



# Community Forum

*held in conjunction with the 2<sup>nd</sup> EMI technical conference*



## 2012 Book of Abstracts

First published in 2012 in the Netherlands by EGI.eu  
Science Park 140  
1098 XG Amsterdam  
The Netherlands  
[www.egi.eu](http://www.egi.eu)  
[contact@egi.eu](mailto:contact@egi.eu)

ISBN 978 90 816927 0 0

Copyright © 2012 by Members of the EGI-InSPIRE Collaboration

This work is licenced under the Creative Commons  
Attribution-Noncommercial 3.0 Licence.

To view a copy of this licence, visit

<http://creativecommons.org/licenses/by-nc/3.0>

or send a letter to Creative Commons, 171 Second Street,  
Suite 300, San Francisco, California, 94105, USA.

The work must be attributed by attaching the following reference to the copied  
elements: Copyright © Members of the EGI-InSPIRE Collaboration, 2012.

See [www.egi.eu](http://www.egi.eu) for details of the EGI-InSPIRE project and the collaboration.

Using this document in a way and/or for purposes not foreseen in the licence  
requires the prior written permission of the copyright holders.

EGI-InSPIRE (European Grid Initiative: Integrated Sustainable Pan-European  
Infrastructure for Researchers in Europe) is a project co-funded by the European  
Commission as an Integrated Infrastructure Initiative within the 7<sup>th</sup> Framework  
Programme. EGI-InSPIRE began in May 2010 and runs for 4 years.

The Organisers and the Publisher have made every effort to provide accurate  
and complete information in this Book. However, changes or corrections may  
occasionally be necessary after publication.



Editor: Sara Coelho  
Design: Viviane Li  
Printed in the Germany by Neumann Druck OHG

# EGI COMMUNITY FORUM 2012

---

held in conjunction with the  
2<sup>nd</sup> EMI TECHNICAL CONFERENCE



26–30 MARCH 2012, MUNICH, GERMANY

# Contents

Index of authors . . . . .	9
----------------------------	---

## USERS AND COMMUNITIES 12

ScalaLife—providing scalable software services and expertise to Life Sciences . . . . .	14
Added value of new features of the ATLAS computing model and a shared Tier-2 and Tier-3 facilities from the community point of view . . . . .	14
SuperB evaluation of Dirac distributed infrastructure . . . . .	15
Earth Science VRC activities . . . . .	15
Development and deployment of distributed e-VLBI components . . . . .	16
CTACG—Cherenkov Telescope Array Computing Grid . . . . .	16
e-Science on weather simulation using the WRF model . . . . .	17
High-throughput interactive in-silico drug discovery platform: GVSS-2 . . . . .	17
Achievements and perspectives of the biomed technical team . . . . .	18
Experiment representation at the WLCG Tier-1 centre GridKa . . . . .	18
The evolution of the EGI user support platform . . . . .	19
Outreach to new EGI communities . . . . .	19
EGI user requirements—processes, tools and achievements . . . . .	20
Extending WeNMR e-infrastructure outside Europe . . . . .	20
Support for non-LHC VOs by GridPP in the UK . . . . .	21

## SOFTWARE SERVICES FOR USERS AND COMMUNITIES 22

MILU, the three middleware user interfaces . . . . .	24
Exploiting grid resources for data simulation by using a general purpose framework . . . . .	24
A general purpose grid and cloud portal to simplify scientific communities' integration into distributed computing infrastructures . . . . .	25
Cloud, ready for bioinformatics? . . . . .	25
Workflows for the simulation of organic light-emitting diodes . . . . .	26
A grid execution model for computational chemistry applications using GC3Pie and AppPot . . . . .	26
The Biovel Project: robust phylogenetic workflows running on the grid . . . . .	27
A data engine for grid science gateways enabling easy transfer and data sharing . . . . .	27
The grid goes 'social': extending authentication to members of social network . . . . .	28
A generic gateway and its customisation in different scientific areas . . . . .	28
The MoSGrid community—from national to international scale . . . . .	29
Workflow and data management for nuclear magnetic resonance . . . . .	29
Interactive information extraction based on distributed data management for D-Grid projects . . . . .	30
WisNetGrid service layer—enabling creative research workflows . . . . .	30
SOMA2—application oriented molecular modelling workflows in www-browser . . . . .	31
InSilicoLab: a user workspace implementing e-science principles . . . . .	31
Using SHIWA workflow interoperability tools for neuroimaging data analysis applications . . . . .	32
Okeanos IaaS . . . . .	32



