations such as EGI could leverage it from setting up the global EMIR registry which offers robustness while replicating among other global EMIR registries.

Wider impact of this work

It is indispensable to have autonomous service registries which enables robust mechanisms to discover the services in federations like EGI. The talk will primarily focus on proposed deployment scenarios of EMIR in EGI and NGIs.

Printable Summary

In modern Grid infrastructures where multi-middleware services exists, a unified mechanism to service discovery is fundamental. These services often belong to specific type of middleware, therefore a number of (middleware) specific service registries has to be setup to enable service look-up and discovery. Within EMI, a common service registry (EMIR) has been developed which offers a unified mechanism of discovering the federation wide heterogeneous Grid services. This however relieve the operators from configuring and deploying existing/emerging service registries. The EMIR has been designed to create distributed registry infrastructure, where multiple EMIR servers can be connected hierarchically and/or in P2P fashion to offer robust Grid wide service discovery.

Primary author: SHIRAZ MEMON, ahmed (JUELICH)

Co-authors: Mr SZIGETI, Gabor (Niif); MARTON, Ivan (NIIF Institute); FIELD, Laurence (CERN)

Presenters: FIELD, Laurence (CERN); SHIRAZ MEMON, ahmed (JUELICH)

Session Classification: New middleware products

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 69

Type: **Presentation**

New Types of Accounting Records

Wednesday, September 19, 2012 11:00 AM (20 minutes)

Printable Summary

EGI has been planning to record accounting of additional types of resource, in addition to the long-standing cpu usage record. This session will describe several strands of work in EGI, EMI, and other projects to define and collect usage records for storage, virtual machine/cloud use, MPI jobs, and applications.

Description of the work

CPU accounting data has been gathered in EGI, EGEE, and the Worldwide LHC Computing Grid since 2004. Over the last couple of years EGI has been sponsoring and encouraging the accounting of other resources such as the use of virtual machine in clouds, storage use, parallel jobs and the use of applications.

This presentation will give a status update on this work from a variety of sources, including work on StAR, a storage accounting usage record in EMI and OGF and the accounting of VM use in the EGI Cloud task force.

Wider impact of this work

EGI has been the focus of worldwide accounting, acting as a central repository for international VOs and storing records for them from other infrastructures too.

Primary author: Dr GORDON, John (STFC)

Presenter: Dr GORDON, John (STFC)

Session Classification: Operations Workshops

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 70

Type: **Presentation**

Transporting Accounting Data using SSM

Wednesday, September 19, 2012 12:00 PM (15 minutes)

Description of the work

SSM was written to use the EGI broker network to send and receive messages securely. It uses X509 certificates for encryption and signing of data in transit. It has been extensively tested and is now used as part of the new APEL server. It can be used for sending any text-based data, and will be used for other types of accounting.

Work is in progress to improve SSM by simplifying its message sequence and defining its protocol precisely, allowing easier interoperability with other accounting systems in EGI.

Wider impact of this work

The SSM can be used as a transport layer for any message content, and is tested with the EGI broker network. Although it was designed for accounting, it could be used for other purposes.

Printable Summary

SSM is a program written in python which allows secure transfer of data over the STOMP protocol, which can be configured to send and receive data between arbitrary endpoints. It is used in production for publishing accounting data to the new APEL server. This presentation describes the current status of SSM and plans for its development.

Primary author: Mr ROGERS, Will (STFC)

Presenter: Mr ROGERS, Will (STFC)

Session Classification: Operations Workshops

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 71

Type: **Presentation**

A bioinformatics user point of view of cloud computing

Tuesday, September 18, 2012 2:40 PM (20 minutes)

Description of the work

I will list some recent activities that have been computationally demanding and their computing environments. In the presentation I will include some wider examples beyond our own work.

-The SUPERFAMILY database pipeline (supfam.org) runs in the AWS EC2 cloud environment. We use this to process the results of genome and other sequencing. The flexibility allows us to use a large number of instances to return results quickly whilst being idle most of the time. Our demands fluctuate wildly, favouring such a flexible funding model. Furthermore the rate at which sequence data is being produced is rising faster than Moore's law, so we see cloud computing as the most scalable solution for the future.

-To give context to genome sequencing we must build phylogenetic trees of nearly two thousand species. The search space of a binary tree of this size is effectively infinite, but by using heuristic maximum likelihood algorithms we are able to calculate a biologically reasonable topology. Investigations of these trees were carried out on the Compute Canada national resource, accessing the order of 10^7 CPU hours.

-Next generation sequencing technology presents a significant computational challenge in assembling the overlapping raw fragments of data that are produced, into contiguous sequence. We carried out the assembly of about 70 full human transcriptomes, each requiring a minimum of 100GB of RAM taking many days. We also attempted an assembly of the combined set using 1TB of RAM. This was conducted on the Institute of Cancer Research HPC facility in London.

-Without going into details, constructing protein coding genes out of assembled DNA or RNA sequence is another challenge facing bioinformatics in the face of the deluge of next generation sequence data. We have written prototype software in Hadoop, developed on a local cluster, for effectively carrying out protein analysis directly on assembled DNA/RNA effectively bypassing the need for gene prediction.

Link for further information

<http://www.cs.bris.ac.uk/~gough>

Wider impact of this work

The motivation for presenting at this forum is twofold: first to give attendees (by describing use cases) an insight into the computational requirements of bioinformatics, the ways in which it is currently being deployed in different computing environments and give some orientation of the field of bioinformatics and the way it moves; secondly to bring back from the meeting information to disseminate to the community of bioinformatics developers and users. As a member of three influential consortia in bioinformatics, various biotechnology/medical national funding panels and international conference committees, I will contribute to bringing these applied areas forward in their use of cloud computing.

The wider impact is thus in influencing both communities and accessing, via this forum a channel of two-way communication. Promotion of cloud computing in the life sciences and raising awareness of bioinformatics to the developers of software and infrastructure will aid coordinated growth.

Printable Summary

Our research group has been making use of HPC and distributed computing for bioinformatics since 1999. We committed the majority of our research to cloud computing at the start of 2010, and see this as our long-term future.

This presentation will describe the various bioinformatics use cases in our research group and some from the wider field of bioinformatics in general. This includes e.g.: DNA and protein sequence analysis, phylogenetics, next generation sequence assembly, personal medicine, etc. The presentation will also describe some of the computational resources we use for bioinformatics including e.g.: AWS EC2 cloud, Compute Canada national distributed resource, local Hadoop, and national HPC.

This presentation is aimed at those interested in hearing stories of use cases for cloud computing in bioinformatics, and potential directions in which the field could be moving in the future with respect to its computational requirements.

Primary author: GOUGH, Julian (University of Bristol)

Co-authors: Mr SMITHERS, Ben (University of Bristol); Dr VAVOULIS, Dimitris (University of Bristol); Mr OATES, Matt (University of Bristol)

Presenter: GOUGH, Julian (University of Bristol)

Session Classification: Providing cloud services

Track Classification: Virtualised Resources: challenges and opportunities (Michel Drescher: track leader)

Contribution ID: 72

Type: **Presentation**

GOCDB, GLUE2 and Information Discovery

Wednesday, September 19, 2012 4:00 PM (30 minutes)

Description of the work

see abstract

Wider impact of this work

see abstract

Printable Summary

The GOCDB data model is comparable to a sub-set of the GLUE2 specification. A current GOCDB requirement is to render the GOCDB data in a GLUE2 compatible format. To achieve this, we have engaged with the GLUE2 working group to help advance the GLUE2 XML rendering. We will provide a status update of this work. In addition, we will also assess the feasibility and requirement of the Regional-Publishing GOCDB model that would involve publishing of EGI scoped data to a parent GOCDB; its main use-case is to allow customisation of the Local scoped data. With the introduction of a GLUE2 rendering, we would like to explore the overlap (if any) between GOCDB and other information systems. We would also like to explore any avenues for interoperability between the different information discovery systems.

Primary authors: CASSON, John; MEREDITH, david (STFC)**Presenter:** MEREDITH, david (STFC)**Session Classification:** Operations Workshops**Track Classification:** EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 73

Type: **Presentation**

SOA3: an architecture for service oriented authentication, authorization and accounting in distributed environment

Wednesday, September 19, 2012 2:50 PM (20 minutes)

Description of the work

The user list is stored in an LDAP directory: every record contains userId, password and a set of generic attributes. Every user can be associated to one or more roles and one or more groups. SOA3 User Management Service provides REST CRUD operations for managing stored identities, roles or groups. SOA3 Authentication Service is composed by two modules:

username/password Module, which matches incoming usernames/passwords with the information stored on the LDAP directory

SAML Access Module which grants the accesses parsing SAML assertions: this module acts as a SAML Service Provider providing SAML based Identity Federation.

The Authentication Service also includes a SAML Identity Provider and Attribute Authority: when the access is granted, a SAML Authentication Assertion is returned containing a set of attributes associated to the user. The associations to group and roles are considered attributes in this context. This Assertion can be used both for Identity Federation and to propagate the attributes in the domain for authorization purposes.

The Authorization Service implements Attribute Based Access Control model with policies written in XACML. The service checks if a request must be permitted or denied basing on the attributes of the requester: the attributes are transported as SAML assertions associated to the request. The assertions can be directly associated, e.g. adding an assertion to the SOAP security header, or indirectly, adding only the reference to be resolved by SAML Artifact Resolution Profile.

The Accounting service is built around the Usage Tracker service whose goal is to keep track of resource usage by receiving and archiving usage records. It provides CRUD operations on usage records. Core accounting components are agnostic on the type of the resources and their properties by adopting an extensible data model. Specializations of this generic model have, however, been realized to ease the management of resource types actually in use.

Wider impact of this work

The work is a complete, standard based security architecture. It has been designed to be used in distributed environment. The possibility to introduce the security as an infrastructure service provides a great added value on modularity and maneageability.

A strong importance has been given to the performance aspect: for example LDAP directory has been chosen for this reason.

Printable Summary

Service Oriented Authentication, Authorization and Accounting (SOA3) provides the three main functionalities of a security system according to the Security as a Service model. All the services

are exposed by RESTful interfaces providing high adaptability to different context, especially distributed environments and Cloud environments.

The architecture is composed by:

- * an Authentication Service, including an User Management module and a SAML ID-Federation Module
- * a XACML based Authorization Module
- * an Accounting Module

Primary author: FORMISANO, Ciro (Engineering Ingegneria Informatica)

Co-authors: TRAVAGLINO, Ermanno (Engineering Ingegneria Informatica); BALRAJ, Kanchanna (Engineering Ingegneria Informatica Spa); Mr FABRIANI, Paolo (Engineering Ingegneria Informatica S.p.A.)

Presenter: FORMISANO, Ciro (Engineering Ingegneria Informatica)

Session Classification: AAI Workshop

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 74

Type: **Presentation**

Volunteer clouds to extend the resources of EMI middleware based VOs

Tuesday, September 18, 2012 2:20 PM (20 minutes)

Description of the work

The Generic BOINC Application Client (GBAC) is a virtualization based wrapper. It aims to be a generic framework providing virtualized environments for various distributed computing infrastructures (DCIs). GBAC is implemented using the DC-API Meta API and does not rely on any middleware specific functionalities, thus it is possible to use it on any DCIs that are supported by DC-API. In the talk we refer to the BOINC version of GBAC for demonstrating its concepts and internals.

Volunteer cloud is the term used for a cloud type infrastructure based on volunteer resources. Thanks to the combination of BOINC and virtualization, now it is possible to launch virtual machines on volunteer resources. BOINC primarily supports VirtualBox [9], which means any client machine having preinstalled VirtualBox can participate in executing virtual machines.

In this volatile and unreliable environment deploying an Infrastructure-as-a-Service cloud infrastructure faces several challenges. The first challenge is motivating donors not just to donate their CPU and GPU resources but to provide more access to their computers.

Second, volatility and availability problem of volunteer resources is another major challenge. Since any volunteer host may be shut down any time fault tolerance of appliances becomes a key challenge. In volunteer computing this is mitigated in two ways. First, redundancy is used, meaning more than one resource handles the same task and if one fails the other(s) are still able to produce the result. Second, applications periodically checkpoint themselves, in case of a restart the application can continue from the last checkpoint. Overall, checkpointing and redundancy has to be carefully designed and optimized for network bandwidth. Moreover, forecasting the availability of the volunteer hosts is also a key factor to increase the efficiency of the service migration.

Both the basic technical solutions as well as the challenges will be explained in the talk.

Link for further information

<http://www.lpds.sztaki.hu/products>

<http://edgi-project.eu/>

Wider impact of this work

The GBAC approach enables run any parameter sweep application on a BOINC project that registers the GBAC application. In this way volunteer BOINC systems can be used for a much wider set of applications than before without any application porting effort. This is particularly important in two scenarios.

The first scenario comes from the EDGI project where SG VOs can be extended with volunteer and local DG systems. If these DG systems register the GBAC application then any PS application running in the SG VO can be automatically transferred to and executed in the connected DG systems.

No application porting effort is required. This opens new horizon for extending SG systems with large set of volunteer and local DG resources.

In the second scenario universities would like to set up volunteer DG systems for their researchers and students. In order to ensure the flexibility of these university DG systems the GBAC approach can be applied as we do in a Hungarian university DG program.

Printable Summary

EMI middleware (ARC, gLite, UNICORE) based VOs can be extended with cheap and already available resources as volunteer clouds. The new GBAC (Generic BOINC Application Client) technology developed by MTA SZTAKI in the framework of the EU FP7 EDGI project enables that volunteer desktop machines can appear as a volunteer cloud resources for EMI-based VOs. The talk will explain the technical solution how to turn desktop machines into virtualized resources and use them as clouds. The talk will show how large-scale parameter sweep simulations can exploit this technology in a transparent way.

This solution shows many similarities with SaaS clouds in the sense that the SG systems are extended with new resources on demand. The main difference is that these additional resources are collected from a volunteer system and hence the users do not have to pay for the use of these resources.

Primary author: Prof. KACSUK, Peter (MTA SZTAKI)

Co-author: KOVACS, Jozsef (MTA SZTAKI)

Presenter: Prof. KACSUK, Peter (MTA SZTAKI)

Session Classification: Providing cloud services

Track Classification: Virtualised Resources: challenges and opportunities (Michel Drescher: track leader)

Contribution ID: 75

Type: **Workshop**

Towards an integrated information system - third workshop

Thursday, September 20, 2012 11:00 AM (1h 30m)

Description of the work

The main topics of the workshop will be:

- Status of the implementation of GLUE2 in the various services of the different middleware stacks.
- The integration between the middleware service discovery infrastructures and the operational tools.
- Usability and performance of the current information discovery services and new developments.
- Feedback from user communities about the information system usage and related requirements.

Wider impact of this work

As the previous workshops on this topic, the output of this session will be a detailed snapshot of the current status of the information system related developments for the different projects, the actions and the planned activities for the next months.

Since Information system is a widely used service, its evolution will influence the operations activities as well as the user community workflows.

Printable Summary

This is the third joint EGI/EMI/IGE workshop that focuses on the need of an EGI integrated infrastructure for middleware services discovery across heterogeneous grid infrastructures - based on ARC, gLite, GLOBUS, UNICORE –and virtualized resources.

The workshop will assess the status of the activities proposed during the previous workshops, which include developments in middleware services and operations tools, as well as high level policies and standards definition and application.

The workshop will define the actions to continue the work towards the implementation of a common strategy among the projects, and it is a good opportunity to discuss with the developers and resource providers use cases and requirements.

Primary author: BURKE, Stephen (EGLEU)

Presenters: KONYA, Balazs (EMI project); CESINI, Daniele (INFN); SAVERCHENKO, Ilya (BADW); BURKE, Stephen (EGLEU)

Session Classification: Operations

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 76

Type: **Presentation**

Managing Virtual Research Environments in Hybrid Data Infrastructures

Tuesday, September 18, 2012 11:22 AM (22 minutes)

Description of the work

Science is increasingly global, multidisciplinary and networked. It needs access to a large amount of datasets that come in all forms and shapes from huge international experiments to cross-laboratory, single laboratory, or even from a multitude of individual observations. The exploitation of datasets originally maintained by several organizations usually spread worldwide represents the new challenging requirement. It is not even possible to think to a future where few selected standards, best practices, and policies will be widely adopted by the science ecosystem. Heterogeneity will likely continue to exist even in the presence of an increasing multidisciplinary approach to science.

Hybrid Data Infrastructures (HDI) born to deal with such heterogeneity. They integrate several technologies for data management, access and analysis while providing transparent access to heterogeneous computational and storage platforms. Moreover, HDI preserving management capabilities such as monitoring, accounting, and secure access become a credible approach in this new challenging scenario.

The gCube software system implements the HDI approach. It offers a data-management-capability-delivery model in which computing, storage, data and software are made accessible by the infrastructure and are exploited by users using a thin client (namely a web browser), through dedicated on-demand Virtual Research Environments (VRE).

VREs are declaratively and dynamically build while abstracting on the location, provenance and interfaces of the resources. gCube technology implements a user friendly Software as a Service framework where the data, the application services, and the storage and computing resources needed by a scientist are automatically aggregated and made available through a web based interface. The aggregated resources are also monitored to guarantee the VRE service while guaranteeing secure and controlled access.

Link for further information

D4Science : www.d4science.org

gCube web site: www.gcube-system.org

gCube documentation: gcube.wiki.gcube-system.org/gcube

Wider impact of this work

In this emerging new science, the HDI maintainer aggregates computational and storage resources from a variety of providers including commercial Cloud providers, while resource consumers buy them only for the time needed for their exploitation and use them to build their virtual environments.

Thanks to D4Science HDI and the gCube technology for VRE creation, for example a newly Biodiversity VRE:

- aggregates datasets from the Ocean Biogeographic Information System, the Global Biodiversity Information Facility, the Ocean Monitoring and Forecast initiative, the NCBI, the World Register of Marine Species, and others
- offers new and unprecedented computing capabilities obtained accessing transparently to resources provided by D4Science contributors and by Windows Azure Cloud computing platform
- integrates technologies for managing ecological niche modelling, time series harmonization, statistical data analysis with R, and data mining with WEKA.

Printable Summary

The global and networked needs of the science produce and need to exploit huge quantities and varieties of data. To confirm this impression, a recent study, promoted by The Royal Society of London, highlighted how science is becoming increasingly global, multidisciplinary and networked. Singleton technological platforms are no longer able to address the data and processing requirements of the emerging data-intensive science characterized by a predominant data distribution and by evolving user communities. A novel approach, the Hybrid Data Infrastructure (HDI), integrates several technologies, including Grid and Cloud, and offers the necessary management capabilities required by geographically disperse user communities and data centers.

If HDI provides the data processing and analysis computing capabilities, Virtual Research Environments allow to manage the elastic and secure aggregation of users in focused communities living for the time needed to achieve their objective.

Primary authors: MANZI, Andrea (CERN); Dr CANDELA, Leonardo (CNR - ISTI); PAGANO, Pasquale (CNR - ISTI)

Presenter: PAGANO, Pasquale (CNR - ISTI)

Session Classification: Research Infrastructures

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 77

Type: **Demonstration**

WNoDeS: how virtualization can get both users and resource center administrators satisfied

Description of the work

WNoDeS integrates Grid and Cloud provisioning through virtualization. Resource brokering is done through a tight integration with an LRMS, allowing flexible policies to access resources; since these are allocated through an LRMS regardless of the requested interface, sites can enrich their offerings with IaaS provisioning exploiting current assets, know-how and tools. Also, static partitioning can be avoided, allowing optimization of resource utilization.

Grid jobs can be run on VMs using a GLUE Schema attribute to specify virtual image name and characteristics such as RAM or cores. For VM instantiation an API (being made compliant to the OCCI 1.1 specification) and a CLI are provided.

A WNoDeS feature called mixed mode allows using physical resources as both traditional batch nodes and, at the same time, as HVs for VM instantiation. This lets sites to introduce features such as VM and cloud computing support on traditional resources without disrupting existing services and allows to efficiently decide which workloads are to be virtualized and which should be run instead on non-virtual hardware.

The demo will show how WNoDeS was used to support some real use cases:

- how a virtual image is managed, how a user can ask that jobs be executed on a VM running a given image and how mixed mode supports on the same box jobs requiring a virtual environment and jobs that don't. This will be applied to a case describing the work to support the Auger experiment through an adaptation of their exp. framework involving DBs accessed from and encapsulated in VMs.

- how WNoDeS can run Grid jobs on custom VMs for the WeNMR project. Their CING (Common Interface for NMR structure Generation) sw has many external dependencies making it suitable for a virtual set up. The demo will show how the VCING machine is instantiated and how a selection mechanism allows WeNMR users to run jobs on the VM.
- how it is possible to directly instantiate VMs using the OCCI API and the WNoDeS CLI.

Link for further information

<http://web.infn.it/wnodes>

<https://wiki.italiangrid.it/twiki/bin/view/Cloud/WebHome>

Wider impact of this work

WNoDeS has been in production use at INFN Tier-1 site, at CNAF, Bologna, for over three years and has been installed at several Italian centers. Scalability of the WNoDeS LRMS-based brokering system has been demonstrated for several thousands of concurrent virtual machines running millions of production jobs. WNoDeS 2, the latest version, has been released as part of the EMI-2 distribution and further developments are already planned to improve the product. In particular, we expect important new features related to flexible support of dynamic virtual networks and

to the integration of third party solutions to implement intercloud operations to be introduced. WNoDeS characteristics like OCCI support, mixed mode and integration with local and Grid computing, demonstrated through real use cases in this session, will allow a seamless introduction of virtual and Cloud computing services into existing centers, extending the customer base and optimizing resource exploitation.

Printable Summary

The demo will show how the WNoDeS virtualization framework was applied to real use cases. In particular, how WNoDeS can be used to introduce Cloud-based services and to support custom VMs for local and Grid users alike. The demo is based on actual production usage at the INFN Tier-1 CC at CNAF, Bologna and describes how WNoDeS has fulfilled the needs of 2 important user communities:

- the Auger astro-particle physics experiment;
- a use case part of the WeNMR e-infrastructure.

It will also show how direct instantiation of VMs can be accomplished through an OCCI-compliant CLI.

Finally, it will show how the WNoDeS mixed mode feature can be used to efficiently decide which workloads are to be virtualized and which instead should be run on non-virtual hardware. This lets the centers to progressively introduce virtualization technologies and related services without disrupting existing customers and, at the same time, achieving maximum exploitation of the available computing resources.

Primary authors: ANDREOTTI, Daniele (INFN CNAF); DALLA TORRE, Gianni (INFN CNAF)

Co-authors: ITALIANO, Alessandro (INFN CNAF); CESINI, Daniele (INFN); SALOMONI, Davide (INFN); Dr DONVITO, Giacinto (INFN); GAIDO, Luciano (INFN); Dr VERLATO, Marco (INFN); MAZZUCATO, Mirco (INFN); VENTURI, Valerio (INFN CNAF)

Presenters: ANDREOTTI, Daniele (INFN CNAF); DALLA TORRE, Gianni (INFN CNAF)

Track Classification: Virtualised Resources: challenges and opportunities (Michel Drescher: track leader)

Contribution ID: 78

Type: **Presentation**

ScienceSoft - Open Software for Open Science - Call For Participation to Our Community

Tuesday, September 18, 2012 4:30 PM (10 minutes)

Description of the work

TBD

Wider impact of this work

TBD

Printable Summary

ScienceSoft - Open Software for Open Science is a new initiative promoted by EMI in collaboration with EGI, StratusLab, iMarine, OpenAIRE and a number of other projects and SMEs. The goal is to explore the feasibility and advantages of creating an open source community for software specific to scientific communities. Also, collecting community requirements and propose realistic solutions while

making the activities of producing and using open source software for science more transparent and collaborative across communities and projects. Another part is to implement a sustainable business model based on existing successful examples (Apache, Eclipse, Drupal, SourceForge, etc). ScienceSoft is currently still in a preparation phase and open for input from new members and actors. This contribution will inform about ScienceSoft including a strong call for participation to members of our community.