Data-intensive processing with Hadoop and friends at BiG Grid . . . . .	.33
AppPot: bridging the grid and cloud worlds. . . . .	.33
Cloud federation over the world wide grid . . . . .	.34
The VRC-driven GISELA Science Gateway . . . . .	.34
Challenges in data management and running DNA sequencing experiments on grid . . . .	.35
CRAB: a user friendly application for distributed data processing for the Compact Muon Solenoid experiment at the LHC . . . . .	.35
Building a 'grid of clouds' –how one HEP experiment is evaluating strategies to incorporate 'The Cloud' into existing grid infrastructures. . . . .	.36
gridCake and gridCamp: making the grid easier . . . . .	.36
Handling of network and database instabilities in CORAL . . . . .	.37
Web4Grid, web interface for grid jobs . . . . .	.37
Operating a public cloud with StratusLab . . . . .	.38
Beehub, facilitating scientific collaboration. . . . .	.38
Supporting grid-enabled GPU workloads using rCUDA and StratusLab . . . . .	.39

## MIDDLEWARE SERVICES

40

Improvements to the EMI build and test tools . . . . .	.42
Advanced data staging in the ARC Computing Element . . . . .	.42
EMI Release Management Highlights–EMI 1 Updates and EMI 2 Status. . . . .	.43
Advancements in EMI Testing Infrastructure implementation . . . . .	.43
SIENA: grid and cloud standards for e-Science and beyond . . . . .	.44
Quality Control in EMI . . . . .	.44
Performance testing of distributed computational resources in the software development phase. . . . .	.45
Common framework for extracting information and metrics from multiple change trackers . . . . .	.45
EMI Execution Service implementation in ARC . . . . .	.46
EMI common authentication library . . . . .	.46
Extended grid data management for UNICORE–accessing files in SRM and LFC . . . . .	.47
Federated access to data storage with Moonshot . . . . .	.47
The gCube data transfer facilities. . . . .	.48
Lessons learned from UNICORE EMI-ES Adoption towards improved open standards . . .	.48
EMIR: An EMI service registry for federated grid infrastructures . . . . .	.49
dCache, agile adoption of storage technology. . . . .	.49
EMI Data, preparing for the finish line. . . . .	.50
The ARC Information System: overview of a GLUE2 compliant production system . . . . .	.50
Software for distributed systems–the EMI product portfolio . . . . .	.51



Standardisation strategies of the European Middleware Initiative. . . . .	51
T-StoRM: a StoRM testing framework . . . . .	52
Grid Engine batch system integration in the EMI era . . . . .	52
CREAM Computing Element: status and new developments . . . . .	53
Using L&B to monitor Torque jobs across a national grid. . . . .	53

## **OPERATIONAL SERVICES AND INFRASTRUCTURE 54**

---

User-centric monitoring of the analysis and production activities within ATLAS and CMS VOs using the Experiment Dashboard system . . . . .	56
The GGUS report generator. . . . .	56
The role of GGUS in delivering highly reliable WLCG operations . . . . .	57
Site Status Board: a flexible monitoring system developed in close collaboration with user communities. . . . .	57
Hybrid cloud-based grid infrastructure: experience and future . . . . .	58
User community input to new types of accounting . . . . .	58
Improving resilience of T0 grid services . . . . .	59
DDM Site Services: a solution for global replication of HEP data. . . . .	59
Grid/cloud computing platform for medical imaging and signal processing in clinical decision system. . . . .	60
Consistency between grid storage elements and file catalogue for the LHCb experiment. . . . .	60
Service Availability Monitoring (SAM) . . . . .	61
Accounting of storage resources usage for the LHCb VO . . . . .	61

Gratia: new development in grid accounting . . . . .	62
StratusLab: use cases, features and sustainability. . . . .	62
Monitoring in CORAL . . . . .	63
The Green Computing Observatory . . . . .	63
Messaging service and client software . . . . .	64
Optimising the usage of multi-Petabyte storage resources for LHC experiments . . . . .	64
Experience in grid site testing for High Energy Physics with HammerCloud . . . . .	65
Experiences running a HPC Cloud. . . . .	65
Providing resources through cloud and grid interfaces by means of the standard CREAM CE and the WNoDeS cloud solution . . . . .	66

## **COORDINATION AND COMMUNICATION 68**

---

A business model approach for a sustainable grid infrastructure in Germany . . . . .	70
How do research teams commission original software applications and ensure that they do what they want them to do? . . . . .	70
The EGI Software Vulnerability Group and EMI . . . . .	71
Communication: the key to future funding . . . . .	71
Setting up a Finnish NGI site–site perspective. . . . .	72
Strategic plan for a scientific cloud computing infrastructure for Europe . . . . .	72
Linking authenticating and authorising infrastructures in the UK NGI (SARoNGS) . . . . .	73
Tweaking the certificate lifecycle for the UK e-Science CA . . . . .	73
PLGrid PLUS: towards domain-specific infrastructure for supporting international research collaboration . . . . .	74