Primary author: Mr RIEDEL, MORRIS (JUELICH SUPERCOMPUTING CENTRE)

Co-authors: DI MEGLIO, Alberto (EMI); STEFAN, Peter (NIIFI)

Presenters: DI MEGLIO, Alberto (EMI); Mr RIEDEL, MORRIS (JUELICH SUPERCOMPUTING CENTRE)

Session Classification: Software services for community building and support

Track Classification: Community and Co-ordination (Sergio Andreozzi: track leader)

Contribution ID: 79

Type: **Presentation**

EMI and Open Standards: A Brief History of the Future

Tuesday, September 18, 2012 12:00 PM (15 minutes)

Description of the work

TBD

Wider impact of this work

TBD

Printable Summary

EMI is one of the drivers of the Open Grid Forum (OGF) by providing numerous inputs into ongoing specification work and by chairing actively several working groups. While several concrete specification proposals have been given into OGF (e.g. Compute Accounting Record, Storage Accounting Record, or EMI Execution Services), the standardization of these specifications is work-in-progress. Also work on OASIS SAML and XACML profiles will be shown in this presentation. The relationships of the EMI activities to OGF activities and future outcomes will be presented.

Primary author: Mr RIEDEL, MORRIS (JUELICH SUPERCOMPUTING CENTRE)

Co-authors: KERR NILSEN, Jon (UIO); FUHRMANN, Patrick (DESY); MEMON, Shahbaz (JUELICH); SHIRAZ MEMON, ahmed (JUELICH)

Presenter: Mr RIEDEL, MORRIS (JUELICH SUPERCOMPUTING CENTRE)

Session Classification: New middleware products

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 80

Type: **Workshop**

Resource Centre Forum

Wednesday, September 19, 2012 11:00 AM (5 minutes)

Description of the work

This is the III meeting of the Resource Centre Forum.

Wider impact of this work

Fostering of collaboration between resource providers and resource administrators, improve operational support to multidisciplinary science.

Printable Summary

The Resource Centre Forum puts together experts and representatives from the EGI Resource Centres to discuss operational and technical issues that concern various topics such as deployment of licensed software, resource allocation to new user groups, deployment of GPGUS, deployment of batch systems such as Grid Engine, etc.

During the Resource Centre Forum we will discuss the creation of a permanent group of experts to address site-specific VO operational issues.

The ultimate goal of the Resource Centre Forum is to improve support to multiple disciplines, foster collaboration between Resource Centres from different NGIs and to allow sharing of expertise between them.

The III Resource Centre Forum meeting is aimed at VO operations managers, to site representatives and NGI operations teams.

Primary author: Dr FERRARI, Tiziana (EGLEU)

Presenter: Dr FERRARI, Tiziana (EGLEU)

Session Classification: Operations - Resource Centre Forum

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 81

Type: **Presentation**

EMI Product Use Cases - Supporting Science in Distributed Systems

Tuesday, September 18, 2012 2:00 PM (22 minutes)

Description of the work

TBD

Wider impact of this work

TBD

Printable Summary

The EMI project is a collaboration among ARC, dCache, gLite, and UNICORE and all these middle-ware technologies serve the needs for scientific users with distributed systems since many years. EMI continues and documents these use cases under the umbrella of EMI product use cases clearly illustrating where and how EMI products are used across a wide variety of scenarios. The EMI products use cases are the key to understand and communicate the current usage of EMI products in practice and attract interest by new user communities that have similar use cases (e.g. such as those arising from ESFRI RIs with computationals and storage demands). This presentation will provide an overview of the use case scenarios specifically in the context of user communities and so called Virtual Research Environments (VREs) and give a glimpse on the answer to 'quo vadis EMI?'

Primary author: Mr RIEDEL, MORRIS (JUELICH SUPERCOMPUTING CENTRE)

Co-authors: Dr RYBICKI, Jędrzej (Juelich Supercomputing Centre); STEFAN, Peter (NIIFI); BALA, Piotr (UWAR); MEMON, Shahbaz (JUELICH); SHIRAZ MEMON, ahmed (JUELICH)

Presenter: Mr RIEDEL, MORRIS (JUELICH SUPERCOMPUTING CENTRE)

Session Classification: Research Infrastructures

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 82

Type: **Presentation**

WS-PGRADE/gUSE generic DCI gateway framework for EGI user communities

Thursday, September 20, 2012 11:30 AM (20 minutes)

Description of the work

In this talk we introduce WS-PGRADE/gUSE, a generic DCI gateway framework and backend service stack that can be easily customized to build application-specific science gateways. The most distinguishing feature of WS-PGRADE/gUSE compared to other generic DCI gateway frameworks is that it is workflow-centric, i.e., it provides services to build workflow applications that can be executed on various DCIs. The provided execution mechanism enables the simultaneous execution of workflow nodes placed on parallel workflow branches. These nodes can be executed in parallel in different DCIs enabling the organization of very efficient parallel, multi-DCI workflow applications. WS-PGRADE/gUSE provides all the services that are needed to create, execute and monitor these workflows. Therefore the three most important features of WS-PGRADE/gUSE are as follows:

1. Workflow support
2. Enabling multi-DCI workflow execution
3. Enabling the customization of the framework towards application-specific science gateways

In the current talk we focus on these most important features of WS-PGRADE/gUSE, i.e., we show how it allows creating complex workflow scenarios, and enables running them on a diverse set of DCIs. We show two different customization methodologies of WS-PGRADE/gUSE that will be used within the EU FP7 SCI-BUS project to create customized science gateways.

As an example to create customized science gateways a new Autodock gateway based on WS-PGRADE/gUSE has been opened for biologist to access the EDGeS@home volunteer desktop grid resources. This is a good example for the communities to create specific gateways by customizing WS-PGRADE/gUSE.

Link for further information

<http://www.guse.hu/>

<https://sourceforge.net/projects/guse/>

Wider impact of this work

WS-PGRADE/gUSE is an open source software based on apache license and can be downloaded from sourceforge. Since its publication on sourceforge in February 2011 more than 2000 downloads have been done from 42 countries.

The WS-PGRADE/gUSE framework has already been used by several user communities to customize for their need because its very flexible workflow system and its capability of submitting and managing these workflows on a large variety of different DCIs.

It is the core gateway framework of the SCI-BUS EU FP7 project that will customize from gUSE 17

different application-specific gateways for various user communities including seismology, astrophysics, helio-physics, chemistry, biology, medical science, etc.

WS-PGRADE/gUSE is also the basis of the SHIWA Simulation Platform that enables combining many different kind of workflows (Taverna, Askalon, Moteur, GWES, Galaxy, WS-PGRADE, Kepler, Triana, etc.) into a single meta-workflow and execute them on various DCIs.

Printable Summary

The WS-PGRADE/gUSE generic DCI gateway framework has been developed to support large variety of user communities. It provides a generic purpose, workflow-oriented graphical user interface to create and run workflows on various DCIs including clusters, grids, desktop grids and clouds. The framework can be used by NGIs to support various small user communities who cannot afford to develop their own customized science gateway. The WS-PGRADE/gUSE framework also provides two API interfaces (Application Specific Module API and Remote API) to create application-specific science gateways according to the needs of different user communities. The paper describes in detail the workflow concept of WS-PGRADE, the DCI Bridge service that enables access to most of the popular European DCIs and the Application Specific Module and Remote API concepts to generate application-specific science gateways.

Primary author: KACSUK, Peter (MTA SZTAKI)

Co-author: FARKAS, Zoltan (MTA SZTAKI)

Presenter: FARKAS, Zoltan (MTA SZTAKI)

Session Classification: Science Gateways

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 83

Type: **Presentation**

APEL as a Global Accounting Repository

Wednesday, September 19, 2012 11:40 AM (20 minutes)

Description of the work

APEL has been collecting CPU accounting data for EGI, EGEE, and the Worldwide LHC Computing Grid since 2004. As a precursor to developing a distributed server architecture for APEL a major redesign started in 2011. One element of this was a new transport layer SSM (described in another abstract). SSM uses the production EGI Messaging Infrastructure rather than a dedicated broker. The migration of OSG, DGAS, SGAS, and few site-specific accounting systems is under way and this led to the engagement with a variety of other accounting systems, including, Globus, ARC, MAPPER, and EDGI.

The next stages of evolution will be for the APEL client to use this infrastructure and thus migrate the bulk of EGI sites. Then a NGI version of the APEL repository can be deployed for NGIs to manage their own job-level data and send summaries on to the global repository.

Wider impact of this work

This work has widened the accounting coverage of the EGI ecosystem to well beyond EGI itself. International VOs can accumulate accounting data from a variety of middleware solutions and thus a wider coverage of sites and e-infrastructures.

Printable Summary

The APEL Central Accounting Repository has gone through a major redevelopment in the last year. The ability for non-APEL systems to publish accounting records over the EGI Messaging Bus was developed for the infrastructures running their own accounting systems and a few sites. As a result of this development a number of other infrastructures have planned to send their accounting data to APEL, thus widening the accounting coverage of the global e-infrastructure. The new transport layer SSM has also been used to transport new types of accounting usage records for prototypes and tests. This has extended accounting coverage in scope as well as geographical range.

This presentation describes the progress made in the past year by the EGI Accounting Infrastructure, lists the various systems now publishing and the roadmap for others.

Primary author: Ms PACKER, Alison (STFC)

Presenter: Ms PACKER, Alison (STFC)

Session Classification: Operations Workshops

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 84

Type: **Presentation**

Experience in community building and user championing

Tuesday, September 18, 2012 2:40 PM (20 minutes)

Description of the work

Community building and user championing are a major part of the business of EGI and this is reflected in the range of interconnected activities that support this mission. EGI obviously depends upon users for its continued success and growth. To this end there is a distinction between the role of support for existing users which is now part of the operational services whereas out reach to new and emerging communities is undertaken by a twin approach of Community Outreach and more specifically, Technical Outreach to New Communities (TONC). Community Outreach is focussed on identifying these new, emerging and potential communities and reaching out to them in the form of participation at their community events and also in inviting and encouraging them or at least their significant representatives to attend EGI's Forums. Attendance at the Forums enables opportunities to occur the communities to interact with the developers, providers and other communities that make up the EGI community.

Once these initial introductions have been made, specific technical interactions can be established and developed in the form of Virtual Team (VT) projects with self-selecting NGI representatives and partner projects, specialist Task Forces and even new fully fledged funded projects. NGI International Liaisons (NILs) have been established at each of the NGIs to handle such non-operational initiatives and develop the reach out activities with the user communities. The work of the NILs will be presented and some examples given of their role and successes to date. To extend these activities another role is being developed, that of EGI Champion. These will be identified from amongst the user communities and will typically be associated with a particular institution. As experts in their field they will be ideally placed to champion the benefits of using the infrastructure to a new cluster of users across Europe and beyond.

Link for further information

<http://www.egi.eu>

Wider impact of this work

The main impact of this work is to grow the usage of the infrastructure in a variety of ways. The breadth of user communities has been broadened with the addition of new VRCs both officially recognised through MoUs and others who are developing structures around which to build their community. VT projects have been responsible for implementing in discrete phases a range of new technological solutions for users and partners. The non-operational NILs have already in a relative short space of time, proved the importance of co-ordination in this area. The UCB meetings continue to be a valuable forum for those communities that wish to feed new requirements into the evolution process and also who wish to discuss aspects of the community support process in a public arena. The benefits all of these shared initiatives are that such discussions can acquire further momentum as a result of the sharing of needs and solutions.

Printable Summary

The second of three sessions on the theme of Community and Co-ordination, this session focuses on experience in community building and user championing and covers VRCs, NGIs, ESFRI projects and other initiatives. Within EGI-INSPIRE this activity primarily falls under the Community Engagement work package and in particular the tasks Community Outreach and Technical Outreach to New Communities (TONC). In addition to these activities the function and operation of the User Community Board (UCB) will also be presented. Many of the user community representatives within the UCB are from recognised Virtual Research Communities but not all. In addition to the structure of the Forums, Technical and Community, two new entities have been created during the second year of EGI-INSPIRE - namely NILs and VTs. NIL stands for NGI International Liaisons and VT for Virtual Teams. NILs are responsible for non-operational matters and may participate in Virtual Teams (VTs) to achieve short time-frame goals.

Primary author: BREWER, Stephen (EGLEU)

Presenter: BREWER, Stephen (EGLEU)

Session Classification: Experience in community building

Track Classification: Community and Co-ordination (Sergio Andreozzi: track leader)

Contribution ID: 86

Type: **Presentation**

OpenNaaS, a toolkit for IP Networks as a Service

Tuesday, September 18, 2012 3:00 PM (20 minutes)

Description of the work

OpenNaaS was born with the aim to create an open source software project community that allows several stakeholders to contribute and benefit from a common NaaS software stack. OpenNaaS offers a versatile toolset for the deployment of NaaS oriented services. With an eye on versatility and smooth integration, OpenNaaS offers a powerful remote command line, as well as web-service interfaces. This web-service interface will offer the possibility to both build a GUI and integrate it with existing middleware applications already deployed in the virtual research organisations.

Initially four research communities will benefit from OpenNaaS powered services: the Danish Health Data Network, the High Performance Networks Group from the University of Essex, the Irish NGL, and NORDUnet through their cloud services.

The Grid-Ireland Operations Centre (<http://grid.ie>) in Trinity College Dublin runs the Irish NGL. Grid and cloud computing use-cases are being explored by the Grid-Ireland Operations Centre. As Grid-Ireland has resources distributed across Ireland and connected via HEAnet it is an ideal test bed for OpenNaaS.

The Grid-Ireland research explores novel approaches to sharing and connecting computing and storage resources. This includes creating a distributed computing cloud between institutions with resources federated at the network level, with an intended use as a platform for agent-based computing. This raises issues relating to adding and removing sites amongst others. A similar approach is taken for integrating computing resources into grid batch systems, transferrable to use of academic clouds for grid computing.

These use cases will be demonstrated at the EGI Technical Forum 2012.

Link for further information

<http://www.mantychore.eu/> <http://www.opennaas.org/>

Wider impact of this work

Mantychore FP7 will carry out pre-operational deployments of the IP network service at UNI-C, NORDUnet and HEAnet. Part of the project effort will be dedicated to consolidate and enhance the community of providers (NRENs but also commercial) and users of the IP network service.

In general, by working with research communities in higher-education and research institutions Mantychore will gather real-world requirements for OpenNaaS connecting dynamic and user-controlled networks to traditionally managed networks.

Printable Summary

Dynamic and multi-domain provisioning of network resources has been a long-standing research area, and the rise of grid and cloud computing brings even more challenges. In order for operators

to be able to deploy innovative Networks as a Service offerings, the Mantychore FP7 project has created the OpenNaaS framework. Grid and cloud computing provide some of the first use cases for OpenNaaS, making it possible to federate –at the network level –computing and storage resources between grid resource centres or to create a distributed computing cloud spanning several organisations.

This 20-30 minute presentation will give infrastructure operators and user communities with unique network requirements an overview of OpenNaaS capabilities and the development roadmap and will demonstrate relevant use cases, all with a view to gathering additional requirements and interest in deploying the software.

Primary author: O'CALLAGHAN, David (TCD)

Co-author: MINOVES, Pau (TCD)

Presenters: O'CALLAGHAN, David (TCD); MINOVES, Pau (TCD)

Session Classification: Providing cloud services

Track Classification: Virtualised Resources: challenges and opportunities (Michel Drescher: track leader)

Contribution ID: 88

Type: **Workshop**

Introduction to the EGI/EUDAT/PRACE workshop

Tuesday, September 18, 2012 2:00 PM (10 minutes)

Wider impact of this work

Fostering of collaboration between different resource infrastructures, improve operational support to multidisciplinary science.

Description of the work

This workshop aims at consolidating the collaboration between EGI-InSPIRE, EUDAT and PRACE.

Printable Summary

EGI, EUDAT and PRACE are pan-European infrastructures offering differentiated types of resources whose integrated offering aims at facilitating e-Science in Europe. Seamless provisioning of services across EGI, EUDAT and PRACE requires some level of integration in different areas, such as security policies and procedures, support structures, accounting, monitoring and data management.

During the workshop different integration strategies will be explored and cross-infrastructure data management use cases will be presented.

During the EGI-EUDAT-PRACE workshop we will

1. present the current status and next plans for the integration of the helpdesk systems;
2. present the current status and next steps for the integration of the EGI and PRACE accounting systems to allow the collection of accounting information from multiple infrastructures;
3. investigate use cases for data management, transfer across different infrastructures and usage of Globus Online: we will review existing requirements and identify emerging ones.

The workshop is aimed at user communities requiring the coupled usage of HTC and HPC resources, operations tool developers and the EGI, EUDAT and PRACE operations communities.

Primary author: Dr FERRARI, Tiziana (EGLEU)

Presenter: Dr FERRARI, Tiziana (EGLEU)

Session Classification: Operations -EGI/EUDAT/PRACE workshop

Contribution ID: 89

Type: **Workshop**

EGI/EUDAT workshop

Description of the work

This workshop aims at consolidating the collaboration between EUDAT and EGI-InSPIRE

Wider impact of this work

Fostering of collaboration between different resource infrastructures, improve operational support to multidisciplinary science.

Printable Summary

EUDAT and EGI are pan-European infrastructures offering differentiated types of resources whose integrated offering aims at facilitating e-Science in Europe. Seamless provisioning of services across EUDAT and EGI requires some level of integration in different areas, such as security policies and procedures, support structures, accounting and monitoring. During the workshop different integration strategies will be explored for the definition of a common roadmap. The workshop is aimed at resource infrastructure representatives, the operations community and operational tool developers.

Primary author: Dr FERRARI, Tiziana (EGLEU)

Presenter: Dr FERRARI, Tiziana (EGLEU)

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 90

Type: **Workshop**

Future of EGI Operations - workshop

Thursday, September 20, 2012 2:00 PM (5 minutes)

Description of the work

This workshop aims at involving Operations experts, Resource Infrastructure representatives and Resource Centre representatives in order to start defining an EGI operations roadmap for the next project years.

Wider impact of this work

Evolution of the operations architecture, contribution of operations to the implementation of the EGI strategic plan.

Printable Summary

EGI Operations rely on a set of tools, policies, procedures and services that allow seamless operations of heterogeneous infrastructures, which differ considerably in complexity, software deployed and supported, user communities supported etc.

In this workshop we will address the problems of the current operations model and we will discuss ideas on how to advance its architecture to make it more sustainable and suitable to Resource Infrastructures with different needs. In addition, we will discuss:

1. the operations sustainability of NGI and EGI operational services after EGI-InSPIRE
2. areas of innovations that have a good potential of being funded through Horizon 2020
3. the impact of middleware sustainability plans on EGI

The session seeks participation of operations experts, Resource Infrastructure representatives and Resource Centre representatives who will thus have a chance to contribute to the definition of the operations roadmap for the next project years.

Primary author: Dr FERRARI, Tiziana (EGLEU)

Presenter: Dr FERRARI, Tiziana (EGLEU)

Session Classification: Operations

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 91

Type: **Presentation**

Demonstration of the Federated Clouds Task Force test bed and services

Tuesday, September 18, 2012 4:00 PM (1h 30m)

Description of the work

This will be a 30 min Demonstration and Presentation.

Wider impact of this work

Research team leaders and representatives of research collaborations will be able to evaluate the advantages of a federated public sector Cloud infrastructure over a full-time full-fee commercial Cloud integration.

Printable Summary

The Federated Clouds Task Force will demonstrate the most recent proceedings in setting up a federated Clouds infrastructure in EGI, so that team leaders and representatives of research collaborations can evaluate the advantages of a federated public sector Cloud infrastructure over a full-time full-fee commercial Cloud integration. The demonstration will include a presentation about the functional components and services that comprise the federated Clouds infrastructure and the areas that will be technically demonstrated.

Primary author: Dr TURILLI, Matteo (Oxford e-Research Centre)

Presenter: Dr TURILLI, Matteo (Oxford e-Research Centre)

Session Classification: Providing cloud services

Track Classification: Virtualised Resources: challenges and opportunities (Michel Drescher: track leader)

Contribution ID: 92

Type: **Workshop**

Federated Clouds workshop

Thursday, September 20, 2012 2:30 PM (1 hour)

Description of the work

The session will be a 1 hour long discussion workshop.

Wider impact of this work

Enables communication between all players in the federation of clouds that EGI should support and/or work with.

Printable Summary

After the demonstration of the current state of the Federated Clouds technology and testbed, an open workshop will be held to reflect the current status with any interested party, and to elicit requirements and gaps to satisfy stakeholder use cases and goals for engaging with a federated Clouds infrastructure. Resource Providers will be able to evaluate which further actions need to be taken before they can join the Cloud infrastructure federation, both on the federation's side and the provider's side. Commercial cloud providers may learn about which services and technical interfaces they need to support to allow seamless Cloud bursting from a federated public sector infrastructure to commercial providers and vice versa. This workshop is intended for resource providers, their stake holders in the scientific research community and commercial cloud providers.

Primary authors: Dr TURILLI, Matteo (Oxford e-Research Centre); BREWER, Stephen (EGIEU)

Presenters: Dr TURILLI, Matteo (Oxford e-Research Centre); BREWER, Stephen (EGIEU)

Session Classification: Virtualised Resources

Track Classification: Virtualised Resources: challenges and opportunities (Michel Drescher: track leader)

Contribution ID: 94

Type: **Internal Project Meeting**

EGI Security Policy Group meeting

Friday, September 21, 2012 1:30 PM (1h 30m)

Description of the work

A closed meeting of SPG to continue its ongoing work of developing and maintaining security policies for use in EGI. The detailed agenda will be available nearer the time on the SPG wiki.

Wider impact of this work

Security policies are essential for the operation of the EGI infrastructure and for interoperation with others.

Printable Summary

A closed meeting of the EGI Security Policy Group (SPG) to continue its ongoing work.

See more detailed SPG agenda at:

<https://indico.egi.eu/indico/conferenceDisplay.py?confId=1161>

Primary author: KELSEY, David (STFC)

Presenter: KELSEY, David (STFC)

Session Classification: EGI InSPIRE

Track Classification: Community and Co-ordination (Sergio Andreatto: track leader)

Contribution ID: 95

Type: **Presentation**

Deploying User Oriented Services in IBERGRID

Thursday, September 20, 2012 4:40 PM (20 minutes)

Description of the work

During the last year, IBERGRID has put in place a set of strategies and tools to serve the user needs, based on two main branches: “enhancement of user activities” and “enhancement of user support”. The “enhancement of user activities” focus in providing the necessary information to the user so that he is empowered to take decisions. In this sense, sites have been classified in terms of the computing and storage capabilities they offer. This will allow users to know what to expect at each site, and to redirect their tasks to sites which, by configuration, offer more dedicated resources. On the other hand, a pilot infrastructure based on DIRAC is currently being tested which will allow to decrease the barriers of current and future users in using the infrastructure.

The “enhancement of the user support” focus on providing the proper structure to the current support activities. User support shifts are focused on serving users sending their questions via a recently deployed RT regional helpdesk. The user support teams have at their disposal a wide set of VO monitoring tools, like a VO SAM connected to a VO Operational Dashboard, and a VO Admin Dashboard that works as an aggregator of relevant VO information. Such tools allow shifters to take the decision if the issue is an operation problems or a misuse of the user. The RT regional helpdesk is connected to GGUS which provides the functionality to redirect user complains to sites, if at the end the problems can only be solved by the sites operational team. These are just two concrete examples of the efforts in place but many other implemented and on-going processes will be presented and discussed during the talk.

Wider impact of this work

The IBERIAN strategies in place have as ultimate goal the fostering of the regional infrastructure usage. A very important effort has been made to boost the research activities in the region, but the most important challenge is to provide the necessary conditions so that researchers maintain a sustainable and continue activity. In this sense, we will also present, share and discuss our plans on how preserve the researchers we currently gained during the last year, being able to correspond to their expectations and challenges imposed to the user support and infrastructure managers almost on a daily basis.

Printable Summary

In moving towards a sustainable infrastructure, IBERGRID seeks increased users both by local communities and by general international users. Such goals can only be achieved by improving the satisfaction of users by providing them mechanisms and tools they need ofr their research activities. Simultaneously, such mechanisms and tools have also to be provided with a scope to make them easy for adoption by both current and new user communities, with minimum interference on their research activities. This talk presents the efforts developed by the Portuguese and Spanish user support teams during the last year, awarded with important achievements such as a huge

increase of the regional communities activities, where one of the Iberian Virtual Organizations reached EGI top-10 since January 2012, and by a very sustainable participation of local Iberian researchers in major international research projects, not only in High Energy Physics but also in other multidisciplinary domains.

Primary author: BORGES, Goncalo (LIP)

Co-authors: Dr BLANQUER, Ignacio (UPVLC); Dr PLASENCIA, Isabel (CSIC)

Presenter: BORGES, Goncalo (LIP)

Session Classification: VREs - Community Contributions

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 96

Type: **Presentation**

Performance Prediction of Cloud Applications Using Dwarf Benchmarks

Thursday, September 20, 2012 11:20 AM (20 minutes)

Description of the work

We have developed a benchmark suite for measuring the performance of virtualised hardware based on the earlier work carried out in the TORCH project [1]. This collection uses the concept of algorithmic patterns, in the literature referred to as “dwarfs”. “A dwarf is an algorithmic method that captures a pattern of computation and communication”[2]. The proposed dwarf taxonomy [2] consists of 13 frequently occurring patterns that can be considered as building blocks of applications, such as dense matrix operations, map reduce, particle and spectral computations. This feature makes them especially suitable for performance modelling. For validating the prediction technique we have included applications from different domains such as molecular dynamics (Gromacs [3]), rendering (Blender [4]) and video transcoding (FFMPEG [5]). We have carried out extensive benchmarking on BonFIRE [6] project resources and on facilities of commercial Cloud providers such as Amazon EC2, Rackspace, CloudSigma, GoGrid and ElasticHosts. As a mathematical apparatus for modelling initially we used polynomials and later the Moore-Penrose inverse matrix [7] that enabled us to construct multi-dimensional models by combining the data from all dwarfs and has significantly improved the accuracy of predictions. Our work has demonstrated that the concept of algorithmic dwarfs can be successfully used not only for the characterisation of Cloud resources but also for performance modelling. We have also investigated the problem of reliability of long term experiments and the impact of various levels of contention on processes running on virtual resources.

Link for further information

<http://eprints.soton.ac.uk/273157/>

Wider impact of this work

The concept of algorithmic dwarfs has been presented a few years ago. However, the modelling power of this approach has not been put to a rigorous test until now. We have collected data from several cloud providers and validated this concept on four applications drawn from different computational domains. We believe that our work can be used by several stakeholders in Cloud market for:

- a) making better provisioning decisions: deploying the infrastructure resources required for a given application QoS rather than over-provisioning.
- b) making better application scheduling decisions: knowing the application runtime with a good reliability permits more intelligent scheduling.
- c) determining the optimal application configuration: the performance of complex applications and business or industrial data processing workflows with many components can be greatly affected by their configuration.

Printable Summary

The EC-FP7 BonFIRE project offers a multi-site Cloud facility of geographically distributed testbeds of heterogeneous resources for Future Internet experimentation. IT Innovation has taken part in the project as one of three driving experiments, focusing on QoS-oriented service engineering for federated Clouds. In this presentation, we will discuss the challenges we have addressed pertaining to resource provisioning for applications that have specific QoS requirements and where consumers of Cloud resources want to avoid lock-in to any specific IaaS provider. This is a significant challenge due to the limited and varying information IaaS providers disclose about infrastructure resources. This presentation covers the use of application-focused benchmarks, referred to as Dwarfs, for characterising compute resources and for application performance modelling. This work benefits both researchers and software developers involved in the SaaS, Paas and IaaS areas of Cloud computing.

Primary author: Mr BONIFACE, Michael J. (IT Innovation Centre, University of Southampton)

Co-authors: Mr PAPAY, Juri (IT Innovation Centre, University of Southampton); Mr PHILLIPS, Stephen C. (IT Innovation Centre, University of Southampton); Mr ENGEN, Vegard (IT Innovation Centre, University of Southampton)

Presenter: Mr PAPAY, Juri (IT Innovation Centre, University of Southampton)

Session Classification: Virtualised Resources

Track Classification: Virtualised Resources: challenges and opportunities (Michel Drescher: track leader)

Contribution ID: 97

Type: **Poster**

GaaS: toward a more “elastic” and sustainable grid environment

Description of the work

The GaaS model combines the advantage of providing users with an usage model that is familiar to the traditional Grid, with the possibility of flexible management of computational resources in a IaaS-like manner. Hence, our model can be classified as a Platform-as-a-Service for extending Grid environments with elastic (i.e., virtual) resources. By using GaaS, users can define new Grid Sites, add computational resources to existing Grid Sites and modify the resources aggregation scheme, i.e., site queues. All GaaS services may be requested by users, with certain roles, from the UI. GaaS actually consists of four services to (a) add new computing resources (WN service: GaaS_WNS); (b) aggregate existing computing resources in a new queue (queue service: GaaS_QS); (c) add a new grid site for an existing VO (grid site service: GaaS_GSS); (d) create a suited runtime environment for a set of applications on existing or new computing resources (application environment service: GaaS_AES). The GaaS prototype, deployed on the S.Co.P.E. infrastructure at University of Naples, is based on the gLite-EMI middleware and on the OpenNebula cloud management system with Xen hypervisor to create virtual machines (VM) that host WNs, CE and IS. The main efforts in the prototype development were (i) the definition of templates for gLite-EMI services configuration, and (ii) the enabling of their fast provisioning. We designed our VM disk provisioning system in order to provide fast VM creation and avoid as much data copy as possible. Our solution is based on the GNU/Linux’s Logical Volume Manager (LVM) system and a set of grid roles configuration templates. This approach allows a really fast creation of snapshots starting from a reference LV and the later configuration of the VM with the requested grid role.

Wider impact of this work

GaaS flexibility provides several advantages to traditional Grid infrastructures in different scenarios: WNs can be customized with software tailored to a given set of users, as well as queues can be configured to fulfill a specific computation needs. Moreover, GaaS support the creation of complete Grid sites in order to, i.e., enable a community that has to share resources for the life time of a project, to avoid the burden of configuring from scratch all the required services and resources. Finally it must be taken in account that GaaS, being an on-demand solution, can be also useful to optimize the energy efficiency of large scale systems reducing their overall operational cost and improving their sustainability.

Printable Summary

The work on GaaS (Grid as a Service) started two years ago in order to find a solution to make distributed computing environments, based on Grid model, more “elastic”and more sustainable by exploiting the features of the Cloud model. The Grid model is basically static and users cannot define new Grid sites, add computational resources to existing ones and modify the resources aggregation scheme in accordance to their needs; it is also not possible to dynamically modify the

resources number on the basis of the real system workload, in the name of saving energy and to achieve a more efficient and sustainable environment. During the poster session we will describe the steps taken to design and to deploy GaaS prototype on the S.Co.P.E. production infrastructure at University of Naples.

The presentation is aimed at both scientists who use and technicians who manage grid environments because they will gain an understanding of how this work can be re-used for their own.

Primary author: Mr BARONE, G.B. (Università degli Studi di Napoli Federico II)

Co-authors: Mr BOTTALICO, D (Università degli Studi di Napoli Federico II); Mr CARRACCIUOLO, L (Università degli Studi di Napoli Federico II); Mr BIFULCO, Roberto (University of Naples Federico II); Dr BOCCIA, Vania (Università degli Studi di Napoli Federico II)

Track Classification: Virtualised Resources: challenges and opportunities (Michel Drescher: track leader)

Contribution ID: 98

Type: **Internal Project Meeting**

GGUS Advisory Group - Introduction

Friday, September 21, 2012 11:00 AM (10 minutes)

Description of the work

The development of GGUS, the tool behind the EGI helpdesk and the EGI technology helpdesk, has been guided by advisory groups since the days of EGEE-II, when first the ESC (Executive Support Committee) and later on the USAG (User Support Advisory Group) were fulfilling this role. With the start of EGEE, this role was shared between the OTAG (Operational Tools Advisory Group) and the USAG (User Services Advisory Group). After the USAG was discontinued it was decided that a GGUS specific advisory body should be founded, bringing on board all the stakeholder from EGI, NGIs, VRCs and Technology Providers. The group will meet on a regular basis to discuss the high level strategic direction the development of GGUS should be taking.

Link for further information

<https://indico.egi.eu/indico/sessionDisplay.py?sessionId=58&confId=1019#20120921>

Wider impact of this work

The EGI helpdesk is one of the central tools used to communicate within the EGI ecosystem, inside and across the borders of projects, infrastructures and user communities. As such it is vital that the general strategic direction of the development is discussed with all the stakeholders. This face to face meeting aims at kick starting the new advisory group.

Printable Summary

This is the first face to face meeting of the recently founded GGUS Advisory Group. GGUS (Global Grid User Support) is the tool behind the EGI helpdesk. Since the EGI helpdesk is used for various purposes by a large number of different user groups, the establishment of an advisory group regularly discussing seemed necessary. Amongst other things, the meeting aims to further detail the composition of the group, its processes and frequency of convening. There will be presentations on current development and strategic topics.

This is a closed meeting - anyone intending to attend should contact the Chairman, Dr Torsten Antoni.

Primary author: Dr ANTONI, Torsten (KIT-G)

Presenter: Dr ANTONI, Torsten (KIT-G)

Session Classification: GGUS Advisory Board

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 99

Type: **Workshop**

WS on GlobusOnline/GridFTP integration into a Cloud-Installation

Friday, September 21, 2012 1:30 PM (1h 30m)

Description of the work

The integration of Clouds and Grids is at the core of this project. With an IBM CloudBurst installation containing two different cloud stacks for evaluation, research and testing purposes, the presented activities show the integration of single-sign-on systems with Clouds and basic Grid technologies. The first step of this project was to create a common share with user-specific directories to enable easy transfer to and from all virtual machines of one user.

The project will be presented in the following steps:

- Systemstructure (10 min)
- GridFTP-Installation&Configuration (10 min)
- MyProxy-Installation&Configuration (5 min)
- GlobusOnline Integration (5 min)

Wider impact of this work

Through a centralized share per user and the integration of GridFTP file transfer to larger amounts of virtual machines in Infrastructure-as-a-Service cloud environments for users file management and file transfers to and from clouds are improved enormously.

Printable Summary

RISC Software GmbH runs an IBM CloudBurst instance which has been enabled for GRID Usage. The Company created a storage server that speaks GridFTP, MyProxy and is also enabled for GlobusOnline usage and this can be accessed via individual cloud instances (virtual machines). Once deployed, special cloud images automatically hook into the authentication infrastructure and mount the exported data share from the file server. This means that the cloud users get an easy way to access their data via GridFTP and GlobusOnline without having the complexities of setting this up themselves. This presentation is aimed at scientists, software developers and resource providers who will gain an understanding of how this work can be re-used for their own virtualized computing platform.

Primary authors: Mr KRIEGER, Michael (Leibniz-Rechenzentrum (LRZ)); Mr HENNERBICHLER, Wolfgang

Presenter: Mr HENNERBICHLER, Wolfgang

Session Classification: Virtualised Resources

Track Classification: Virtualised Resources: challenges and opportunities (Michel Drescher: track leader)

Contribution ID: 100

Type: **Presentation**

dCache: challenges and opportunities when growing into new communities

Tuesday, September 18, 2012 11:40 AM (20 minutes)

Description of the work

The presentation will give a summary of experiences gained from working to support non-HEP communities, such as XFEL and LOFAR and how adopting standards has been key in facilitating how dCache may be used by other communities. The reasons how (and why) the HEP community are currently using non-standard protocols are mentioned along with the outlook for the future. The talk will describe challenges when moving from the (relatively) well-organised and structured HEP communities to the more chaotic collaborations. There are challenges both within dCache (e.g., how to authenticate end-users and authorise their activity) and challenges at the boundary of dCache and other services (e.g., facilitating data movement and complex data management). The talk will describe how recent changes in dCache target these problems and how future changes are aimed to solve them. The presentation will also describe how dCache is opening up to support a more community-driven approach; a more inclusive approach will allow more flexible adoption. A new software license, along with spinning off dCache components as separate projects, allows for easier re-use within other projects. Some examples of successes are provided. In addition, reformulating the existing flexibility in dCache as support for plugins provides a low-overhead entry-point for community involvement. The presentation will also include details from the “dCache labs” project. These are experimental new features that the dCache team will be providing as technological previews, without promising that they will be adopted by dCache core; examples include our ongoing investigation with integrating shared file-systems. This is done to allow dCache administrators to provide feedback on features.

Link for further information

<http://www.dcache.org/>

Wider impact of this work

Different communities are now starting to generate HEP-like quantities of data that they now need to store and process, communities that hitherto could fit their experimental data on single server solutions are now faced with a data explosion. Many of the problems from this data explosion are problems that the HEP community has already experienced and found solutions, some of which are encoded within dCache; for example, the ability to split responsibility for digital curation and data access from scientific analysis allows a common approach across different disciplines, with a corresponding economy-of-scale. By adopting existing solutions and, importantly, by adjusting existing solutions so they are adoptable, the effort required for these scientific communities to support their Big Data problems is reduced, allowing them to focus on their scientific programme.

Printable Summary

The dCache project started in 2000 in order to provide a common commodity storage front-end to different tape systems for use within the High-Energy Particle Physics (HEP) community. Since then, the project has morphed into a generic storage system, providing both disk-only and disk-and-tape systems. Although dCache is mostly heavily used by the HEP communities, it is being adopted by ever increasing number of communities beyond. This talk will describe some of the challenges in adapting dCache to support a wider community of users, the recent changes in dCache to accommodate these challenges and our plans to support an increasingly wide user-base. This presentation is aimed at people from non-HEP communities that are looking for solutions to their Big Data problems, administrators who are interested in supporting a wider range of users with their existing dCache infrastructure, and people generally interested in challenges when dealing with data in large communities.

Primary author: MILLAR, Paul (DESY)

Presenter: MILLAR, Paul (DESY)

Session Classification: New middleware products

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 101

Type: **Presentation**

Compute Client Consolidation and Harmonization in the EMI Project

Tuesday, September 18, 2012 2:30 PM (20 minutes)

Description of the work

This activity was started to survey the existing compute client solutions within and outside of EMI, and to recommend a way of harmonizing and consolidating them. We evaluated the new EMI-ES Clients, the existing WMS and CREAM clients, the UNICORE/X clients, the ARC clients, the L&B clients and SAGA. We surveyed their available platforms, features and capabilities (job execution, management, brokering, workflows, credential handling), users and user communities, programming languages, documentation status, developer efforts and post-EMI plans. Then we collected several scenarios for possible consolidation, and ranked them, recommending the three best ones. Then after the decision of the Technical Board of the EMI project the chosen one is put into the development plans of the final year of the EMI project.

Link for further information

https://twiki.cern.ch/twiki/bin/view/EMI/EmiJra1T2Compute_Client

Wider impact of this work

The consolidation and harmonization of the compute clients can result in a better quality client library shared and maintained by all the interested parties together. The developers of gLite, UNICORE and ARC can work together to provide a much better experience for command line users and application developers, by making this common library easier to use and also more powerful by providing access to the now execution service interface of EMI, and because this would be the only client providing this access, it has big impact of the future of the EMI-ES execution interface.

Printable Summary

The consolidation and harmonization of the compute client tools and APIs is an ongoing activity within the EMI project to reduce the maintainable code and to provide a higher quality software. This presentation will describe the decisions and the progress we made so far. It is aimed at software developers writing applications using client APIs, and at command line users of grid middlewares, who will gain an understanding how this development can affect their work.

Primary author: HAGEMEIER, Bjoern (JUELICH)

Co-authors: MERZKY, Andre (Center for Computation and Technology, Louisiana State University); SUSTR, Zdenek (CESNET); NAGY, Zsombor (NIIF)

Presenter: HAGEMEI,ER, Bjoern (JUELICH)

Session Classification: New middleware products

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 102

Type: **Presentation**

Commercial applications of open source middleware: the EMI and DCore experience

Thursday, September 20, 2012 2:05 PM (20 minutes)

Wider impact of this work

The experience collected during the collaboration of EMI and DCore is extremely valuable as a concrete proof that the commercial exploitation of open source software is possible also in the area of middleware in particular and of software for scientific applications in general. If the model produces the expected results, the collaboration could become an important success story and motivate further collaborations on similar activities.

Description of the work

The EMI sustainability plan includes a number of activities focused on commercial exploitation of its open source middleware. DCore is a Swiss-based private company with operations in several European countries. EMI and DCore together have started about a year ago a collaboration to exploit a few of the major EMI services as part of the DCore commercial services. During the past 12 months EMI and DCore together have developed a business model with the potential of creating a revenue generating mechanism that benefits both the commercial company and the academic Institutes involved in the collaboration. A number of important use cases have been analysed and concrete prototypes of commercially exploitable services and applications have been developed.

Printable Summary

The commercial exploitation of the open source middleware services developed by EMI is one of the main objectives of the EMI project and a fundamental part of its sustainability plan. This presentation describe the work done by EMI in collaboration with DCore System SA, a Swiss technology company with a strong focus on open source software, data management and security. The long-term objectives of the collaboration and the potential revenue generating mechanisms are described together with a number of concrete examples of the results achieved so far. This presentation is targeted at development managers in research Institutes interested in knowing more about potential commercial exploitation of distributed software services. It is also targeted at representatives of commercial companies interested in finding out the challenges and opportunities of open source software exploitation

Primary authors: DI MEGLIO, Alberto (EMI); ESTRELLA, Florida (CERN); Mr HAKALY, Robert (EMI)

Presenters: DI MEGLIO, Alberto (EMI); Mr HAKALY, Robert (EMI)

Session Classification: Sustainability of Technology Providers

Track Classification: Community and Co-ordination (Sergio Andreozzi: track leader)

Contribution ID: 103

Type: **Presentation**

Astro Community and DCIs: user requirements, use-cases, test-beds & VRCs

Tuesday, September 18, 2012 2:20 PM (20 minutes)

Description of the work

The session, chaired by the current coordinators of the A&A HUC, will initially illustrate where the A&A HUC is now within the EGI ecosystem, trying to identify what still hampers the adoption of DCIs by the A&A HUC; this discussion will also include the status of proposals for new A&A VRCs, those presented in the past as well as those brought for the first time in this session. The second part of the session then will be dedicated to discuss and analyse the pool of user requirements, use-cases and test-beds proposed by the research leaders, scientists and software developers who attend the session; for each of the proposed use-cases and test-beds it will be discussed their scientific background, their ability to represent a whole class of applications and workflows and all aspects related to their porting on various DCI related platforms. The session should also be useful to identify astronomical Institutes, projects and groups able to bring detailed workflow proposals to be integrated in Science Gateways and used through them. The final part of the session will be dedicated to conclusions and next steps to be undertaken to activate new VRCs and to port the analysed use-cases, test-beds, applications and workflows on DCIs.

Wider impact of this work

If accepted and successful, it is expected that the outcome of the session will significantly impact the A&A HUC in relation to DCIs. The success of the session, in fact, means a progress in identifying and creating new VRCs and new applications and workflows to port on DCIs, and this is crucial to improve the effectiveness of the coordinating action within the HUC. Attendees will be invited to present one or more use-cases and test-beds related to their scientific work to be discussed and analysed during the session in order to understand how they can be ported on various DCIs. The purpose of the session is therefore to encourage a discussion of the technical aspects related to concrete proposals related to the creation of new VRCs and the porting of applications and workflows on DCIs rather than to merely host a number of presentations.

Printable Summary

A&A has been recognized as one of the HUCs in EGI. Its coordination aims at promoting the adoption of DCIs within it. To achieve this objective the creation of a number of well-focused A&A-related VRCs in EGI is mandatory. To this end several meetings of the community have been organized in the past. Now this process has to continue; it is necessary to further push the activity aimed at gathering user requirements, use-cases and test-beds, applications and workflows to port and run on DCIs, all of them useful to increase the number of potential users and to identify and help the creation process of candidate VRCs. The required 270-minutes session is aimed at research and project leaders, scientists and software developers and will: a) describe the steps taken so far and the current status of the HUC; b) discuss about new candidate VRCs and the related

creation process; c) discuss the requirements and the pool of use-cases and test-beds brought by the attenders of the session.

Primary author: VUERLI, Claudio (INAF)

Presenter: VUERLI, Claudio (INAF)

Session Classification: Experience in community building

Track Classification: Community and Co-ordination (Sergio Andreozzi: track leader)

Contribution ID: **107**Type: **Demonstration**

MPI in EGI

Description of the work

The Virtual Team project was created by EGI-InSPIRE to detect significant issues with support for MPI jobs within EGI. The project started in November 2011 and finished in May 2012. During its six month lifetime the project has collaborated with different user communities, NGIs user support teams, middleware technology providers and resource providers to identify issues and to establish services by which MPI applications can work successfully in the European Grid Infrastructure. The work spanned across a number of technical areas and these will be all covered in the demonstration:

- Documentation: Improved documentation has been prepared in the EGI wiki for site administrators and for application developers. These provide guidance to to configure and to use MPI resources correctly.
- Nagios probes: New monitoring probes for the EGI Service Availability Monitor has been defined. These will be implemented and put into production by the Heavy User Community and Operations teams.
- Information system: The typical problems with the registration of MPI resources have been collected and reported to Operations. The Nagios probes have been designed to be able to detect these problems.
- MPI VO: A new VO which includes only correctly configured MPI sites have been setup on the production infrastructure. The VO can be used to port MPI applications to EGI. During the demo MPI members will show how many MPI resources are available in EGI and how to use them. Real MPI applications will be sent to show the capabilities of the VO.
- Batch systems: Issues with interfacing MPI applications and some of the local batch job schedulers of EGI have been collected and addressed.
- Accounting: Issues with collecting accounting information about parallel applications have been collected and reported to responsible technology developers and providers with request for addressing.

Wider impact of this work

This demo will shown the current MPI features and the work done by the MPI Virtual Team to fix MPI issues in EGI. It will show that parallel jobs can successfully use resources from the NGIs that participate in EGI.

Printable Summary

The EGI community receives dedicated support from NGIs through the EGI-InSPIRE 'Heavy User Community' activity yet despite this, there have been significant issues which have detracted from user uptake and satisfaction of the available MPI services. The EGI community therefore conducted a six month 'Virtual Team project' to address and resolve these issues. As a result, improved and expanded MPI support and services have been implemented for both resource providers and users. This contribution will showcase the improved MPI services and will demonstrate a set of scientific

applications that are already benefiting from the project's results. The demonstration aims to promote the EGI MPI services and increase their uptake by resource providers and by developers of parallel scientific applications.

Primary authors: SIMON, Alvaro (FCTSG); Dr SIPOS, Gergely (EGLEU)

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: **108**Type: **Workshop**

EGI Scientific Publications Repository

Description of the work

The ultimate goal is to improve the evidence of the EGI scientific impact. This session will describe the proposed processes, policies and tools to be established.

Link for further information

https://wiki.egi.eu/wiki/VT_Scientific_Publications_Repository

Wider impact of this work

It is aimed at NGIs/VOs representatives and all people interested in demonstrating the scientific impact of their research or infrastructure and will give the opportunity to provide feedback to the defined plan before starting the implementation.

Printable Summary

The Scientific Publications Repository project started in June 2012 in order to define tools, policies and processes to track the scientific outcome that was possible thanks to the usage of the European Grid Infrastructure (EGI). The ultimate goal is to improve the evidence of the EGI scientific impact. This session will describe the proposed processes, policies and tools to be established. It is aimed at NGIs/VOs representatives and all people interested in demonstrating the scientific impact of their research or infrastructure and will give the opportunity to provide feedback to the defined plan before starting the implementation.

Primary author: ANDREOZZI, Sergio (EGLEU)

Track Classification: Community and Co-ordination (Sergio Andreozzi: track leader)

Contribution ID: 109

Type: **Poster**

Testing GlusterFS in a Grid Site for SuperB collaboration

Description of the work

In the proposed poster we show the results of the test obtained by a practical experience made on an high performance hardware. The setup which we used is mostly close to one possible production cluster. The aim of the tests is understand at first the global reliability of this kind of configuration, the performance and the scalability limits of GlusterFS. We proceeded in a systematically way by measuring the network occupation during large data transfer between the nodes, then we measure how the GlusterFS demon affect the CPU of each node. Finally we use standard I/O benchmark to measure the read/write limits.

After general evaluation we measure realistic behaviour to respect the SuperB usage, by sending burst of analysis job over simulated data coming from the previous MonteCarlo production created by the collaboration and store in the GlusterFS area.

Wider impact of this work

The interest for GlusterFS is due to the high performance, the possibility to setup hard disk in JBOD, the standard posix interface and the auto replication feature inter and intra cluster. Moreover, the Gluster community roadmap plans to make available soon new API to individuate the disks that contain a single file, that open the possibility to implement a file affinity scheduling in the future.

Printable Summary

The aim of this work is to present our experience and future plans to use GlusterFS at site level. The goal is to create a storage area that is shared between storage element and worker nodes in a Grid Cluster. GlusterFS is an open source cluster file system that allows creating a single namespace and managing the data redundancy and data replication automatically. In order to evaluate the system performance, we have setup a test infrastructure composed of 12 servers over a 10Gbit/s network. Each node shares its disk space through GlusterFS, and mounts a unique file system as the sum of every single 'brick'. The performed tests show interesting results in terms of performance, CPU usage, Network, I/O and Grid Integration.

The presented activities are performed within the SuperB R&D program to study the future computing model.

Primary author: Prof. PARDI, Silvio (INFN and UNINA)

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 110

Type: **Poster**

Grid Computing for Empirical verification of the even Goldbach conjecture.

Description of the work

The Empirical Goldbach Conjecture verification has been started and coordinates by Tomás Oliveira e Silva since 2005. Before October 2011 the code had run on several supercomputing machines and on the spare time of many computers until reach the $26 \cdot 10^{17}$. After that, we porting the code in Grid by creating a set of scripts able to manage the complete computation lifecycle: generate dynamically the next numbers to analyze, create the JDL file, then submit the job in Grid, retrieve the output and resubmit the failed jobs. In particular to do that, we take advantage from the Parametric-Job features provided by the WMS component of gLite. About of the 50% of the jobs submitted through the Grid ran on the SCOPE datacentre in Naples, while the rest of the work has been computed in other infrastructure within the Italian Grid Infrastructure (IGI) and in many other site of the EGI by using different virtual organizations: unina.it and infngrid.

Link for further information

<http://www.ieeta.pt/~tos/goldbach.html>

Wider impact of this work

This work represents a success experience to porting a new application on Grid by scratch and take advantage by the resource availability of the EGI infrastructure. Moreover this work shows a new application domain for the distributed infrastructure, represented by the computer algebra and can be a starting point to individuate a larger Mathematician community.

As regarding the mathematical point of view, our verification represents actually the world record of computation. A recent proof (by Terence Tao) that any odd number larger than 1 is the sum of at most 5 primes uses the fact that the Goldbach Conjecture is true up to $4e14$. It seems likely that a (future) proof that any odd number larger than 1 is the sum of 3 primes will require a verification of the Goldbach conjecture to a higher limit perhaps $4e18$ will suffice!

Printable Summary

In this work we present our experience in use of Grid technologies to undertake empirical verification of the Goldbach Conjecture. The large availability of resources provided by EGI infrastructure and the advanced use of job submission techniques allowed us to improve the rate of “verified numbers” per day. After the porting in grid of the analysis code, the total amount of verified numbers in 7 month has been of $1,410^{18}$ against the 110^{18} verified in the previous 21 months. In 2012 we reached the desired verification limit ($4 \cdot 10^{18}$), that overcome the previous world record of computation ($4 \cdot 10^{14}$). Moreover we discovered the largest prime of a minimal Goldbach partition. In our activities, no counter-example of the conjecture was found but several statistical behaviours have been confirmed. Full Mathematical considerations are contained in a more extensive work,

while this poster proves that we can take advantage from the Grid to enlarge the knowledge in the math filed.

Primary author: Prof. PARDI, Silvio (INFN and UNINA)

Presenter: Prof. PARDI, Silvio (INFN and UNINA)

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: **111**

Type: **not specified**

Discussion

Tuesday, September 18, 2012 12:05 PM (25 minutes)

Presenter: GATER, Catherine (EGLEU)

Session Classification: Marketing and Communications Strategies

Contribution ID: 113

Type: **Presentation**

DECIDE: An e-Infrastructure for studying neurodegenerative diseases

Tuesday, September 18, 2012 12:06 PM (24 minutes)

Description of the work

The presentation will go through the motivations and objectives of the project, and will present an overview of the lessons learnt over the project lifetime, in particular over the 6 months since the service is available to users. We will recap the building blocks of the service (which is offered via a Science Gateway where access is mediated by identity federations), the specific Grid services adopted (AMGA, gLibrary, SecureStorage), the methodology for porting applications to the Science Gateway, the strategy adopted to make sure the developed service would be acceptable by the targeted communities. Moreover, the presentation will also discuss other aspects, like training and sustainability roadmap, which are critical in ensuring the developed service will reach a critical mass of users and will survive after the project's end.

Wider impact of this work

.

Printable Summary

FP7 project DECIDE has successfully built an e-infrastructure and an e-service dedicated to researchers and clinicians working in the field of Alzheimer's and other neurodegenerative diseases. The service has been designed taking into account the specific requirements of these communities, in terms of ease of use, data protection and confidentiality, access control to applications, definition and publication of standardized operational procedures. The project started in September 2010 and a first set of applications is available to users since April 2011.

Presenter: GALEAZZI, Fulvio (GARR)

Session Classification: Research Infrastructures

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 114

Type: **Presentation**

EGI Applications Database

Tuesday, September 18, 2012 4:00 PM (10 minutes)

Description of the work

.

Wider impact of this work

.

Printable Summary

The EGI Applications Database (AppDB) stores information about grid-based computing tools for scientists to use. The scope of the database embraces all scientific fields, from resources to simulate exotic excitation modes in physics, to applications for complex protein sequences analysis.

The applications filed in AppDB are finished products, ready to be used. Using applications already enabled for EGI means that you, as a scientist, don't have to spend your research time adapting software to the grid.

The goal for AppDB is to inspire scientists less familiar with programming to use the grid and its resources. The database is also meant to avoid duplication of effort across the user community.

The database has been online since 1 July and is the natural successor to EGEE's database, which was developed by Italian and Greek teams during that project's third and final phase.

Primary authors: Dr SIPOS, Gergely (EGI.EU); CHATZIANGELOU, Marios (IASA)

Presenter: CHATZIANGELOU, Marios (IASA)

Session Classification: Software services for community building and support

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 115

Type: **Presentation**

EGI Training Marketplace

Tuesday, September 18, 2012 4:10 PM (10 minutes)

Description of the work

.

Wider impact of this work

.

Printable Summary

The EGI Training Marketplace is a service to coordinate training across communities, projects and national teams. The new Training Marketplace enables trainers to advertise events, materials and resources, and researchers using EGI to locate, access and comment on training material, events and resources that meet their needs. If the specialist course, resource or material that you are looking for is not available in the marketplace, send us an email so we can develop it with our partners!

Primary author: DEVEREUX, Claire (STFC)

Presenter: DEVEREUX, Claire (STFC)

Session Classification: Software services for community building and support

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 116

Type: **Presentation**

EGI Client Relationship Management System

Description of the work

.

Wider impact of this work

.

Printable Summary

The merger of the NA2-NA3 activities was driven by the need to give more focus and resources to outreach to new user communities. A coordinated outreach activity needs a tool where NGIs and EGI.eu can record contacts, leads, conversations. It was decided to setup a Client Relationship Management (CRM) system for this purpose.

Primary author: BORGES, Goncalo (LIP)

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: **117**

Type: **not specified**

Discussion

Tuesday, September 18, 2012 4:50 PM (40 minutes)

Session Classification: Software services for community building and support

Contribution ID: **118**

Type: **Presentation**

EGI virtual machine image marketplace

Tuesday, September 18, 2012 4:20 PM (10 minutes)

Description of the work

.

Wider impact of this work

.

Printable Summary

.

Presenter: Dr TURILLI, Matteo (Oxford e-Research Centre)

Session Classification: Software services for community building and support

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 119

Type: **Presentation**

EGI Requirements Tracker

Description of the work

.

Wider impact of this work

.

Printable Summary

Researchers using the infrastructure and operations teams can submit their service requirements as tickets to the EGI Requirements tracker (RT) dashboard. The tickets will be handled by the EGI.eu user support and operations teams and categorised by community (for example NGIs, projects, virtual organisations) or according to their status (new, open, accepted, resolved).

The RT system will help the user support team to keep track of what needs to be done and lets the whole EGI community know. It's a global view of what European scientific communities need from EGI. Requirements will be posted in an open ticketing system, which makes the gathering and solution provisioning processes open and transparent. This allows the community to check if what they need has already been required by other users, teams or NGI staff and will save time for user communities, for technology providers and avoid duplication of efforts.

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 120

Type: **Presentation**

Experiences with ELIXIR in the Czech Republic

Tuesday, September 18, 2012 2:22 PM (22 minutes)

Description of the work

.

Wider impact of this work

.

Printable Summary

Approximately two years ago activities leading to the setup of national ELIXIR node started in the Czech Republic. Under the coordination of the Institute of Organic Chemistry and Biochemistry of the Czech Academy of Science, 5 universities and several research institutes joined forces with the national e-infrastructure players (representing network, grids, clouds, data, and high performance computing) to define the framework for the establishment of the national ELIXIR node. The collaboration between scientific bioinformatics community and the e-Infrastructure bodies lead to the joint setup of the first ELIXIR hardware node (space and electricity provided by UOCHB, hardware by CESNET/NGI and experience by CESNET and CERIT-SC). The future plans include further collaboration on the expansion of the current e-Infrastructure architecture to also encompass needs and expectations of the ELIXIR community. We expect this to be a positive example of successful collaboration between e-infrastructure architects and non-HEP scientists.

Primary authors: Dr VONDRASEK, Jiri (UOCHB AV CR); MATYSKA, Ludek (CESNET)

Presenter: MATYSKA, Ludek (CESNET)

Session Classification: Research Infrastructures

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: **121**

Type: **not specified**

Discussion

Contribution ID: 122

Type: **not specified**

Discussion

Contribution ID: 123

Type: **Presentation**

Centralized user management and Single Sign On for the WeNMR gateway through the WeNMR Virtual Research Community

Wednesday, September 19, 2012 12:15 PM (15 minutes)

Description of the work

The WeNMR services are hosted by project partners throughout Europe. Historically these are stand-alone services, each equipped with their own user management and authentication systems. This fragmented approach is however undesirable from an user perspective, as it requires separate login credentials for all services, as well as from a management perspective. We thus decided to equip the WeNMR VRC with a centralized user management and authentication system that allows VRC members to use all services with their VRC login credentials. At the heart of the system is a secure SQL authentication database that contains a synchronized copy of the VRC user records. A custom Drupal module in the VRC manages the synchronization to ensure that all portals have access to up-to-date user data when they query the SQL database for authentication. In this procedure, the passwords are communicate only as MD5 hash, for increased security. When creating a new VRC account, a secure connection (ssl) is used to check if the user's personal certificate (installed in the browser) has been issued by a CA recognized by EGI. This is done using the public key of the certificate. If this check is successful, then the server extracts the subject from public key and checks also the VOMS account list. The certificate subject and expiration date are stored in the database within the access credentials.

Wider impact of this work

Furthermore, the aforementioned Drupal module provides WeNMR partners with an administration interface to register a new service. WeNMR VRC members are then able to manage their service subscriptions through a convenient dashboard on their VRC account web page. In registering a new service the WeNMR partner can use the Access Control List (ACL) to fine grain the service sign-up process with additional requests to the user such as obtaining a separate license agreement.

Finally, the user management and authentication API allows service portals to store information about the users activity in the SQL database. Such information includes the total number of service requests by the user or the state of an active server request. Presented via the account dashboard, this information provides a convenient means for the WeNMR members to keep up-to-date information on the status of the services they are using and provides the WeNMR project with detailed means of tracking usage of the services.

Printable Summary

WeNMR is both a three years project funded under the European Commission's 7th Framework Programme (e-Infrastructure RI-261571) and a Virtual Research Community supported by EGI. WeNMR aims at bringing together research teams in structural biology into a virtual research

community at a worldwide level and provide them with a platform integrating and streamlining the computational approaches necessary for NMR and SAXS data analysis and structural modelling. The WeNMR Virtual Research Community (VRC) provides access to the services, a knowledge base, help center and various information and (social) communication channels. Access to the services is provided through easy-to-use web interfaces using the computational resources provided by the grid through the eNMR Virtual Organization (VO). With over 430 registered users (June 2012) the eNMR VO is the largest in the life sciences.

Primary author: Mr VAN DIJK, Marc (WeNMR)

Co-authors: BONVIN, Alexandre (eNMR/WeNMR (via Dutch NGI)); GIACHETTI, Andrea (CIRMMP); ROSATO, Antonio (CIRMMP); Dr VERLATO, Marco (INFN)

Presenter: Mr VAN DIJK, Marc (WeNMR)

Session Classification: AAI Workshop

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 124

Type: **Presentation**

Experiences with integrating science gateways with identity federations

Wednesday, September 19, 2012 11:00 AM (15 minutes)

Description of the work

.

Wider impact of this work

.

Printable Summary

INFN Catania and COMETA have done a considerable amount of work to integrate their Science Gateways with SAML-based systems and register them as Service Providers of several national Identity Federation and of the eduGAIN inter-federation. Benefits and issues will be presented and discussed. A special focus will be given to problems encountered in the day-by-day administration of the Science Gateway for users with federated identities.

Primary author: FARGETTA, Marco (INFN)

Co-author: BARBERA, Roberto (University of Catania and INFN)

Presenter: FARGETTA, Marco (INFN)

Session Classification: AAI Workshop

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 125

Type: **Presentation**

Generic User Management for Science Gateways via Virtual Organizations

Wednesday, September 19, 2012 11:15 AM (15 minutes)

Description of the work

The MoSGrid science gateway provides features to intuitively utilize UNICORE DCIs via gUSE services (grid User Support Environment) for domain-specific applications. Many middlewares, and in particular UNICORE, use X.509 certificates for authentication. The certificates are often available in the end users' web browser. Therefore, the security mechanisms in CAS have been extended to improve the automatic login into Liferay via certificates stored in web browsers. This simplifies the authentication for the users. Additionally, the certificates are not only used for the authentication to the science gateway, but also for the creation of short-lived credentials to gain access to suitable DCIs. Each DCI checks whether a credential is valid against its own user management. In German grid infrastructures, user accounts are typically maintained via a central database, which includes information about all valid grid certificates and associated VOs. Portal frameworks like the award-winning Liferay do not support the automatic use of VO membership information. Currently, users have to apply for the MoSGrid user role via an email to gain access to all provided features and all DCIs supporting the MoSGrid VO. Administrators of the science gateway check whether an applying user is member of the VO and, if applicable, assign suitable roles to the associated user account. To improve this situation, another extension to CAS is currently being developed. CAS will check the user provided certificate information and passes it on to Liferay, which checks it against VO membership information, previously fetched from the central database. On the one hand, the list of members of a VO is smaller than all valid certificates. Thus each VO can be efficiently used as a user authorization list. On the other hand, this information will be used by the Liferay user management in order to improve the user experience by showing and hiding features depending on VO membership.

Wider impact of this work

Impact Authentication and authorization via X.509 certificates and VO memberships are standardized for supporting communities with suitable DCIs, not only in Germany but also on international level. The extension of Liferay with CAS to enable logging in via a certificate imported in the browser and to make use of VO membership information is generally applicable. The source code will be published under an open source license and, thus, can be re-used for other Liferay science gateways. So far, users are guided in taking the necessary steps towards first simulations by documentation and authorization involves checking each user individually by the science gateway administrator. In the near future the automatic usage of VO membership information will allow for a more compact documentation, much easier user authorization handling, and the automatic access to suitable DCIs.

Printable Summary

In general, science gateways provide features to access domain-specific applications on distributed computing infrastructures (DCIs). Independent of the technology used for authentication of a user to the science gateway, the membership in a virtual organization (VO) mostly defines which DCIs are available to the user. The MoSGrid project (Molecular Simulation Grid) has developed a workflow-enabled science gateway based on Liferay and WS-PGRADE. It helps users in the complex tasks of configuration and performance of molecular simulations on DCIs. In order to improve the user experience, the login process is being optimized. Currently, a Centralised Authentication Service (CAS) automatically authenticates users according to their X.509 certificate stored in the web browser. To significantly improve the usability, works are underway to authorize users based on their membership in a VO. Thus, users are automatically offered suitable features and DCIs.

Primary authors: GRUNZKE, Richard; GESING, Sandra (University of Tuebingen); SCHLEMMER, Tobias (University of Tübingen)

Co-authors: BIRKENHEUER, Georg; Dr KRÜGER, Jens; Prof. KOHLBACHER, Oliver; Dr MUELLER-PF-EFFERKORN, Ralph (Technische Universität Dresden)

Presenters: GRUNZKE, Richard; GESING, Sandra (University of Tuebingen); SCHLEMMER, Tobias (University of Tübingen)

Session Classification: AAI Workshop

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 126

Type: **Presentation**

Using online CAs and MyProxy servers for identity mapping

Wednesday, September 19, 2012 11:30 AM (15 minutes)

Description of the work

.

Wider impact of this work

.

Printable Summary

Since one of the major obstacles that very often prevents new users from using the grid infrastructure is the AA complexity, IGI has developed an access portal based on an external authentication layer providing strong user identification by means of an accredited identity federation. It is interfaced to an online CA, which provides users with digital certificates in a transparent way; a long term proxy is then created and saved into a MyProxy server in an encrypted form. Through the portal users can then submit either simple jobs or complex workflows to the grid without the need for directly managing their X509 certificates. A prototype of the portal is operational and used by some Italian VOs but the number of users is expected to increase in the coming months.

Primary authors: GAIDO, Luciano (INFN); BENCIVENNI, Marco (INFN)

Co-author: VERONESI, Paolo (INFN)

Presenter: BENCIVENNI, Marco (INFN)

Session Classification: AAI Workshop

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 127

Type: **Presentation**

Report on Federated Identity Management workshops

Wednesday, September 19, 2012 11:45 AM (15 minutes)

Description of the work

.

Wider impact of this work

.

Printable Summary

Driven by their common needs for federated identity management (FIM) representatives from a variety of research communities, including photon/neutron facilities, social science & humanities, high-energy physics, atmospheric science, bioinformatics and fusion energy, have come together to discuss how to address FIM-related issues with the objective to define a common policy and trust framework for Identity Management based on existing structures, federations and technologies. Four workshops have been held since the summer of 2011 with the involvement of these communities, various technology and community projects and initiatives. The set of requirements, recommendations and plans that emerged through these workshops will be presented in the talk.

Primary author: KELSEY, David (STFC)

Presenter: KELSEY, David (STFC)

Session Classification: AAI Workshop

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 128

Type: **Presentation**

Identity Developments in the NRENs community

Wednesday, September 19, 2012 12:00 PM (15 minutes)

Printable Summary

This talk will report on the latest developments in the identity space carried out within the NRENs community. The talk will highlight areas where synergy between different communities is important and will report on any initiative aimed to strengthen cross-communities collaboration.

Primary authors: SCHOFIELD, Brook (TERENA); FLORIO, Licia (TERENA)

Presenter: FLORIO, Licia (TERENA)

Session Classification: AAI Workshop

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 129

Type: **Presentation**

Integrating non-web based services with identity federations

Wednesday, September 19, 2012 3:10 PM (20 minutes)

Description of the work

Additionally, it is preferable to hook non web-based services into (web-based) federations that have already been established. In this talk we present FACIUS (Federated Access Control Integration for Universal Services), an approach to integrate non web-based services into widely deployed SAML federations. The focus of our concept lies on deployability, i.e., the minimization of interference with existing service deployments, the perpetuation of usability and the immutability of already deployed SAML identity providers. The approach has already been successfully applied to federate SSH access to cluster/grid resources based on a SAML federation consisting of nine universities of the state of Baden-Württemberg (Germany) in the context of the bwIDM project.

Wider impact of this work

.

Printable Summary

Federated authentication and authorisation infrastructures (AAIs) for web-based services are already well established and yield advantages such as enhanced usability and improved quality of identity information. Non web-based services such as grid resources could benefit from federated identity management in a similar way web-based services do. However, web-based approaches such as the Security Assertion Markup Language (SAML) commonly rely on the versatile user interface as well as the HTTP protocol implementation of web browsers and are thus not seamlessly applicable for non web-based services. Furthermore, a clean slate approach to build federation technologies for non web-based services is not viable, as it has to be taken into account that most services to be federated are already deployed and operational.

Primary author: KOEHLER, Jens**Presenter:** KOEHLER, Jens**Session Classification:** AAI Workshop**Track Classification:** Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 130

Type: **Presentation**

Manifesto for Secure AdHoc Data Sharing

Tuesday, September 18, 2012 11:44 AM (22 minutes)

Description of the work

We should all go back to the drawing board! The manifesto for secure data sharing is this: 1. Free the ordinary users! Let them decide to share. 2. User-friendly data sharing interface should not require IT skills. 3. Data sharing must be easy, efficient and take seconds. In our talk we will introduce you to AdHoc. AdHoc allows to share resources securely, easily and quickly, based on the Grid security control. It is integrated with Globus and VOMS. And it implements the Manifesto.

Wider impact of this work

.

Printable Summary

A decade ago, the idea of the Virtual Organizations (VO) was born: dynamic groups of mutual trust, that people (or processes) could set up on the fly. Soon, the first Virtual Organizations were implemented... or were they really? Nope. Bureaucracy won. In corporations and universities alike, there is one security bottleneck: the human administrator with his procedures. On the pretext of technology requirements, in the fear of responsibility, we have killed productivity and forgotten the premise of on-the-fly, secure yet easy sharing.

Primary author: HELLER, Helmut (BADW)

Co-authors: FRANK, Anton (BADW); PLASZCZAK, Pawel (GridwiseTech); CZECHOWSKI, Sebastian (GridwiseTech)

Presenter: PLASZCZAK, Pawel (GridwiseTech)

Session Classification: Research Infrastructures

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 131

Type: **Presentation**

Providing short-lived X.509 certificates with the GridCertLib Java library

Wednesday, September 19, 2012 2:00 PM (15 minutes)

Description of the work

..

Wider impact of this work

.

Printable Summary

GridCertLib is a Java library leveraging a Shibboleth-based authentication infrastructure and the SLCS online certificate signing service, to provide short-lived X.509 certificates and Grid proxies. The main use case envisioned for GridCertLib, is to provide seamless and secure access to Grid/X.509 certificates and proxies in web applications and portals: when a user logs in to the portal using Shibboleth authentication, GridCertLib can automatically obtain a Grid/X.509 certificate from the SLCS service and generate a VOMS proxy from it. We give an overview of the architecture of GridCertLib and briefly describe its programming model. Its application to some deployment scenarios is outlined, as well as a report on practical experience integrating GridCertLib into portals for Bioinformatics and Computational Chemistry applications, based on the popular P-GRADE and Django software.

Primary author: MAFFIOLETTI, sergio (UZH)

Presenter: MAFFIOLETTI, sergio (UZH)

Session Classification: AAI Workshop

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 132

Type: **Presentation**

Authorization in Grids using an Electronic ID Card

Wednesday, September 19, 2012 2:30 PM (20 minutes)

Printable Summary

Two years ago Germany was the 10th European country to introduce an electronic ID card. Apart from the usual information on the surface the ID card contains an integrated chip carrying information usable for online authentication and digital signatures. This information is meant to simplify the use of eGovernment services and should therefore be capable to provide a better usability of Grid services as well. In this talk we will outline possible fields of application and give a short overview over the technologies involved. Finally we will discuss issues that may influence a quick uptake.

Primary author: FRANK, Anton (BADW-LRZ)

Co-authors: METZGER, Stefan (BADW-LRZ); WEBER, Tom (University Munich); HOMMEL, Wolfgang (BADW-LRZ)

Presenter: FRANK, Anton (BADW-LRZ)

Session Classification: AAI Workshop

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 133

Type: **Presentation**

Increasing the attractiveness of EGI for computational chemistry

Tuesday, September 18, 2012 3:00 PM (15 minutes)

Description of the work

Virtual Organisations that the EGI community offers in the 'Computational Chemistry and Material Sciences' domain; Identifying key scientific applications that could benefit from resources and services of the National Grid Infrastructures, and assists the integration of these applications with these services through focused technical projects. The work aimed at increasing the number of EGI beneficiaries from the computational chemistry and material sciences domain by improving the portfolio of services that EGI provides for researchers and application developers in this field. The talk will present the achievements of the seconded application expert and the lessons learnt by EGI from this work.

Wider impact of this work

.

Printable Summary

EGI.eu as an organisation that coordinates the work of the European Grid Infrastructure community is involved in a number of projects. These projects require on occasion skills and effort beyond which EGI.eu has available internally. To respond to these needs, and allow the exchange of experiences between EGI.eu and NGI/EIRO staff, EGI.eu established a secondment programme that allows technical staff from organisations affiliated to EGI.eu as participants or associated participants, to work at EGI.eu on these projects. An application expert from the Lithuanian NGI, with scientific background in computational chemistry has been working at EGI.eu for three months in 2012 in the secondment programme. Her task involved the review of existing applications, tools and

Primary author: TAMULIENE, Jelena (EGIEU)

Presenter: TAMULIENE, Jelena (EGIEU)

Session Classification: Experience in community building

Track Classification: Community and Co-ordination (Sergio Andreozzi: track leader)

Contribution ID: 135

Type: **Presentation**

Grid Oversight in Service Level Agreement environment.

Thursday, September 20, 2012 4:20 PM (20 minutes)

Description of the work

The analysis of enhancing Service Level Management framework towards Service Level Agreements is done from the perspective of Central Operator on Duty work including changes to operations procedures & tools.

Wider impact of this work

The paper may be interesting to Operations Management and anyone involved in current Grid Oversight structure, namely members of Central Operator on Duty and Regional Operator on Duty teams. It highlights some thoughts for the future of European Grid Infrastructure.

Printable Summary

The role of Grid Oversight activity is currently to ensure the infrastructure presents expected level of performance with respect to operations activities. This includes handling incidents in a timely manner, achieving service availability/reliability thresholds. Two of new metrics included in Operations Level Agreement framework are also monitored by Grid Oversight - namely Top Level BDII availability and efficiency of Regional Operator on Duty team. It seems a natural consequence on the way of increasing the European Grid Infrastructure maturity that Service Level Agreement will be developed. This paper evaluates impact and role of Grid Oversight in such environment.

Primary authors: KRAKOWIAN, Malgorzata (CYFRONET); RADECKI, Marcin (CYFRONET)

Presenter: RADECKI, Marcin (CYFRONET)

Session Classification: Operations

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 136

Type: **Presentation**

"One size does not fit all" or making the grid more versatile

Thursday, September 20, 2012 4:00 PM (20 minutes)

Description of the work

Over the last decade or so since the start of the DataGrid project we have been able to build distributed infrastructure that offers a fairly good quality of service to its users. However, in its present state, the infrastructure is very much tailored to very specific use cases, middleware and deployed services. We envisage that in order to attract more user communities, the infrastructure should become more versatile. Efforts in this direction are underway in terms of incorporating other middlewares like Globus and UNICORE into the infrastructure and at the moment people are also looking at integrating Cloud resources.

Since the infrastructure appears to be rather rigid, this process is rather time consuming and take a considerable amount of effort. Separate integration taskforces are formed to deal with all kinds of issues related to operations, security, accounting etc. that play a role in incorporating new middleware and new technologies into the infrastructure.

We envisage that in the future, EGI will focus on treating change as a "business as usual". This meaning that that it becomes a day to day routine to incorporate new services etc. without the need for creating taskforces to guide this process. EGI should facilitate any middleware or type of (sub)infrastructure a user community or NGIs on behalf of user communities desire by providing the glue needed to tie together the multinational/multi-NGI resources in order to enable international cooperation between individual users, small user groups and organized multinational user communities. This glue being the operational infrastructure containing the operational and security procedures and policies, security models, operational tools, user support and the development of these, accounting, grid oversight etc. EGI should become more software/infrastructure/technology agnostic and should provide a generic platform for collaboration

Wider impact of this work

The presentation may be of interest to Operations Management

Printable Summary

Over the last decade or so since the start of the DataGrid project we have been able to build distributed infrastructure that offers a fairly good quality of service to its users. However, in its present state, the infrastructure is very much tailored to very specific use cases and very monolithic and rigid in its structure. It is very well suitable for a particular kind of applications but not for others. In order to attract more user communities, we believe the infrastructure should become more versatile and it should be easier to incorporate new middlewares or infrastructures. In this presentation we discuss ideas on how EGI can facilitate any middleware or type of (sub)infrastructure a user community or NGIs on behalf of user communities desire by providing the glue needed to

tie together the multinational/multi-NGI resources in order to enable international cooperation between individual users, small user groups and organized multinational user communities.

Primary author: TROMPERT, Ron (SARA)

Presenter: TROMPERT, Ron (SARA)

Session Classification: Operations

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 137

Type: **Presentation**

Globus Tools Market

Tuesday, September 18, 2012 4:40 PM (10 minutes)

Description of the work

.

Wider impact of this work

.

Printable Summary

.

Presenter: Mr HOFMANN, Matthias (TU-Dortmund)

Session Classification: Software services for community building and support

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: **139**

Type: **not specified**

Introductory remarks

Presenter: Dr FERRARI, Tiziana (EGLEU)

Contribution ID: 140

Type: **not specified**

GGUS/PRACE RT integration: preliminary tests and future steps

Tuesday, September 18, 2012 2:10 PM (20 minutes)

Presenter: MORGOTTI, Marcello (CINECA)

Session Classification: Operations -EGI/EUDAT/PRACE workshop

Contribution ID: **141**

Type: **not specified**

User accounting integration

Tuesday, September 18, 2012 2:30 PM (20 minutes)

Presenter: Dr GORDON, John (STFC)

Session Classification: Operations -EGI/EUDAT/PRACE workshop

Contribution ID: 142

Type: **not specified**

EUDAT operations - perspective and initial achievements

Tuesday, September 18, 2012 2:50 PM (40 minutes)

Printable Summary

This contribution will present the landscape of the envisaged Collaborative Data Infrastructure of EUDAT, which includes community centres, data centres and communities with data challenges.

The presentation will also provide information about the planned operations tasks (e.g. user community support, user level documentation, 1st level support, helpdesk, registration of communities and centres, identity and access management (IAM), and about EUDAT service provisioning.

Presenter: Dr REETZ, Johannes (Garching Computing Centre of the Max Planck Society / MPI for Plasma Physics)

Session Classification: Operations -EGI/EUDAT/PRACE workshop

Contribution ID: 143

Type: **not specified**

EGI/PRACE data management use cases (TBC)

Printable Summary

Analysis of VERCE/DRIHM/LifeWatch data transfer use cases

Contribution ID: 144

Type: **not specified**

Introduction

Thursday, September 20, 2012 2:00 PM (5 minutes)

Presenter: ANDREOZZI, Sergio (EGIEU)

Session Classification: Sustainability of Technology Providers

Contribution ID: 145

Type: **Presentation**

The sustainability plan of VisIVO

Thursday, September 20, 2012 2:45 PM (15 minutes)

Description of the work

VisIVO consists of several components: VisIVO Desktop, VisIVO Server and VisIVO Science Gateway. VisIVO supports high-performance, multidimensional visualization of very large-scale astrophysical datasets. Users can obtain meaningful visualizations rapidly while preserving full and intuitive control of the relevant visualization parameters. VisIVO was designed and implemented through subsequent phases. Recent developments of VisIVO took place in the context of EGI and IGI, the Italian NGI, to make VisIVO fully-compliant with distributed e-Infrastructures and capable to visualize data produced by jobs running on DCIs of different nature. The first design and development of VisIVO took place at INAF –Astrophysical Observatory of Catania; here the tool continues to be maintained and further developed so that their capabilities continue to grow. To this end the VisIVO Project Lead group operating at the Catania Observatory established several collaborations over time, in the past with CINECA, the most important supercomputing Centre in Italy and now with EGI and IGI. The kernel of the Project Lead group is constituted by staff personnel, strongly motivated to maintain VisIVO fully efficient and compatible with new emerging technologies. This model, based on a group of permanent staff supported by both external personnel and non-permanent staff hired through new collaborations and well-targeted projects could be the right way to ensure the long term sustainability of tools and services on which e-Infrastructures are based. The tool moreover is maintained on sourceforge, and collaborative co-development is welcome as highlighted in the governance model document of the project currently under preparation. The complete code documentation is given and the C++ class structure easily allow any programmer to write a new class or to add features to an existing class. Some developer notes are provided to allow the programmer to facilitate the code development.

Wider impact of this work

The long term sustainability of tools and services have a strong impact on research. Tools and service brought into production must be maintained and kept up to date beyond the end of the project that delivered them. The development and maintenance of VisIVO do not depend solely on funds of a specific project but on permanent INAF resources (especially human) and on a well-motivated base of end users that rely on it to achieve their scientific goals. The long term sustainability model put in place for VisIVO could be taken into consideration by other groups and Institutes being in charge of tools and services on which European DCIs depend. The INAF Astrophysical Observatory of Catania (OACT) gives full support to this activity. A project between INAF-OACT and the University of Portsmouth will be signed in the next few months aiming at establishing a long term common development plan for data visualization and exploration which includes the usage of workflows and Science Gateways.

Printable Summary

VisIVO is an integrated suite of tools and services specifically designed for the Virtual Observatory. This suite constitutes a software framework for an effective visual discovery in currently available (and next-generation) very large-scale astrophysical datasets. One of the components of this suite, namely VisIVO Server, has been recently rewritten to make it fully compliant with the emerging distributed e-Infrastructures. This work is part of the planned activity for task TSA3.5 of the EGI-InSPIRE project. As for any other tool and service developed in the context of a specific project, funded through the EU FP7 channel or through other channels, the problem related to the long-term sustainability, i.e. the ability of the tool or service to survive beyond the end of the funding project is of crucial importance. This talk shows how this issue has been approached and solved for VisIVO so that end users can now use the tool being confident that its long term maintenance is ensured.

Presenter: BECCIANI, UGO (INAF)

Session Classification: Sustainability of Technology Providers

Track Classification: Community and Co-ordination (Sergio Andreozzi: track leader)

Contribution ID: 146

Type: **not specified**

Common strategy for Community-Building tools

Thursday, September 20, 2012 3:00 PM (10 minutes)

Presenter: Dr SIPOS, Gergely (EGLEU)

Session Classification: Sustainability of Technology Providers

Contribution ID: 147

Type: **not specified**

Q&A / Discussion

Thursday, September 20, 2012 3:10 PM (20 minutes)

Session Classification: Sustainability of Technology Providers

Contribution ID: 148

Type: **Presentation**

The sustainability plan of VisIVO

Description of the work

VisIVO consists of several components: VisiVO Desktop, VisIVO Server and VisIVO Science Gateway. VisIVO supports high-performance, multidimensional visualization of very large-scale astrophysical datasets. Users can obtain meaningful visualizations rapidly while preserving full and intuitive control of the relevant visualization parameters. VisIVO was designed and implemented through subsequent phases. Recent developments of VisIVO took place in the context of EGI and IGI, the Italian NGL, to make VisIVO fully-compliant with distributed e-Infrastructures and capable to visualize data produced by jobs running on DCIs of different nature. The first design and development of VisIVO took place at INAF –Astrophysical Observatory of Catania; here the tool continues to be maintained and further developed so that their capabilities continue to grow. To this end the VisIVO Project Lead group operating at the Catania Observatory established several collaborations over time, in the past with CINECA, the most important supercomputing Centre in Italy and now with EGI and IGI. The kernel of the Project Lead group is constituted by staff personnel, strongly motivated to maintain VisIVO fully efficient and compatible with new emerging technologies. This model, based on a group of permanent staff supported by both external personnel and non-permanent staff hired through new collaborations and well-targeted projects could be the right way to ensure the long term sustainability of tools and services on which e-Infrastructures are based. The tool moreover is maintained on sourceforge, and collaborative co-development is welcome as highlighted in the governance model document of the project currently under preparation. The complete code documentation is given and the C++ class structure easily allow any programmer to write a new class or to add features to an existing class. Some developer notes are provided to allow the programmer to facilitate the code development.

Wider impact of this work

The long term sustainability of tools and services have a strong impact on research. Tools and service brought into production must be maintained and kept up to date beyond the end of the project that delivered them. The development and maintenance of VisIVO do not depend solely on funds of a specific project but on permanent INAF resources (especially human) and on a well-motivated base of end users that rely on it to achieve their scientific goals. The long term sustainability model put in place for VisIVO could be taken into consideration by other groups and Institutes being in charge of tools and services on which European DCIs depend. The INAF Astrophysical Observatory of Catania (OACT) gives full support to this activity. A project between INAF-OACT and the University of Portsmouth will be signed in the next few months aiming at establishing a long term common development plan for data visualization and exploration which includes the usage of workflows and Science Gateways.

Printable Summary

VisIVO is an integrated suite of tools and services specifically designed for the Virtual Observatory. This suite constitutes a software framework for an effective visual discovery in currently available

(and next-generation) very large-scale astrophysical datasets. One of the components of this suite, namely VisIVO Server, has been recently rewritten to make it fully compliant with the emerging distributed e-Infrastructures. This work is part of the planned activity for task TSA3.5 of the EGI-InSPIRE project. As for any other tool and service developed in the context of a specific project, funded through the EU FP7 channel or through other channels, the problem related to the long-term sustainability, i.e. the ability of the tool or service to survive beyond the end of the funding project is of crucial importance. This talk shows how this issue has been approached and solved for VisIVO so that end users can now use the tool being confident that its long term maintenance is ensured.

Primary author: BECCIANI, UGO (INAF)

Presenter: BECCIANI, UGO (INAF)

Track Classification: Community and Co-ordination (Sergio Andreozzi: track leader)

Contribution ID: 149

Type: **Demonstration**

VisIVO Software and gLite: A Set of Tools for Large Scale Dataset Exploration

Description of the work

VisIVO tools are developed by INAF and by the University of Portsmouth (UK).

VisIVO consists of several tools: a) VisIVO Desktop - a stand-alone application for interactive visualization on standard PCs; b) VisIVO Server –a command line set of programs for high performance visualization; c) VisIVO Science Gateway –a WS-PGRADE gateway supporting services based on the VisIVO Server functionality, developed in SCI-BUS project; d) VisIVO Mobile, an iOS application that allows the user to execute VisIVO on mobile devices; e) VisIVO Library to produce a set of images or movies directly using VisIVO with its internal data arrays without the need to produce intermediate files.

All these tools are based on three main components: Importer, Filter and Viewer. VisIVO Importer converts user-supplied datasets into VBTs (VisIVO Binary Table), the internal data format of VisIVO. VBTs are employed by VisIVO Filter modules for data exploration and by VisIVO Viewer for the final data rendering.

Through VisIVO, users can explore and display large datasets and generate movies; they can rapidly obtain meaningful visualizations while preserving full and intuitive control of the relevant visualization parameters.

The development of VisIVO is included in several EU funded international projects: a) EGI-InSPIRE, to complete the VisIVO porting on gLite middleware; b) EDGI where the VisIVO porting on Desktop Grids is completed and the validation phase is in progress; c) SCI-BUS to develop the new VisIVO Science Gateway and the VisIVO Mobile application; d) Muon Portal: a national project funded by the Italian Ministry of Research using EU funds, for the inspection of standard containers using the research conducted in Astrophysical and in Nuclear Physics (<http://muoni.oact.inaf.it:8080>). Moreover a new project to conduct a pilot experiment for the usage of VisIVO in the visitors' centres using the VisIVO Mobile application is going to start.

Link for further information

<http://visivo-server.oact.inaf.it:8080>

Wider impact of this work

VisIVO was initially developed to simulate the Large Scale Structure of the Universe and the exploration of related data. It is particularly suitable in exploring N-Body simulations and many physical problems can be described using particles.

The simplified usage of VisIVO combined with its different components allow user to explore complex datasets and make it attractive for all users working with particles and volumes to describe the physical phenomena. The development of the VisIVO Library moreover offers new opportunities to the user, especially for the investigation of possible problems at run time, by checking the current status of the simulation trying to understand if everything is correct.

All this explains the wide potential impact of VisIVO as it is a tool suitable for many disciplines and also for common citizens.

The future work aims at fully porting VisIVO Server on Grid environments enhancing significantly in this way the performances of the tool.

Printable Summary

Modern scientific research needs to exploit computer graphics and scientific visualization tools to appropriately display datasets and thus allow scientists to perform efficient visual discovery.

VisIVO is a suite of software tools aimed at creating 3D customized views of many types of data with no limits imposed on the data size. The software allows datasets to be explored and the creation of images and movies starting from files provided by users. VisIVO was originally written to visualize astronomical data but can now be used in many other scientific fields including chemistry, nuclear physics and biomedical.

In this demonstration VisIVO will be used to explore a complex dataset and in particular, VisIVO Server as a set of programs will produce images and movies directly from complex datasets. We will also demonstrate the features offered by the VisIVO Science Gateway and by the Mobile application. Finally, a simple use of the VisIVO Library will be presented.

Primary author: BECCIANI, UGO (INAF)

Presenter: BECCIANI, UGO (INAF)

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 150

Type: **not specified**

Introduction and Goals

Tuesday, September 18, 2012 4:00 PM (10 minutes)

Description of the work

The ultimate goal is to improve the evidence of the EGI scientific impact. This session will describe the proposed processes, policies and tools to be established.

Wider impact of this work

It is aimed at NGIs/VOs representatives and all people interested in demonstrating the scientific impact of their research or infrastructure and will give the opportunity to provide feedback to the defined plan before starting the implementation.

Printable Summary

The Scientific Publications Repository project started in June 2012 in order to define tools, policies and processes to track the scientific outcome that was possible thanks to the usage of the European Grid Infrastructure (EGI). The ultimate goal is to improve the evidence of the EGI scientific impact. This session will describe the proposed processes, policies and tools to be established. It is aimed at NGIs/VOs representatives and all people interested in demonstrating the scientific impact of their research or infrastructure and will give the opportunity to provide feedback to the defined plan before starting the implementation.

Presenter: ANDREOZZI, Sergio (EGLEU)

Session Classification: EGI Scientific Publications Repository workshop

Contribution ID: 151

Type: **not specified**

EC Policy on Open Access and the OpenAIRE Initiative

Tuesday, September 18, 2012 4:10 PM (30 minutes)

Primary author: PAGANO, Pasquale (CNR - ISTI)

Presenter: PAGANO, Pasquale (CNR - ISTI)

Session Classification: EGI Scientific Publications Repository workshop

Contribution ID: 152

Type: **not specified**

VT Activity Report and Proposed Workplan for EGI

Tuesday, September 18, 2012 4:40 PM (20 minutes)

Primary author: ANDREOZZI, Sergio (EGLEU)

Presenter: ANDREOZZI, Sergio (EGLEU)

Session Classification: EGI Scientific Publications Repository workshop

Contribution ID: **153**

Type: **not specified**

Interactive Discussion

Tuesday, September 18, 2012 5:00 PM (30 minutes)

Session Classification: EGI Scientific Publications Repository workshop

Contribution ID: **154**

Type: **not specified**

Talk 2

Contribution ID: 155

Type: **not specified**

MAPPER status and data management use cases

Tuesday, September 18, 2012 4:00 PM (20 minutes)

Presenter: SAVERCHENKO, Ilya (BADW)

Session Classification: Operations -EGI/EUDAT/PRACE workshop

Contribution ID: 156

Type: **not specified**

Virtual Earthquake and seismology (VERCE): data management use case

Tuesday, September 18, 2012 4:20 PM (20 minutes)

Primary author: SCHWICHTENBERG, Horst (FRAUNHOFER)

Presenter: SCHWICHTENBERG, Horst (FRAUNHOFER)

Session Classification: Operations -EGI/EUDAT/PRACE workshop

Contribution ID: 157

Type: **not specified**

High Performance Grid Computing for Molecular and Materials Science and Technology

Tuesday, September 18, 2012 4:40 PM (20 minutes)

Presenters: COSTANTINI, Alessandro (INFN); LAGANA, Antonio (UNIPG)

Session Classification: Operations -EGI/EUDAT/PRACE workshop

Contribution ID: 158

Type: **not specified**

Globus Online for HTC/HPCe: use cases and security aspects

Tuesday, September 18, 2012 5:00 PM (20 minutes)

Presenter: Mr TUECKE, Steve (University of Chicago / Argonne National Laboratory)

Session Classification: Operations -EGI/EUDAT/PRACE workshop

Contribution ID: **159**

Type: **not specified**

Conclusions and next steps

Tuesday, September 18, 2012 5:20 PM (10 minutes)

Presenter: Dr FERRARI, Tiziana (EGLEU)

Session Classification: Operations -EGI/EUDAT/PRACE workshop

Contribution ID: 160

Type: **Demonstration**

EGI Applications Database demonstration –The path towards a DCI software registry

Description of the work

One of the most prominent new features of the EGI Applications Database is the release of v1.0 of its REST API, supporting authenticated write access. Third party application providers can make use of the API by forwarding their users' EGI SSO credentials, or by creating an AppDB system account to act on behalf of their users, in order to modify content and to read content that is not open to the general public. Developments related to the REST API's write access include amongst others:

- API key management feature from within the portal, under user preferences
- Documentation, complete with examples and sample use cases [R1]

Other significant developments that took place in the last 6 months and will be demonstrated relate to the curation and quality control of information stored in the database. The new services in this area include:

- A broken link notification subsystem, which automatically sends out e-mail notifications and reminders to application owners if a link referenced from the profile of their application is not reachable
- A mechanism that identifies application entries that have not been updated recently and may become obsolete.
- A mechanism to enable the community classify software entries into different high level categories. Besides the already known 'scientific application' and 'software developer tool' categories new categories, such as middleware service, operational tool, workflow system, etc. can be also added.
- A dissemination tool, which allows managers to create and send out informative e-mail messages, to lists of users based on criteria of their choice, such as a specific NGI, discipline, etc.

Link for further information

<http://appdb.egi.eu>

Wider impact of this work

The main aim of the demo session is to assist research leaders, scientists and software developers to gain a better understand of:

- How they can make best use of the AppDB service to find new users, to keep in touch with the existing users and, to gather feedback and satisfaction information from users about the software.
- What are the different interfaces and user roles one could use the system with.
- What are the most significant features developed by the AppDB development team during the period from the EGI UF 2012 up till the present.

Printable Summary

The EGI Applications Database (AppDB) is a service that retrieves information about scientific applications, developer tools and about the programmers and scientists who developed and use them. This demonstration presents the latest developments of the system, the features that enable the database to grow to a registry and platform used by all software developers and consumers in EGI. The demo is aimed at software developers from scientific communities, National Grid Infrastructures, middleware groups, operational tool providers, and anyone working on software for distributed computing.

The following AppDB features will be demonstrated:

- Federating databases and portals
 - AppDB write API principles and notions
 - Authentication methods
- Notifications and feedback
 - RSS feeds and e-mails, customised mass-messaging
 - Commenting and rating
- Embedding AppDB third party websites
 - Using the gadget editor
 - Deploying a Gadget instance

Primary authors: NAKOS, Alexander; Dr SIPOS, Gergely (EGLEU); CHATZIANGELOU, Marios (IASA); VASSILIS KARAGEORGOS, William (GRNET)

Presenters: NAKOS, Alexander; Dr SIPOS, Gergely (EGLEU); CHATZIANGELOU, Marios (IASA); VASSILIS KARAGEORGOS, William (GRNET)

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: **161**

Type: **Demonstration**

Training Marketplace

Description of the work

text to come

Wider impact of this work

text to come

Printable Summary

text to come

Primary author: DEVEREUX, Claire (STFC)

Presenter: DEVEREUX, Claire (STFC)

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 162

Type: **not specified**

GPGPU status of deployment and user requirements

Wednesday, September 19, 2012 11:20 AM (30 minutes)

Description of the work

GPGPU surveys:

User survey: https://wiki.egi.eu/wiki/VT_GPGPU#GPGPU_User_Community

Resource provider survey: https://wiki.egi.eu/wiki/VT_GPGPU#GPGPU_Resource_Centre

Wider impact of this work

Information collected through the survey will be discussed to define a EGI technical roadmap. Information will be also aggregated at a national level and provided to NGI representatives to allow the definition of NGI specific action plans.

Printable Summary

This contribution will provide information about the current deployment status of GPGPU technologies, the deployment issues faced by administrators, and the user requirements. Information presented will be extracted from the surveys launched in August by the GPGPU Virtual Team.

The purpose of this contribution is to understand how the users currently use, or intend to use, GPGPUs or other accelerated devices in a distributed environment such as grids or clouds, or whether they shall use grid/cloud resources at all for this purpose. In particular, we would like to determine if there is a specific need from European research collaborations and their partners beyond Europe for a more tightly integrated GPGPU capability within the European Grid Infrastructure.

Presenter: WALSH, John (TCD)

Session Classification: Operations - Resource Centre Forum

Contribution ID: **163**

Type: **not specified**

Introduction and Key Findings

Friday, September 21, 2012 11:00 AM (30 minutes)

Presenters: Dr KARAGIANNIS, Fotis (Independent); Prof. COHEN, Sandra (Athens University of Economics and Business)

Session Classification: e-Fiscal - Cost Analysis of the European Computing e-Infrastructures

Contribution ID: **164**

Type: **not specified**

Benchmarking Results: EGI & HPC vs. Amazon EC2

Friday, September 21, 2012 11:30 AM (20 minutes)

Presenter: Dr IQBAL, Kashif (ICHEC)

Session Classification: e-Fiscal - Cost Analysis of the European Computing e-Infrastructures

Contribution ID: 165

Type: **not specified**

Reflections on the Experience and Way Forward

Friday, September 21, 2012 11:50 AM (20 minutes)

Presenter: Mr HEIKKURINEN, Matti (Emergence Tech Ltd)

Session Classification: e-Fiscal - Cost Analysis of the European Computing e-Infrastructures

Contribution ID: **166**

Type: **not specified**

Q&A

Friday, September 21, 2012 12:10 PM (20 minutes)

Session Classification: e-Fiscal - Cost Analysis of the European Computing e-Infrastructures

Contribution ID: **167**

Type: **not specified**

Welcome and Overview

Presenter: Mrs NARDI, Nadia (ENG)

Contribution ID: 168

Type: **not specified**

Resource allocation policies in EGI

Wednesday, September 19, 2012 11:50 AM (20 minutes)

Description of the work

Operations Management Board survey on resource allocation:

https://wiki.egi.eu/wiki/EGI_Operations_Surveys#Resource_Allocation_Policies

Printable Summary

Policies for resource allocation in EGI vary greatly at a national level and Resource Centre level. However, common processes are needed to meet offer with the demand of international research collaborations, in particular to support opportunistic usage patterns adopted by various international Virtual Collaborations.

We will review the status of policies according to the input collected from NGI representatives, and we will discuss problems to be addressed.

Presenters: Mr SZEPIENIEC, T. (CYFRONET); Dr FERRARI, Tiziana (EGLEU)

Session Classification: Operations - Resource Centre Forum

Contribution ID: **169**

Type: **not specified**

Resource Centre cooperation: interest groups and ideas of innovation towards Horizon 2020

Wednesday, September 19, 2012 12:10 PM (20 minutes)

Session Classification: Operations - Resource Centre Forum

Contribution ID: 170

Type: **not specified**

Integrating OSG Operational Services

Thursday, September 20, 2012 4:40 PM (20 minutes)

Printable Summary

This contribution will address experiences with the CA switch and integrating it with OSG existing Operational Services (Topology DB, Monitoring Services, Wiki's, Ticketing, etc.), and will provide information about the OSG plan towards the adoption of SHA-2 in its PKI.

Presenter: QUICK, Rob (Open Science Grid)

Session Classification: Operations

Contribution ID: 171

Type: **not specified**

Status and perspectives of operations in the Asia Pacific region

Thursday, September 20, 2012 5:00 PM (20 minutes)

Presenter: YEN, Eric (ASGC)

Session Classification: Operations

Contribution ID: 172

Type: **not specified**

Future directions of EGI operations

Thursday, September 20, 2012 2:45 PM (45 minutes)

Presenter: Dr FERRARI, Tiziana (EGLEU)

Session Classification: Operations

Contribution ID: 173

Type: **not specified**

Sustainability of NGI operations and assessment of EGI global operations services

Thursday, September 20, 2012 2:05 PM (40 minutes)

Description of the work

Link to sustainability survey: https://wiki.egi.eu/wiki/EGI_Operations_Surveys#Sustainability_of_the_EGI.2FNNGIs_operati

Printable Summary

We will analyze the current perspectives and sustainability of NGI operations after April 2014 and the status of EGI global operations services according to the input gathered from NGIs through a survey on operations sustainability.

We will also discuss ideas of innovation for operations towards Horizon 2020.

Presenter: Dr FERRARI, Tiziana (EGIEU)

Session Classification: Operations

Contribution ID: 174

Type: **not specified**

Concluding remarks

Thursday, September 20, 2012 5:20 PM (10 minutes)

Session Classification: Operations

Contribution ID: 175

Type: **not specified**

Speech from EC Project Officer

Contribution ID: 176

Type: **not specified**

ERINA+ Methodology

Wednesday, September 19, 2012 11:10 AM (30 minutes)

Presenter: BELLINI, Francesco (Eurokleis)

Session Classification: e-Infrastructure Impact Assessment Methodologies

Contribution ID: 177

Type: **not specified**

ERINA+ Tools for Data Gathering

Wednesday, September 19, 2012 11:40 AM (35 minutes)

Presenters: BELLINI, Francesco (Eurokleis); BENEDIKT, Josef (ZSI); Dr PASSANI, antonella (T6 ecosystems)

Session Classification: e-Infrastructure Impact Assessment Methodologies

Contribution ID: 178

Type: **not specified**

Q&A / Open Discussion

Wednesday, September 19, 2012 12:15 PM (15 minutes)

Session Classification: e-Infrastructure Impact Assessment Methodologies

Contribution ID: 179

Type: **not specified**

Welcome message

Tuesday, September 18, 2012 9:00 AM (10 minutes)

Presenter: MATYSKA, Ludek (CESNET)

Session Classification: Plenary

Contribution ID: **180**

Type: **not specified**

IBM Platform Computing

Tuesday, September 18, 2012 9:10 AM (15 minutes)

Presenter: PARKS, Leah

Session Classification: Plenary

Contribution ID: **181**

Type: **not specified**

EU policy developments related to e-infrastructures

Tuesday, September 18, 2012 9:25 AM (15 minutes)

Presenter: GLINOS, Kostas

Session Classification: Plenary

Contribution ID: **182**

Type: **not specified**

EGI on the road to Horizon 2020

Tuesday, September 18, 2012 9:40 AM (40 minutes)

Printable Summary

As EGI-InSPIRE approaches its last 18 months, the presentation will reflect on the developments of the last 6 months and examine how they will be shaping our activity in Horizon 2020 and beyond. The endorsement of the EGI 2020 strategy by the EGI Council provides a vision for the EGI community that extends beyond a single project and embraces the need for coordinated community activity that is being initiated through activities taking place this week.

Presenter: NEWHOUSE, Steven (EGLEU)

Session Classification: Plenary

Contribution ID: **183**

Type: **not specified**

One year of The EGI Federated Clouds Task Force

Wednesday, September 19, 2012 9:00 AM (45 minutes)

Presenter: Dr TURILLI, Matteo (Oxford e-Research Centre)

Session Classification: Plenary

Contribution ID: **184**

Type: **not specified**

Cloud-based services for Science

Wednesday, September 19, 2012 9:45 AM (45 minutes)

Presenter: Mr TUECKE, Steve (University of Chicago / Argonne National Laboratory)

Session Classification: Plenary

Contribution ID: 185

Type: **not specified**

What virtual environment and data management for a distributed research infrastructure like ELI?

Thursday, September 20, 2012 9:00 AM (45 minutes)

Presenter: RUZICKA, Vlastimil

Session Classification: Plenary

Contribution ID: 186

Type: **not specified**

Data Challenges in EISCAT_3D -The European 3-Dimensional Imaging Radar for Atmospheric and Geospace Research

Thursday, September 20, 2012 9:45 AM (45 minutes)

Presenter: TURUNEN, Esa

Session Classification: Plenary

Contribution ID: **187**

Type: **not specified**

Clouds, what next?

Friday, September 21, 2012 9:15 AM (35 minutes)

Presenter: BAGULEY, Joe

Session Classification: Plenary

Contribution ID: **188**

Type: **not specified**

Awards & Prizes presentation

Friday, September 21, 2012 10:00 AM (15 minutes)

Session Classification: Plenary

Contribution ID: **189**

Type: **not specified**

Closing remarks

Friday, September 21, 2012 10:15 AM (15 minutes)

Presenter: NEWHOUSE, Steven (EGLEU)

Session Classification: Plenary

Contribution ID: **190**Type: **Presentation**

D-MON

Wednesday, September 19, 2012 2:30 PM (20 minutes)

Description of the work

E-Infrastructure provided by EGI offers European scientific communities state of the art ICT services that are indispensable for success of their work. The e-Infrastructure is a highly complex and sophisticated system thus, to optimise utilisation of available resources and services, user communities require access to precise, mission-specific information. Such information may comprise a list of services accessible to the respective community, forecasted availability of these services, detailed account on work progress of community users, amount of storage, computing and other resources available to the community. This information is available from multiple sources, such as monitoring, information systems, user management and accounting services. Each individual system holds a subset of the required data that often is not valuable in its raw form. For example, monitoring tools like Nagios collect resource and service availability information – however Nagios can not provide a user with a status summary about services he or she are authorised to access.

To address this challenge we developed a framework called D-MON. D-MON integrates existing information providers and applies filtering and data mining algorithms to generate mission-specific infrastructure reports and views. D-MON focuses on satisfying requirements of e-Infrastructure stakeholders, such as academic users, infrastructure operators, and project coordinators, by providing required information while filtering out unnecessary details. D-MON avoids duplicating functionality implemented in existing information services and provides an added value by federating existing information providers and utilising information available in a e-Infrastructure.

In the presentation a live demonstration of D-MON will be given.

Wider impact of this work

D-MON avoids duplicating functionality implemented in existing information services and provides an added value by federating existing information providers and utilising information available in a e-Infrastructure.

Printable Summary

The e-Infrastructure provided by EGI is a highly complex and sophisticated system thus, to optimise utilisation of available resources and services, user communities require access to precise, mission-specific information. Such information may comprise a list of services accessible to the respective community, forecasted availability of these services, detailed account on work progress of community users, amount of storage, computing and other resources available to the community. This information is available from multiple sources, such as monitoring, information systems, user management and accounting services. To address this challenge we developed a framework called D-MON. D-MON integrates existing information providers and applies filtering and data mining algorithms to generate mission-specific infrastructure reports and views.

In the presentation a live demonstration of D-MON will be given.

Primary author: SAVERCHENKO, Ilya (BADW)

Presenter: SAVERCHENKO, Ilya (BADW)

Session Classification: Operations Workshops

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 191

Type: **Poster**

Scientific Large Scale Data Management and Analysis

Wider impact of this work

Scientific Large Scale Data Management and Analysis

Printable Summary

Scientific Large Scale Data Management and Analysis

Description of the work

Scientific Large Scale Data Management and Analysis

Primary author: JUNG, Christopher (KIT-G)

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: **192**

Type: **not specified**

OSG Operational Tools

Wednesday, September 19, 2012 2:50 PM (20 minutes)

Printable Summary

This contribution will present the status of Open Science Grid operational tools with a focus on Topology DB, Monitoring Services and Ticketing.

Presenter: QUICK, Rob (Open Science Grid)

Session Classification: Operations Workshops

Contribution ID: **193**

Type: **not specified**

SAM status and plans

Wednesday, September 19, 2012 3:10 PM (20 minutes)

Presenter: BABIK, Marian (CERN)

Session Classification: Operations Workshops

Contribution ID: **194**

Type: **not specified**

GSTAT status and plans

Wednesday, September 19, 2012 2:10 PM (20 minutes)

Presenter: YEN, Eric (ASGC)

Session Classification: Operations Workshops

Contribution ID: 195

Type: **not specified**

Operations Portal future plan

Wednesday, September 19, 2012 4:30 PM (20 minutes)

Presenter: LORPHELIN, Cyril (CNRS)

Session Classification: Operations Workshops

Contribution ID: **196**

Type: **not specified**

GGUS status and plans

Wednesday, September 19, 2012 4:50 PM (20 minutes)

Presenter: Dr ANTONI, Torsten (KIT-G)

Session Classification: Operations Workshops

Contribution ID: **197**

Type: **not specified**

Discussion

Wednesday, September 19, 2012 5:10 PM (20 minutes)

Session Classification: Operations Workshops

Contribution ID: 198

Type: **Poster**

bwLSDF - distributed storage management Enhancements of the Large Scale Data Facility towards statewide storage usage within Baden-Württemberg

Description of the work

bwLSDF

Wider impact of this work

bwLSDF

Printable Summary

The bwLSDF research project investigates the advantages and drawbacks of centralized & distributed storage systems and identify the potentials and risks of federal storage usage within the state of Baden-Württemberg. Therefore, we analyze common NAS and SAN technologies regarding their limits of utility and usability in wide area networks. Besides these traditional concepts, we look into new approaches for more flexible storage access and address the emerging challenges of statewide identity management.

As a first step, we are going to introduce the bwSync&Share service, which is planned to be a privacy-aware alternative to the well-known dropbox service. Later on, we intend to focus a more general solution which is based on upcoming cloud storage paradigms including REST and CDMI.

Primary author: Mr SCHLITTER, Nico (KIT)

Presenter: Mr SCHLITTER, Nico (KIT)

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 199

Type: **Poster**

bwIDM: Federated Identity Management for the state of Baden-Württemberg (Germany)

Printable Summary

bwIDM: Federated Identity Management for the state of Baden-Württemberg (Germany)

Description of the work

bwIDM

Wider impact of this work

bwIDM

Primary author: KOEHLER, Jens (Karlsruhe Institute of Technology (KIT))

Presenter: KOEHLER, Jens (Karlsruhe Institute of Technology (KIT))

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 200

Type: **Poster**

bwGRiD - A Distributed Computing Infrastructure for Universities in Baden-Wuerttemberg

Description of the work

bwGRiD

Wider impact of this work

bwGRiD

Printable Summary

bwGRiD - A Distributed Computing Infrastructure for Universities in Baden-Wuerttemberg

Primary author: Mr HERMANN, Sven (KIT)

Presenter: Mr HERMANN, Sven (KIT)

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: **201**

Type: **not specified**

Introduction and Goals

Wednesday, September 19, 2012 11:00 AM (10 minutes)

Presenter: ANDREOZZI, Sergio (EGLEU)

Session Classification: Helix Nebula Workshop: Technical Interoperability

Contribution ID: 202

Type: **Presentation**

Use Cases for using integrated clouds infrastructures

Wednesday, September 19, 2012 11:10 AM (20 minutes)

Description of the work

Sergio Androozzi has been involved in grid computing since 2002, when he joined INFN to work on interoperability aspects. Since 2007, he has co-chaired the GLUE Working Group in OGF and contributed to several standard activities. In June 2010, he joined EGI.eu as the Strategy and Policy Manager to steer the policy development process and support strategic planning of the European Grid Infrastructure. Sergio holds a PhD in Computer Science from the University of Bologna and a MSc in Computer Science Engineering from the University of Pisa.

Printable Summary

E-Infrastructures for science have been built to support scientific collaborations and enable digital research to tackle modern grand challenges. On the other side, the commoditisation of virtualisation technologies and the emerging of cloud computing from commercial providers is becoming attractive to some scientific domains or groups. This presentation introduces a number of initial use cases that would benefit from the integration of commercial and publicly-funded infrastructure. The goal is to stimulate the discussion on technical, legal, policy and business aspects.

Presenter: ANDREOZZI, Sergio (EGIEU)

Session Classification: Helix Nebula Workshop: Technical Interoperability

Contribution ID: 203

Type: **not specified**

The Helix-Nebula Technical Architecture

Wednesday, September 19, 2012 11:50 AM (20 minutes)

Description of the work

Emmanuel Mathot has a master degree in Computer Science specialised in data-mining from Université Catholique de Louvain, Belgium. During 5 years at European Space Agency, he led the development and operations for the Grid processing on demand project of ESA (G-POD) where he worked closely with multi-disciplinary scientific communities on HPC for satellite imagery. He joined Terradue Srl 2 years ago as Technical Leader and Architect for Grid/Cloud solutions.

Printable Summary

This presentation introduces the current knowledge of the Helix Nebula Technology and Architecture Group regarding the need of a federated framework to simplify discovery, access, usage and management of a federated cloud system. Alongside this objective, we aim at providing an integration framework, where current and future suppliers (i.e. cloud service providers) can easily interface their system to attract and receive cloud workload.

Primary authors: MATHOT, Emmanuel (Terradue); BEGIN, Marc-Elian (SixSq)

Presenter: EVANS, Phil (Logica)

Session Classification: Helix Nebula Workshop: Technical Interoperability

Contribution ID: 204

Type: **not specified**

EGI FedCloud Architecture

Wednesday, September 19, 2012 11:30 AM (20 minutes)

Description of the work

Matteo Turilli is a Research Associate at the Oxford e-Research Centre. He holds a DPhil (PhD) in Computer Science from the University of Oxford and he is involved in several projects concerning the research and development of Cloud Computing. Matteo's main research interests are in parallel and distributed computing, software design, specifically as it relates to ethical requirements and the concept of design.

Presenter: Dr TURILLI, Matteo (Oxford e-Research Centre)

Session Classification: Helix Nebula Workshop: Technical Interoperability

Contribution ID: 205

Type: **not specified**

Discussion on the technical interoperability requirements

Wednesday, September 19, 2012 12:10 PM (20 minutes)

Session Classification: Helix Nebula Workshop: Technical Interoperability

Contribution ID: 206

Type: **not specified**

Introduction and Goals

Wednesday, September 19, 2012 2:00 PM (5 minutes)

Presenter: ASERO, Carmela (EGLEU)

Session Classification: Helix-Nebula Workshop: Networking Connectivity

Contribution ID: 207

Type: **Presentation**

Connecting commercial cloud providers with NRENs/GEANT

Wednesday, September 19, 2012 2:05 PM (20 minutes)

Description of the work

Niels Hersoug joined DANTE as General Manager in 2011. A Certified Project Manager and Master of Science and Business Administration, Niels is highly experienced in management and technology. Prior to DANTE, Niels worked in a number of senior management roles in a range of technology-driven companies.

Printable Summary

GÉANT indeed welcomes the Helix Nebula initiative and are prepared and able to connect the partners and by this give the initiative a kick-start into real operation. Niels will in general talk about GÉANT and also dig into how to connect some of the present Helix Nebula partners. Over its lifespan the GÉANT pan-European network has transformed European research, delivering the speed and services to enable greater collaboration and benefits across a range of research projects. Regardless of geographic location, data-intensive research relies on the flexible, reliable, high-capacity networks of the specialist research and education infrastructure community such as GÉANT. GÉANT connects Europe's NRENs –which in turn connect users across sectors such as biotechnology, medicine, radio astronomy and the arts –to bring researchers together, driving regional co-operation across the EU and strengthening European research as a whole. GÉANT has a very diversified European footprint with more than 25 POPs and through those able to reach more than 40 million users across 8000 institutions across 40 European Nations. GÉANT is also very well connected across the world making Europe to the very best connected region in the world. GÉANT is a very reliable network with a availability of 99.999%.

Presenter: Mr HERSOUG, Neils (DANTE)

Session Classification: Helix-Nebula Workshop: Networking Connectivity

Contribution ID: **208**

Type: **not specified**

Discussion

Wednesday, September 19, 2012 3:05 PM (25 minutes)

Session Classification: Helix-Nebula Workshop: Networking Connectivity

Contribution ID: 209

Type: **Presentation**

Personal data processed in cloud infrastructures: main legal aspects

Wednesday, September 19, 2012 4:25 PM (20 minutes)

Description of the work

Enrico Pelino is a Senior Associate at ICT Legal Consulting and an attorney in the Bologna bar since 2003. He is specialised in data protection and IT law, which have been his core interests all over his professional career. In recent years, he has been deeply involved in exploring the tangle of legal implications generated by the large-scale adoption of cloud computing technologies. He also works on such topics as a Fellow of the European Privacy Association (EPA) and the Istituto Italiano per la Privacy (Italian Privacy Institute). As an attorney, he actively practices private and commercial litigation and contract law. He graduated with honours from the University of Parma in 1998 and then had a longstanding collaboration with CIRSFID, a University-linked centre of research based in Bologna. In 2005, he earned his Ph.D. in IT law from the University of Bologna with a dissertation on data protection and control. He took part in a number of high-profile research projects, among which the 2007 Italian PRIN research project (Research Project of National Interest).

Printable Summary

Cloud computing involves a variety of actors related each other through a tangle of outsourcing schemes and a fast and widespread circulation of data: all elements that challenge classical approaches to data storage and even defy the traditional perception of national borders. The main aim of this presentation is to provide the audience with a quick overview of some basic legal questions: identification of the applicable legislation (as cloud services physically involve several countries); lawfulness of extra-EU transfers of personal data; determination of security standards; data subjects' rights to transparency and control, including the right to "data portability"; accountability of the cloud provider. Such topics are either totally dependant or strictly connected with a correct allocation of privacy roles (controller/processor), that have to be viewed in the light of the recent position expressed by the European data protection authorities (the "art. 29 Working Party"). The presentation will also raise attention over some provisions of the forthcoming Regulation on data protection, expected to enter into force in 2014. Finally, it will briefly address some generally shared concerns related to the protection of intellectual property in a cloud environment.

Primary author: Dr PELINO, Enrico (ICT Legal Consulting & EPA)

Presenter: Dr PELINO, Enrico (ICT Legal Consulting & EPA)

Session Classification: Helix-Nebula Workshop: Business Models and Legal Aspects

Contribution ID: **210**

Type: **not specified**

Discussion

Contribution ID: **211**

Type: **not specified**

Introduction and Goals

Wednesday, September 19, 2012 4:00 PM (5 minutes)

Presenter: ANDREOZZI, Sergio (EGLEU)

Session Classification: Helix-Nebula Workshop: Business Models and Legal Aspects

Contribution ID: 212

Type: **Presentation**

Business models for integrated e-infrastructures with commercial clouds

Wednesday, September 19, 2012 4:05 PM (20 minutes)

Printable Summary

Insights to business models i.e. definition, analysis, design and evaluation as foreseen by Work Package 7 of the Helix Nebula project. First the concept of business models is explained on the basis of the Osterwalder model. Secondly, steps to analyse and understand business models are explained. Thirdly, methods to design innovative business models and to evaluate a set of potential models are presented. In the end a few insights on the expected business case calculations are shown. Further, the slides highlight how Work Package 6 could collaborate to address the business case of publicly-funded infrastructures integrated with commercial providers.

Primary author: DOLL, Julia (SAP)

Presenter: SYMONDS, Mick (Atos)

Session Classification: Helix-Nebula Workshop: Business Models and Legal Aspects

Contribution ID: 213

Type: **Presentation**

Discussion / Workshop Wrap-up / Action Summary

Wednesday, September 19, 2012 5:05 PM (25 minutes)

Session Classification: Helix-Nebula Workshop: Business Models and Legal Aspects

Contribution ID: 214

Type: **Presentation**

Cost of e-Infrastructures

Wednesday, September 19, 2012 4:45 PM (20 minutes)

Description of the work

Fotis Karagiannis has 18 years of experience in the ICT research, focusing mainly in Research Networking and Computing e-Infrastructures. He received his PhD in 1998 in the fields of Integrated Communications and Management of Broadband Networks from NTUA, Greece. He participated in multiple European research projects working for commercial or research entities such as now OTEPlus, GRNET, CERN, CESNET, Microsoft Research Cambridge and Microsoft Innovation Center Greece, ATHENA Research Center, and Athens University of Economics and Business. He has worked for major e-Infrastructure projects such as the GEANT, EGEE series, EGI Design Study, PRACE, and the e-IRG support projects.

Printable Summary

The presentation summarises the current findings of the e-FISCAL project about evaluating the cost of e-Infrastructures. These findings are based on questionnaire data received from a mix of computing centres belonging to EGI, PRACE or purely national e-Infrastructures. In addition, a survey of the state of the art in financial analysis of ICT infrastructures will be presented, as well as some initial results of the benchmarking efforts of the project that compare commercial cloud offerings with research e-Infrastructures. The presentation will exploit the workshop as an opportunity to get feedback from the participants, especially on the financial front and the related business models.

Presenter: Dr KARAGIANNIS, Fotis (Independent)

Session Classification: Helix-Nebula Workshop: Business Models and Legal Aspects

Contribution ID: 215

Type: **not specified**

Discussion

Contribution ID: 216

Type: **Demonstration**

OpenNaaS, a toolkit for IP Networks as a Service

Description of the work

OpenNaaS

Wider impact of this work

OpenNaaS

Printable Summary

Dynamic and multi-domain provisioning of network resources has been a long-standing research area, and the rise of grid and cloud computing brings even more challenges. In order for operators to be able to deploy innovative Networks as a Service offerings, the Mantychore FP7 project has created the OpenNaaS framework. Grid and cloud computing provide some of the first use cases for OpenNaaS, making it possible to federate –at the network level –computing and storage resources between grid resource centres or to create a distributed computing cloud spanning several organisations.

Primary author: O'CALLAGHAN, David (TCD)

Presenter: O'CALLAGHAN, David (TCD)

Track Classification: Virtualised Resources: challenges and opportunities (Michel Drescher: track leader)

Contribution ID: 217

Type: **not specified**

Sustainability of NGI operations

Presenter: Dr FERRARI, Tiziana (EGLEU)

Contribution ID: 218

Type: **Presentation**

Sustainability of NGI operations

Printable Summary

We will present the output of the OMB survey that was launched in July 2012, whose purpose is to collect information about the readiness of NGIs towards self-sustaining operational services, and we will discuss the future steps to ensure continuity to operations across EGI.

Presenter: Dr FERRARI, Tiziana (EGLEU)

Contribution ID: **219**

Type: **not specified**

NGI Sustainability Checklist

Presenter: APPLETON, Owen (Emergence Tech Limited)

Contribution ID: 220

Type: **not specified**

Sustainability from an NGI Perspective

Contribution ID: 221

Type: **not specified**

Sustainability of NGI Operations

Thursday, September 20, 2012 11:00 AM (20 minutes)

Printable Summary

We will present the output of the OMB survey that was launched in July 2012, whose purpose is to collect information about the readiness of NGIs towards self-sustaining operational services, and we will discuss the future steps to ensure continuity to operations across EGI.

Presenter: Dr FERRARI, Tiziana (EGIEU)

Session Classification: Sustainability of National Infrastructures

Contribution ID: 222

Type: **Presentation**

NGI Sustainability Checklist

Thursday, September 20, 2012 11:20 AM (20 minutes)

Printable Summary

At present the NGI/EGI ecosystem is working to complete the transit from the project based funding to a set of sustainable governance, operation and funding structures to support long term success. This is an important but complex process and has been the subject of considerable discussion at previous EGI events. Despite hard work most NGIs have found it challenging to find new structures that are effective, achievable and will be supported by their national authorities. One reason for this is that sustainability requires a set of preconditions relatively well accepted in the commercial sector but more difficult for academic institutions. These include clear definitions of 'customers', service portfolios and catalogues, business models and IT Service Management processes. This presentation will introduce these factors and discuss how tackling them will facilitate work on sustainability.

Presenter: APPLETON, Owen (Emergence Tech Limited)

Session Classification: Sustainability of National Infrastructures

Contribution ID: 223

Type: **Presentation**

An NGI Sustainability Perspective: PL-Grid case study

Thursday, September 20, 2012 11:40 AM (20 minutes)

Presenter: Mr SZEPIENIEC, T. (CYFRONET)

Session Classification: Sustainability of National Infrastructures

Contribution ID: 224

Type: **not specified**

Panel Discussion

Thursday, September 20, 2012 12:00 PM (30 minutes)

Session Classification: Sustainability of National Infrastructures

Contribution ID: 225

Type: **not specified**

Closing plenary

Contribution ID: 226

Type: **Presentation**

EGI's Support services for Science Gateway developers

Thursday, September 20, 2012 11:10 AM (20 minutes)

Description of the work

NA

Wider impact of this work

NA

Printable Summary

Science gateways (SGs) are an emerging interface for communities to engage more actively with the European Grid Infrastructure.

The complexity of accessing distributed resources is hidden from researchers, therefore several communities are building gateways for their members.

This presentation will briefly touch on some topics that should become general knowledge for a science gateway enabler/developer, namely:

- Policies and procedures approved and in use by the EGI community;
- Applications Database as a way to advertise SGs and enabling technologies;
- Training marketplace to promote SG related events and training material;
- Surveys as a way to collect user requirements which will shape the evolution of SG technologies;
- RT as the way to feed SG community requirements.

Primary author: FERREIRA, Nuno (EGLEU)

Presenter: FERREIRA, Nuno (EGLEU)

Session Classification: Science Gateways

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 227

Type: **Presentation**

How to integrate portals with EGI monitoring system

Thursday, September 20, 2012 11:50 AM (20 minutes)

Printable Summary

Scientific gateways have become an essential tool for research in the age of e-Science. Their operation and performance has to be monitored in order to ensure quality of service for end users. Such monitoring also has to ensure an integrated overview of the global status of the scientific gateway, as well as the detailed status of the individual scientific gateway layers and components. In addition, it has to enable sending of alerts to administrators when a particular issue is identified, to enable scheduling of downtimes during service maintenance, as well as to produce infrastructure performance statistical reports.

In this talk we will give brief overview of the current production monitoring system in EGI, and discuss possibilities for its extensions to monitoring of scientific gateways. We will try to identify relevant properties of scientific gateways to be monitored, and to describe the corresponding probes.

Primary author: Mr VUDRAGOVIC, Dusan (Institute of Physics Belgrade)

Presenter: Mr VUDRAGOVIC, Dusan (Institute of Physics Belgrade)

Session Classification: Science Gateways

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 228

Type: **not specified**

How to integrate EGI portals with identity federations

Thursday, September 20, 2012 12:10 PM (20 minutes)

Printable Summary

in preparation

Description of the work

in preparation

Wider impact of this work

in preparation

Primary author: BARBERA, Roberto (University of Catania and INFN)

Presenter: BARBERA, Roberto (University of Catania and INFN)

Session Classification: Science Gateways

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 229

Type: **Presentation**

Data handling in scientific gateways on DCIs

Thursday, September 20, 2012 2:20 PM (20 minutes)

Description of the work

.

Wider impact of this work

.

Printable Summary

Science gateways provide easy access to computing resources. In most applications, attention also needs to be given to mechanisms to manipulate data via the web interface.

This involves two perspectives, from the application (domain-specific) and from the DCI.

In this talk we will discuss implications of both perspectives for the design of the data handling functionalities necessary for a science gateway, and presenting some of the existing tools.

Primary author: SANTCROOS, Mark (Academic Medical Center Amsterdam)

Presenter: SANTCROOS, Mark (Academic Medical Center Amsterdam)

Session Classification: Science Gateways

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 231

Type: **not specified**

Science Gateway session - concluding discussion

Thursday, September 20, 2012 3:20 PM (10 minutes)

Presenter: Dr LOVAS, Robert (MTA SZTAKI)

Session Classification: Science Gateways

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 233

Type: **not specified**

Science Gateway session introduction

Thursday, September 20, 2012 11:00 AM (10 minutes)

Presenter: Dr LOVAS, Robert (MTA SZTAKI)

Session Classification: Science Gateways

Contribution ID: 234

Type: **Presentation**

Commercial Networking Solutions - Helix Nebula Supplier

Wednesday, September 19, 2012 2:25 PM (20 minutes)

Description of the work

Udo Schäfer joined Alcatel-Lucent in February 1992. He studied computer science and received his degree from the University of Stuttgart. In different roles within the company he worked on networking solutions as the German research & education network G-Win built with SDH in 2000, the WDM based Géant2 built in 2005 and the IP based ESnet now in 2012. Today he leads the customer marketing team in Alcatel-Lucent's Global Account organization focusing on the Deutsche Telekom Group.

Printable Summary

The limiting factor in the adoption and success of Cloud Computing today is moving from IT towards the WAN infrastructure, where Internet –even if it is high speed –is not any more good enough. It is about guaranteed transfer rates, latency at the speed of light –all together without losing any bit of transferred data. This and inherent security of all data while being in WAN transit can be facilitated by commercial networking solutions.

Presenter: SCHÄFER, Udo (Alcatel-Lucent)

Session Classification: Helix-Nebula Workshop: Networking Connectivity

Contribution ID: 235

Type: **Presentation**

Network Connectivity - Scientific Clouds

Wednesday, September 19, 2012 2:45 PM (20 minutes)

Description of the work

Jurry de la Mar joined T-Systems/Deutsche Telekom group in 1994. He studied Nuclear Physics and received his degree from Free University in Amsterdam. He has started and lead various strategic projects with European Institutions for T-Systems in Europe. In 2008 he took over the sales and account management responsibility for European Institutions in Germany as well as the Galileo and Earth Observation programmes within T Systems' Public Sector. In 2011 he was one of the initiators to create Cloud Computing services for European Science which has now led to Helix Nebula –The Science Cloud consortium. And he is Member of the Supervisory Board of cesah GmbH, the Centre for Satellite Navigation in the State of Hesse, Germany. He began his career with Siemens where he held various international positions in the medical and high-tech industry in Germany, Netherlands and Sweden spanning from product management to sales and operations.

Printable Summary

Network connectivity has become a key element of success for cloud computing deployments. Especially when large distributed computing and/or data-intensive applications are concerned. Apart from building networks to provide ubiquitous bandwidth at acceptable costs industry is focussing also on solutions to simplify management of such complex infrastructures with software defined network (SDN) technologies. The presentation will give insight into current SDN-type developments and how this may benefit future e-Infrastructures and Helix Nebula deployments with flexible cloud network security.

Presenter: DE LA MAR, Jurry (T-Systems International GmbH)

Session Classification: Helix-Nebula Workshop: Networking Connectivity

Contribution ID: 236

Type: **not specified**

Workshop discussions

Presenter: Dr SIPOS, Gergely (EGLEU)

Contribution ID: 237

Type: **not specified**

Workshop discussions

Friday, September 21, 2012 2:36 PM (24 minutes)

Presenter: Dr SIPOS, Gergely (EGLEU)

Session Classification: Workflow community workshop

Contribution ID: 238

Type: **Training**

Genesis II –From the Global Federated File System to Running Jobs

Thursday, September 20, 2012 2:00 PM (1h 30m)

Description of the work

.

Wider impact of this work

.

Printable Summary

Genesis II, a new EGI technology provider, is an open source, standards based grid platform designed to support both high-throughput computing and secure data sharing. This hands-on training introduces participants to using Genesis II access layer tools and sharing capabilities such as the Global Federated File System (GFFS), the queue, and the client user interfaces.

The Genesis II Project is intended to provide the grid user community with easy-to-use tools to exploit grid technology to fulfil their computation and data management needs. The Genesis II platform is an interoperable (via adherence to community standards), and flexible platform for developing and evaluating new grid technologies and models.

Primary author: SARNOWSKA-UPTON, Karolina (University of Virginia)

Presenters: DOUGHERTY, Daniel (University of Virginia); SARNOWSKA-UPTON, Karolina (University of Virginia)

Session Classification: Genesis II tutorial

Track Classification: Resource Infrastructure services (Peter Solagna: track leader)

Contribution ID: 239

Type: **not specified**

Defining the GGUS Advisory Board

Friday, September 21, 2012 11:10 AM (30 minutes)

Printable Summary

Discussing:
Meeting frequency
Chair person
Tools
etc

Presenter: Dr ANTONI, Torsten (KIT-G)

Session Classification: GGUS Advisory Board

Contribution ID: **240**

Type: **not specified**

GGUS Report Generator

Friday, September 21, 2012 11:40 AM (20 minutes)

Presenter: GREIN, Guenter (KIT-G)

Session Classification: GGUS Advisory Board

Contribution ID: **241**

Type: **not specified**

Discussion

Friday, September 21, 2012 12:00 PM (30 minutes)

Session Classification: GGUS Advisory Board

Contribution ID: 242

Type: **Training**

Regional Operator on Duty - Training

Friday, September 21, 2012 11:00 AM (1h 30m)

Description of the work

.

Wider impact of this work

.

Printable Summary

This session is aimed at personnel contributing to NGI Regional Operator on Duty activities (ROD).

There will be three 30 minute slots. The first one will cover an overview of the Operations Portal - the tool used in daily operations support activities, and information on how recent developments influence daily operators work. Likely a new version of the dashboard will be presented.

The second slot is about the status and issues of the Grid Oversight activity. The Central Operator on Duty team (COD) will also provide information on current status and open issues.

The session will close with a Question and Answer session, which will give attendees the opportunity to ask any questions about grid oversight activities.

This session is particularly recommended to ROD personnel, in particular from NGIs that recently joined the production infrastructure.

Primary author: TROMPERT, Ron (SARA)

Presenters: LORPHELIN, Cyril (CNRS); BABIK, Marian (CERN); TROMPERT, Ron (SARA)

Session Classification: Regional Operator on Duty - Training

Track Classification: EGI Operations (Tiziana Ferrari: track leader)

Contribution ID: 243

Type: **Presentation**

EMI STS - Transforming the existing user credentials for the Grid

Wednesday, September 19, 2012 2:15 PM (15 minutes)

Description of the work

TBA

Wider impact of this work

TBA

Printable Summary

Demo/ Talk: The EMI project is responding to the challenges in the Grid credential management with a new general purpose service, the Security Token Service (STS). STS is a service that can be used for transforming an existing security token into another security token format. From the clients' point of view STS is a Web Service that, like any other Web Services, is accessed using the SOAP protocol. This means that any party capable of producing specified request messages and understanding response messages can act as a client if they have an access to a security token supported by STS. By enabling the token transformation, STS can establish a trust relationship between different security and application domains. This presentation also contains a short demonstration of the STS service in action.

Primary author: MIKKONEN, Henri (Helsinki Institute of Physics)

Presenter: MIKKONEN, Henri (Helsinki Institute of Physics)

Session Classification: AAI Workshop

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 244

Type: **not specified**

The SHIWA simulation platform for workflow interoperability and sharing

Friday, September 21, 2012 11:00 AM (22 minutes)

Workflow applications are very widely used to enact scientific applications on the grid infrastructure. Scientific workflow language provides high-level construct which is well suited to describe complex data-intensive applications. Communities may use many different workflow languages and to be able to share the scientific work, they need an interoperability solution. The workflow environment shields the end users from the low-level details related to the use of the grid infrastructure. The SHIWA project provides a multi-systems workflow execution platform and an interoperability solution, called the SHIWA Simulation Platform, that enables the collaboration of user communities that work on similar research field but use different workflow systems. The SHIWA Simulation Platform also includes a repository for publishing and sharing workflow applications. During the presentation both coarse- and fine-grained solutions will be introduced, together with the achievements the SHIWA project. The tools presented during this presentation can be of interest for any community who want to publish and share workflows, and also want to work together with others to ease and extend their research.

Description of the work

.

Link for further information

.

Wider impact of this work

.

Printable Summary

Workflow applications are very widely used to enact scientific applications on the grid infrastructure. Scientific workflow language provides high-level construct which is well suited to describe complex data-intensive applications. Communities may use many different workflow languages and to be able to share the scientific work, they need an interoperability solution. The workflow environment shields the end users from the low-level details related to the use of the grid infrastructure. The SHIWA project provides a multi-systems workflow execution platform and an interoperability solution, called the SHIWA Simulation Platform, that enables the collaboration of user communities that work on similar research field but use different workflow systems. The SHIWA Simulation Platform also includes a repository for publishing and sharing workflow applications. During the presentation both coarse- and fine-grained solutions will be introduced, together with the achievements the SHIWA project. The tools presented during this presentation can be of interest for any community who want to publish and share workflows, and also want to work together with others to ease and extend their research.

Primary author: KISS, Tamas (University of Westminster, London, UK)

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

Session Classification: Workflow community workshop

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 245

Type: **not specified**

Building an European Research Community through Interoperable Workflows and Data - the ER-Flow project

Friday, September 21, 2012 11:22 AM (22 minutes)

Printable Summary

ER-flow will build a European Research Community to promote workflow sharing and to investigate interoperability of the scientific data in workflow sharing. The project will disseminate the achievements of the FP7 SHIWA project, particularly the coarse-grained workflow interoperability based on the SHIWA Simulation Platform. It will target major research communities that use workflows to run their experiments on a regular basis. The project includes four major research communities: Astrophysics, Computation Chemistry, Heliophysics and Life Sciences. They will be supported to run experiments with the simulation platform. Beyond these communities the project will strongly collaborate with the National Grid Infrastructures through EGI.eu in order to identify and involve further major research communities which either already use workflows or which are perspective workflow users. The research communities will select workflows which can be used as pilot workflows in particular research area of a particular research community to demonstrate how to develop, use and share workflows. The project will port these pilot workflows to the simulation platform and publish them in a workflow repository. The pilot workflows first, will demonstrate how to use the simulation platform; secondly, researchers can use these workflows in their experiments; thirdly, they can modify them to create their own workflows. The pilot workflows will help to create a critical mass of workflows to enable workflow sharing inside and between research communities. ER-flow will collect and analyse requirements of the supported research communities towards interoperability of scientific data in the workflow domain. It will investigate existing protocols and standards that support this interoperability. The project will compile a study outlining the above mentioned requirements, protocols and standards and will make recommendations how to achieve interoperability of scientific data in the workflow domain.

Primary author: Mr TERSTYANSKY, Gabor

Presenter: Mr TERSTYANSKY, Gabor

Session Classification: Workflow community workshop

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 246

Type: **not specified**

Multiscale Programming and Execution Tools in the MAPPER project

Friday, September 21, 2012 11:44 AM (22 minutes)

Printable Summary

The subject of this talk are the multiscale programming and execution tools [1,3] developed in the MAPPER project. The MAPPER project responds to the critical need by developing computational strategies, software and services for distributed multiscale simulations across disciplines, exploiting existing and evolving European e-Infrastructure. It is driven by the computational needs of seven exemplary multiscale problems from a variety of disciplines including Systems Biology, Fusion, Physiology, Engineering and nano-Material Science. The presented tools support composition of multiscale applications from existing single scale submodules installed on e-infrastructures [2]. After being composed, such applications are executed. The application structure is described in Multiscale Modelling Language. The tools include: the application composition tool called Multiscale Application Designer (MAD), Registry for application modules description implemented as MAPPER Memory (MaMe) and GridSpace (GS) Experiment Workbench tools supporting high level stage of application execution. We will present the current status of the tools working on the example of multiscale application [3].

[1] K. Rycerz, M. Nowak, P. Pierzchala, E. Ciepiela, D. Harezlak, and M. Bubak: Comparison of Cloud and Local HPC Approach for MUSCLE-based multiscale simulations. In Proceedings of The Seventh IEEE International Conference on e-Science Workshops, Stockholm, Sweden, 5-8 December 2011. IEEE Computer Society, Washington, DC, USA, 81-88.

[2] K. Rycerz, M. Bubak: Building and Running Collaborative Distributed Multiscale Applications. In Large-Scale Computing Techniques for Complex System Simulations Wiley Series on Parallel and Distributed Computing, W. Dubitzky, K. Kurowski, and B. Schott, Eds. Vol. 1. John Wiley & Sons, Chapter 6, 111-130, 2012

[3] K. Rycerz et al: An Environment for Programming and Execution of Multiscale Applications, ACM Transactions on Modeling and Computer Simulation, in review

Primary author: Mrs RYCERZ, Katarzyna

Presenter: Mrs RYCERZ, Katarzyna

Session Classification: Workflow community workshop

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 247

Type: **not specified**

WS-VLAM workflow management system

Friday, September 21, 2012 12:06 PM (24 minutes)

Printable Summary

WS-VLAM is workflow management system which aim to provide and support coordinated execution of scientific workflow on distributed Grid and Cloud resources. This system combines the ability to take advantage of the underlying Grid and Cloud infrastructures and a flexible high-level rapid prototyping environment. WS-VLAM can model the computatin of the workflow graph in two ways: as a stream-based network, and as data flow network. In the stream based network, workflow components stream data between each other in a peer-to-peer fashion this entails that workflow tasks are co-allocated on resources and streaming channels are setup between tasks, the stream-based approach covers applications where communication is time critical. In the second execution models the workflows as a data flow graph tasks can only execute once its data dependencies have been met which means that all preceding tasks have produced some data to the current task and therefore it has data present on all its input port. This allows the task to start execution. The dataflow approach is suitable for a wider spectrum of scientific application especially legacy file based applications where communication between tasks is based on les. In large, long running workflows, dataflow is instrumental for better resource usage since a task at the end of the workflow do not waist time idling waiting for data to be produced but instead are only submitted to resources upon data availability.

Primary author: Mr BELLOUM, Adam

Presenter: Mr BELLOUM, Adam

Session Classification: Workflow community workshop

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 248

Type: **not specified**

Wf4Ever: Supporting Reuse and Reproducibility in Experimental Science

Friday, September 21, 2012 2:14 PM (22 minutes)

Printable Summary

Experimental reproducibility and knowledge sharing and reuse are key in reliable scientific development. The Wf4Ever project addresses these two challenges by proposing models and tools around the concept of workflow-centric research Objects (RO). ROs encapsulate workflows implementing scientific experiments together with the information required to run the experiment and replicate its results, and links it to external resources including both data and services. The notion of RO facilitates a deeper understanding, diagnosis and prevention of workflow decay, especially against changes in external third party resources. Consequently, it supports the maintenance of workflow integrity and therefore its long-term preservation. During this talk we will present the main building blocks around ROs in the Wf4Ever project and will make a quick analysis of the potential impact of these methods and tools in the particular domain of Astrophysics.

Primary author: SÁNCHEZ, Susana

Presenter: SÁNCHEZ, Susana

Session Classification: Workflow community workshop

Contribution ID: 249

Type: **Presentation**

VERCE e-infrastructure

Tuesday, September 18, 2012 3:06 PM (24 minutes)

Description of the work

tab

Wider impact of this work

tba

Printable Summary

place holder for text from Horst Schwichtenberg and Malcolm Praag

Primary author: SCHWICHTENBERG, Horst (FRAUNHOFER)

Presenter: SCHWICHTENBERG, Horst (FRAUNHOFER)

Session Classification: Research Infrastructures

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 250

Type: **not specified**

How to integrate portals with EGI accounting system

Thursday, September 20, 2012 2:00 PM (20 minutes)

Presenter: Dr GRACIANI DIAZ, Ricardo (University of Barcelona)

Session Classification: Science Gateways

Contribution ID: 252

Type: **Presentation**

Science Gateways and Clouds

Thursday, September 20, 2012 3:00 PM (20 minutes)

Printable Summary

Science gateways enable researchers to use domain applications and workflows on DCIs (Distributed Computing Infrastructures) without detailed knowledge of the underlying computer resources. They may be seen as the highest level of cloud services that can be provided to scientific end users. However, the service-oriented approach of clouds also allows assembling the underlying building blocks from the cloud, providing all the associated flexibility, modularity and scalability advantages. This talk gives an introduction to using clouds for scientific computing and to constructing science gateways from cloud building blocks. In particular, the approach of the SCI-BUS (SCientific gateway Based User Support) EU FP7 project is described in detail. Here, WS-PGRADE / gUSE from MTA SZTAKI provides the scientific gateway layer, the CloudBroker Platform the SaaS (Software as a Service) layer and different public or private cloud providers the IaaS (Infrastructure as a Service) layer.

Description of the work

SCI-BUS is a European project supported by the FP7 Capacities Programme under contract no. RI-283481. It aims at developing gateway technology and community gateways to provide researchers with seamless access to major computing, data and networking infrastructures and services, with focus on scientific workflows. SCI-BUS is a collaboration of 15 academic and commercial consortium members and six subcontractors, supporting gateways in various disciplines, such as proteomics, molecular science, seismology, business processes, medicine, rendering, astrophysics, SMEs, software testing, citizen web and heliophysics. SCI-BUS' major generic-purpose gateway technology building blocks are WS-PGRADE / gUSE and the CloudBroker Platform. gUSE is a framework of high-level services to achieve interoperation between DCIs and user communities. Its WS-PGRADE web portal provides scientific workflow management for both application developers and end users, and supports various DCIs including clusters, service grids, desktop grids and cloud. The CloudBroker Platform is a SaaS application store that allows easily offering and using compute-intensive applications on different cloud infrastructures. It can be accessed via a web browser or programmatically, and is available in public pay-per-use as well as hosted or in-house versions. Within the SCI-BUS project, recently an interface to the CloudBroker Platform has been integrated in WS-PGRADE / gUSE, which now allows to easily submit scientific workflow jobs to the cloud. Furthermore, the CloudBroker Platform is being extended to permit users also registering their own public or private cloud resources, giving the freedom to choose IaaS from different providers and tools, such as Amazon Web Services, IBM SmartCloud Enterprise, Eucalyptus and OpenStack.

Link for further information

<http://www.sci-bus.eu>; <http://www.cloudbroker.com>; <https://platform.cloudbroker.com>

Wider impact of this work

Clouds fundamentally change two aspects of computing: 1. the interface between providers and users, and 2. the underlying business model. Science gateways are a way to offer high-level computing services to domain application users, and thus themselves can be designed as cloud services. When also the underlying software and infrastructure resources are obtained from the cloud, a fully modular building and operation of such services is possible, either in-house or outsourced or in a hybrid fashion. This talk will help to better understand how cloud computing impacts computational science and what the SCI-BUS project and the CloudBroker Platform do to ease this transformation.

Primary author: SUDHOLT, Wibke (CloudBroker GmbH)

Presenter: SUDHOLT, Wibke (CloudBroker GmbH)

Session Classification: Science Gateways

Track Classification: Virtual Research Environments (Gergely Sipos: track leader)

Contribution ID: 253

Type: **not specified**

General Introduction to the Hands-On

Tuesday, September 18, 2012 11:00 AM (20 minutes)

Presenter: Dr GABRIEL, Sven (Nikhef)

Session Classification: EGI-CSIRT Security Training Hands-on Wrap-Up

Contribution ID: 254

Type: **not specified**

IT forensics on the battlefield

Tuesday, September 18, 2012 11:20 AM (1 hour)

Presenters: REESE, Heiko (KIT-G); NIXON, Leif (LIU)

Session Classification: EGI-CSIRT Security Training Hands-on Wrap-Up

Contribution ID: 255

Type: **not specified**

Introduction to Part I - "Klingon Attack"

Tuesday, September 18, 2012 12:20 PM (10 minutes)

Presenter: Dr GABRIEL, Sven (Nikhef)

Session Classification: EGI-CSIRT Security Training Hands-on Wrap-Up

Contribution ID: 256

Type: **not specified**

Discussion with participants on Part I

Wednesday, September 19, 2012 4:00 PM (45 minutes)

Presenter: NIXON, Leif (LIU)

Session Classification: EGI-CSIRT Security Training Hands-on Wrap-Up

Contribution ID: 257

Type: **not specified**

Introduction to Part II - "Klingons in the grid"

Wednesday, September 19, 2012 4:45 PM (45 minutes)

Presenters: REESE, Heiko (KIT-G); Dr GABRIEL, Sven (Nikhef)

Session Classification: EGI-CSIRT Security Training Hands-on Wrap-Up

Contribution ID: 258

Type: **not specified**

Introduction

Thursday, September 20, 2012 4:00 PM (5 minutes)

Presenter: BURKE, Stephen (EGLEU)

Session Classification: EGI InSPIRE

Contribution ID: 259

Type: **not specified**

Evolution of the information system as seen by EMI

Presenter: FIELD, Laurence (CERN)

Contribution ID: 260

Type: **not specified**

EMIR multi-tier performance testing

Presenter: SHIRAZ MEMON, ahmed (JUELICH)

Contribution ID: **261**

Type: **not specified**

STS status

Thursday, September 20, 2012 2:05 PM (15 minutes)

Presenter: MIKKONEN, Henri (Helsinki Institute of Physics)

Session Classification: Resource Infrastructure Services

Contribution ID: 262

Type: **not specified**

Exploring the SAML 2.0 ECP-Profile

Thursday, September 20, 2012 2:20 PM (15 minutes)

Primary author: LINDQVIST, Carolina (Helsinki Institute of Physics)

Presenter: LINDQVIST, Carolina (Helsinki Institute of Physics)

Session Classification: Resource Infrastructure Services

Contribution ID: 263

Type: **not specified**

Common AuthN libraries

Thursday, September 20, 2012 2:35 PM (15 minutes)

Primary author: SUSTR, Zdenek (CESNET)

Presenter: KOURIL, Daniel (CESNET)

Session Classification: Resource Infrastructure Services

Contribution ID: 264

Type: **not specified**

Trustmanager and CANI - Joni Hahkala

Thursday, September 20, 2012 2:50 PM (15 minutes)

Primary author: HAHKALA, Joni (Helsinki Institute of Physics/EMI)

Presenter: HAHKALA, Joni (Helsinki Institute of Physics/EMI)

Session Classification: Resource Infrastructure Services

Contribution ID: 265

Type: **not specified**

EMI Hydra and Cloud(s)

Thursday, September 20, 2012 3:05 PM (15 minutes)

Presenter: WHITE, John (University of Helsinki, Finland)

Session Classification: Resource Infrastructure Services

Contribution ID: 266

Type: **not specified**

Argus EES

Thursday, September 20, 2012 3:20 PM (10 minutes)

Primary author: KOEROO, Oscar (FOM)

Presenter: KOEROO, Oscar (FOM)

Session Classification: Resource Infrastructure Services

Contribution ID: 267

Type: **not specified**

Argus tutorial: Installation, configuration and policy writing

Thursday, September 20, 2012 4:00 PM (45 minutes)

Presenter: Mr TSCHOPP, Valery (SWITCH)

Session Classification: Resource Infrastructure Services

Contribution ID: 268

Type: **not specified**

gLExec tutorial: Installation, configuration and policy writing

Thursday, September 20, 2012 4:45 PM (45 minutes)

Presenters: VAN DOK, Dennis (FOM); KOEROO, Oscar (FOM)

Session Classification: Resource Infrastructure Services

Contribution ID: **269**

Type: **not specified**

GLUE2 XML renderings

Thursday, September 20, 2012 4:05 PM (20 minutes)

Presenter: MEREDITH, david (STFC)

Session Classification: EGI InSPIRE

Contribution ID: 270

Type: **not specified**

EMI Security Token Service - Simplified Credential Management

Tuesday, September 18, 2012 12:15 PM (15 minutes)

Printable Summary

The EMI project is responding to the challenges in the Grid credential management with a new general purpose service, the Security Token Service (STS). STS is a service that can be used for transforming an existing security token into another security token format. As the supported security token formats include username/password, SAML assertion, X.509 certificate and VOMS proxy certificate, a user holding any type of such credentials is able to convert them into another supported type, for instance in order to access Grid services. By enabling the token transformation, STS can establish a trust relationship between different security and application domains. The presentation contains a short introduction to the technology, description of the most important use cases and the current status of the implementation.

Presenter: MIKKONEN, Henri (Helsinki Institute of Physics)

Session Classification: New middleware products

Contribution ID: 271

Type: **not specified**

Using EGI AppDB in dissemination

Tuesday, September 18, 2012 11:10 AM (10 minutes)

Presenter: EIGELIS, Karolis (EGLEU)

Session Classification: Marketing and Communications Strategies

Contribution ID: 272

Type: **not specified**

Scientific Publications Repository

Tuesday, September 18, 2012 11:20 AM (10 minutes)

Presenter: ROMIER, Romier (CNRS)

Session Classification: Marketing and Communications Strategies

Contribution ID: 273

Type: **not specified**

The EGI Campus Champions Scheme

Tuesday, September 18, 2012 11:30 AM (10 minutes)

Presenter: MCLENNAN, Richard (EGLEU)

Session Classification: Marketing and Communications Strategies

Contribution ID: 274

Type: **not specified**

ENVRI, communicating environmental research

Tuesday, September 18, 2012 11:40 AM (15 minutes)

Presenter: Mr LEGRE, Yannick

Session Classification: Marketing and Communications Strategies

Contribution ID: 275

Type: **not specified**

Measuring Impact with e-ScienceTalk

Tuesday, September 18, 2012 11:55 AM (10 minutes)

Presenter: Ms QADIR, Zara

Session Classification: Marketing and Communications Strategies

Contribution ID: 276

Type: **not specified**

The Science Gateway

Wednesday, September 19, 2012 2:30 PM (30 minutes)

Presenter: BARBERA, Roberto (University of Catania and INFN)

Session Classification: CHAIN Interoperability Workshop

Contribution ID: 277

Type: **not specified**

Science Gateway Demo and Q&A

Wednesday, September 19, 2012 3:00 PM (30 minutes)

Session Classification: CHAIN Interoperability Workshop

Contribution ID: 278

Type: **not specified**

GISELA and LA experience - Luis Nunez (CLARA)

Wednesday, September 19, 2012 4:00 PM (30 minutes)

Session Classification: CHAIN Interoperability Workshop

Contribution ID: 279

Type: **not specified**

The CHAIN roadmap

Wednesday, September 19, 2012 4:30 PM (30 minutes)

Presenter: MATYSKA, Ludek (CESNET)

Session Classification: CHAIN Interoperability Workshop

Contribution ID: **280**

Type: **not specified**

Final discussion and wrap-up

Wednesday, September 19, 2012 5:00 PM (30 minutes)

Session Classification: CHAIN Interoperability Workshop

Contribution ID: **281**

Type: **not specified**

EGI Security Policy Group Public report

Tuesday, September 18, 2012 11:05 AM (40 minutes)

Presenter: KELSEY, David (STFC)

Session Classification: EGI InSPIRE

Contribution ID: 282

Type: **not specified**

EGI Security Threat Risk Assessment public report

Tuesday, September 18, 2012 11:45 AM (45 minutes)

Presenter: CORNWALL, Linda (STFC)

Session Classification: EGI InSPIRE

Contribution ID: 283

Type: **not specified**

Software Vulnerability issue handling in EGI

Friday, September 21, 2012 11:05 AM (25 minutes)

Presenter: CORNWALL, Linda (STFC)

Session Classification: Operations

Contribution ID: 284

Type: **not specified**

Vulnerability Assessment of Middleware packages supplied by EMI :VOMS Core

Friday, September 21, 2012 11:30 AM (30 minutes)

Presenter: BRUGNOLI, Manuel (UAB)

Session Classification: Operations

Contribution ID: 285

Type: **not specified**

Top 10 Code flaws that cause real vulnerabilities

Friday, September 21, 2012 12:00 PM (30 minutes)

Presenter: HEYMANN, Elisa

Session Classification: Operations

Contribution ID: **286**

Type: **not specified**

IPv6 Reality check

Tuesday, September 18, 2012 3:00 PM (20 minutes)

Presenter: PRELZ, Francesco (INFN)

Session Classification: Network Support

Contribution ID: **287**

Type: **not specified**

Q & A / Discussion

Tuesday, September 18, 2012 3:20 PM (10 minutes)

Presenter: Dr REALE, Mario (GARR)

Session Classification: Network Support

Contribution ID: **288**

Type: **not specified**

Q & A / Discussion

Tuesday, September 18, 2012 5:00 PM (30 minutes)

Session Classification: Network Support

Contribution ID: 289

Type: **not specified**

Adopting the BDII as an information system for Globus services

Thursday, September 20, 2012 4:25 PM (15 minutes)

Presenter: SAVERCHENKO, Ilya (BADW)

Session Classification: EGI InSPIRE

Contribution ID: **290**

Type: **not specified**

Discussion

Thursday, September 20, 2012 4:40 PM (20 minutes)

Session Classification: EGI InSPIRE

Contribution ID: 291

Type: **not specified**

**EGI CSIRT (closed) meeting - agenda at
<https://indico.egi.eu/indico/conferenceDisplay.py?confId=1160>**

Monday, September 17, 2012 2:00 PM (1h 30m)

Presenters: KELSEY, David (STFC); Dr GABRIEL, Sven (Nikhef)

Session Classification: EGI CSIRT (Closed)

Contribution ID: 292

Type: **not specified**

**EGI CSIRT (closed) meeting - agenda at
<https://indico.egi.eu/indico/conferenceDisplay.py?confId=1160>**

Monday, September 17, 2012 4:00 PM (1h 30m)

Presenters: KELSEY, David (STFC); Dr GABRIEL, Sven (Nikhef)

Session Classification: EGI CSIRT (Closed)

Contribution ID: 293

Type: **not specified**

Introduction to the workshop

Tuesday, September 18, 2012 11:00 AM (15 minutes)

Presenter: C. BUSQUETS PÉREZ, Luis

Session Classification: EC Workshop: Distributed Computing Infrastructures for e-Science: Future Perspectives

Contribution ID: 294

Type: **not specified**

Science Computing requirements

Description of the work

DCIs and Science Requirements: summary of Iberian Research Area experience, Isabel Campos

Distributed Computing for Life-Sciences and Medical Research in the Genome-Age”, Andrew Lyall

Cloud Services for scientific institutions, Michael Symonds

Role of software in enabling DC Infrastructures, Oxana Smirnova

Presenters: LYALL, Andrew (EMBL); Dr PLASENCIA, Isabel (CSIC); SYMONDS, Mick (Atos); SMIRNOVA, Oxana (Lund University / NDGF)

Contribution ID: 295

Type: **not specified**

Distributed Computing Services

Presenters: STEIJAERT, Andres; GALIZIA, Antonella; Dr JONES, Bob (CERN); WALLOM, David (OXFORD); NEWHOUSE, Steven (EGLEU)

Contribution ID: 296

Type: **not specified**

Innovation potential and resource constraints in distributed computing

Presenters: Dr KARAGIANNIS, Fotis (Independent); Dr KUROWSKI, Krzysztof (PSNC); ROBERTSHAW, Steve

Contribution ID: 297

Type: **not specified**

DCIs and Science Requirements: summary of Iberian Research Area experience

Tuesday, September 18, 2012 11:15 AM (15 minutes)

Presenter: Dr PLASENCIA, Isabel (CSIC)

Session Classification: EC Workshop: Distributed Computing Infrastructures for e-Science: Future Perspectives

Contribution ID: 298

Type: **not specified**

Distributed Computing for Life-Sciences and Medical Research in the Genome-Age

Tuesday, September 18, 2012 11:30 AM (15 minutes)

Presenter: LYALL, Andrew (EMBL)

Session Classification: EC Workshop: Distributed Computing Infrastructures for e-Science:
Future Perspectives

Contribution ID: 299

Type: **not specified**

Cloud Services for scientific institutions

Tuesday, September 18, 2012 11:45 AM (15 minutes)

Presenter: SYMONDS, Mick (Atos)

Session Classification: EC Workshop: Distributed Computing Infrastructures for e-Science: Future Perspectives

Contribution ID: **300**

Type: **not specified**

Role of software in enabling DC Infrastructures

Tuesday, September 18, 2012 12:00 PM (15 minutes)

Presenter: SMIRNOVA, Oxana (Lund University / NDGF)

Session Classification: EC Workshop: Distributed Computing Infrastructures for e-Science: Future Perspectives

Contribution ID: **301**

Type: **not specified**

Questions and discussion

Tuesday, September 18, 2012 12:15 PM (15 minutes)

Session Classification: EC Workshop: Distributed Computing Infrastructures for e-Science: Future Perspectives

Contribution ID: **302**

Type: **not specified**

Federated e-Infrastructures in 2020

Tuesday, September 18, 2012 2:00 PM (15 minutes)

Presenter: NEWHOUSE, Steven (EGLEU)

Session Classification: EC Workshop: Distributed Computing Infrastructures for e-Science: Future Perspectives

Contribution ID: 303

Type: **not specified**

Cloud Computing Services for Scientific Research Organisations

Tuesday, September 18, 2012 2:15 PM (15 minutes)

Presenter: Dr JONES, Bob (CERN)

Session Classification: EC Workshop: Distributed Computing Infrastructures for e-Science: Future Perspectives

Contribution ID: 304

Type: **not specified**

An e-Science Distributed HPC Infrastructure to support Hydro-Meteorological Research

Tuesday, September 18, 2012 2:30 PM (15 minutes)

Presenter: GALIZIA, Antonella

Session Classification: EC Workshop: Distributed Computing Infrastructures for e-Science: Future Perspectives

Contribution ID: 305

Type: **not specified**

Distributed computing and the adoption of cloud services. Working towards a connected collaboration infrastructure

Tuesday, September 18, 2012 2:45 PM (15 minutes)

Presenter: STEIJAERT, Andres

Session Classification: EC Workshop: Distributed Computing Infrastructures for e-Science: Future Perspectives

Contribution ID: 306

Type: **not specified**

An e-infrastructure eco system for research must include glue not just the trimmings

Tuesday, September 18, 2012 3:00 PM (15 minutes)

Presenter: WALLOM, David (OXFORD)

Session Classification: EC Workshop: Distributed Computing Infrastructures for e-Science: Future Perspectives

Contribution ID: **307**

Type: **not specified**

Questions and discussion

Tuesday, September 18, 2012 3:15 PM (15 minutes)

Session Classification: EC Workshop: Distributed Computing Infrastructures for e-Science: Future Perspectives

Contribution ID: 308

Type: **not specified**

From GridLab to MAPPER - lessons learned over the last decade to address better challenges in distributed computing in the future

Tuesday, September 18, 2012 4:00 PM (25 minutes)

Presenter: Dr KUROWSKI, Krzysztof (PSNC)

Session Classification: EC Workshop: Distributed Computing Infrastructures for e-Science: Future Perspectives

Contribution ID: **309**

Type: **not specified**

• **A policy view on DCIs for H2020: a hybrid future**

Tuesday, September 18, 2012 4:25 PM (25 minutes)

Presenter: Dr KARAGIANNIS, Fotis (Independent)

Session Classification: EC Workshop: Distributed Computing Infrastructures for e-Science: Future Perspectives

Contribution ID: 310

Type: **not specified**

Technology"Transfer: The Missing Link in the Innovation and Sustainability Stories

Tuesday, September 18, 2012 4:50 PM (25 minutes)

Presenter: ROBERTSHAW, Steve

Session Classification: EC Workshop: Distributed Computing Infrastructures for e-Science: Future Perspectives

Contribution ID: **311**

Type: **not specified**

Questions and discussion

Tuesday, September 18, 2012 5:15 PM (15 minutes)

Session Classification: EC Workshop: Distributed Computing Infrastructures for e-Science: Future Perspectives

Contribution ID: 312

Type: **not specified**

Ongoing operational issues of BIOMED VO

Wednesday, September 19, 2012 11:05 AM (15 minutes)

Presenter: MICHEL, Franck (CNRS)

Session Classification: Operations - Resource Centre Forum

Contribution ID: **313**

Type: **not specified**

Discussion

Wednesday, September 19, 2012 4:00 PM (1h 30m)

Presenters: Dr SIPOS, Gergely (EGIEU); SOLAGNA, Peter (EGIEU)

Session Classification: AAI Workshop

Contribution ID: 314

Type: **not specified**

EMI 1 WMS - Installation and Configuration

Wednesday, September 19, 2012 3:00 PM (30 minutes)

Presenter: GIORGIO, Emidio (INFN)

Session Classification: EMI 2 Matterhorn - tutorial for system administrators

Contribution ID: 315

Type: **not specified**

DPM migration from gLite 3.2 to EMI 2

Wednesday, September 19, 2012 4:00 PM (45 minutes)

Presenter: GIORGIO, Emidio (INFN)

Session Classification: EMI 2 Matterhorn - tutorial for system administrators

Contribution ID: 316

Type: **not specified**

Introduction to EMIR : installation, service configuration, and demonstration of a working instance

Wednesday, September 19, 2012 4:45 PM (45 minutes)

Presenter: MARTON, Ivan (NIIF Institute)

Session Classification: EMI 2 Matterhorn - tutorial for system administrators

Contribution ID: 317

Type: **not specified**

CREAM installation, configuration and troubleshooting

Wednesday, September 19, 2012 2:00 PM (1 hour)

Presenter: ANDREETTO, Paolo (INFN)

Session Classification: EMI 2 Matterhorn - tutorial for system administrators

Contribution ID: **318**

Type: **not specified**

Configuring EMI services through YAIM

Wednesday, September 19, 2012 11:10 AM (50 minutes)

Presenter: Dr AIFTIMIEI, Cristina (INFN)

Session Classification: EMI 2 Matterhorn - tutorial for system administrators

Contribution ID: **319**

Type: **not specified**

EMI 2 site BDII installation and configuration

Wednesday, September 19, 2012 12:00 PM (30 minutes)

Presenter: Dr AIFTIMIEI, Cristina (INFN)

Session Classification: EMI 2 Matterhorn - tutorial for system administrators

Contribution ID: **320**

Type: **not specified**

EGI SPG meeting (closed)

Friday, September 21, 2012 3:30 PM (30 minutes)

See agenda at

<https://indico.egi.eu/indico/conferenceDisplay.py?confId=1161>

Presenter: KELSEY, David (STFC)

Session Classification: EGI InSPIRE

Contribution ID: **321**

Type: **not specified**

Manchester

Friday, September 21, 2012 9:50 AM (10 minutes)

Presenter: DEVEREUX, Claire (STFC)

Session Classification: Plenary

Contribution ID: 322

Type: **not specified**

test

Description of the work

test

Wider impact of this work

test

Printable Summary

test

Primary author: MCLENNAN, Richard (EGLEU)

Presenter: MCLENNAN, Richard (EGLEU)

Track Classification: Community and Co-ordination (Sergio Andreozzi: track leader)