European Globus Community Forum (EGCF) . . . . .	74
High Performance grid computing: getting HPC and HTC all together . . . . .	75
GC3Pie: A Python framework for high-throughput computing . . . . .	75
EGI sustainability . . . . .	76
ScienceSoft – open software for open science. . . . .	76
Supporting EGI strategic dissemination activities: the European e-infrastructures Observatory . . . . .	77
Championing e-infrastructure . . . . .	77
How many users does EGI have? . . . . .	78
In case the 'happy accident' doesn't happen: training the next generation of senior managers within European e-Infrastructures . . . . .	78

## Index of authors

<b>A</b>		<b>C</b>		<b>E</b>	
Apostolov, Rossen	14	Cameron, David	42	Elena, Oliver	14
Abad Rodriguez, Andres	42	Capannini, Fabio	53	Eleonora, Luppi	24
Agarwal, Sudhir	30	Casajus, Adria	61	Ellert, Mattias	50
Aiftimiei, Cristina	43	Ceccanti, Andrea	25, 43	Enrico, Vianello	24
Aleksiev, Tyanko	24	Cecchi, Marco	53		
Alvaro, Fernandez	14	Chen, Tim	34	<b>F</b>	
Andrade, Pedro	61	Chien, Jinny	34	Fargetta, Marco	27, 28
Andreetto, Paolo	53	Cinquilli, Mattia	56, 64	Farkas, Zoltan	28
Andreeva, Julia	56, 57	Collados, David	61	Feijoo, Alejandro	58
Andreozzi, Sergio	76	Cornwall, Linda	71	Fella, Armando	15
Antoni, Torsten	56, 70	Costantini, Alessandro	26	Ferguson, Nicholas	44
Armando, Fella	24			Fernandez, Carlos	58
		<b>D</b>		Fiameni, Giuseppe	44
<b>B</b>		Dalla Torre, Gianni	66	Filipovic, Jiri	53
Babik, Marian	61	Denvil, Sebastien	15	Flix, Josep	57
Barbera, Roberto	27, 28, 34	Diaz Alvarez, Ivan	58	Franco-Valiente, José Miguel	36
Barreiro Megino, Fernando	36	Diez, Rubén	58	Freire, Esteban	58
Bencivenni, Marco	25	Dimou, Maria	57	Frizziero, Eric	53
Benedyczak, Krzysztof	46	Dongiovanni, Danilo	43	Fuente, Paloma	61
Berlich, Ruediger	70	Donvito, Giacinto	15, 27, 66	Fuhrmann, Patrick	50
Bertocco, Sara	53	Dorigo, Alvise	53	Furst, Frederic	63
Blanchet, Christophe	25	Dres, Helmut	56		
Borges, Gonçalo	52	Dulov, Oleg	56		
Bozic, Stefan	26	Dvorak, Frantisek	53		
Brewer, Stephen	70	Dzhunov, Ivan	56, 57		
Brunetti, Riccardo	25				



<b>G</b>		Jones, Mike	73	Lanciotti, Elisa	61	Misurelli, Giuseppe	25	<b>S</b>		<b>V</b>	
Gabriel, Amoros	14	Jose, Salt	14	Lapka, Wojciech	61	Mohammed, Kaci	14	Saiz, Pablo	56, 57	Vangelis, Floros	38
Gaido, Luciano	25	Jouvin, Michel	63	Lee, Felix	34	Monforte, Salvatore	53	Salomoni, Davide	66	Venturi, Valerio	25
Gater, Catherine	71	Jozef, Cernak	45	Lengert, Maryline	72	Murri, Riccardo	33	San, Wayne	34	Verlato, Marco	20
Gauthey, Clément	25	Junghans, Martin	30	Levshina, Tanya	62	<b>N</b>		Sanchez, Jorge-A.	77	Veronesi, Paolo	25, 66
Gemuend, Andre	15	Juozapavicius, Algimantas	60	Lin, Simon	17, 17	Nauroy, Julien	63	Santeramo, Bruno	15	Vicario, Saverio	27
Germain-Renaud, Cecile	63	<b>K</b>		Loesch, Christian	47	Newhouse, Steven	76	Sargsyan, Laura	56	Victoria, Sanchez	14
Gesing, Sandra	29	Karavakis, Edward	56, 57, 64	Loomis, Charles	25, 62	Notarangelo, Pasquale	27	Scardaci, Diego	34	Visser, Tom	38
Gianelle, Alessio	53	Kassel, Gilles	63	López García, Álvaro	52	Nowotka, Michal	57	van Schaik, Barbera	35	Vocu, Michal	53
Giordano, Domenico	64	Kenny, Eamonn	45	Loth, Alexander	63	<b>P</b>		Schovancova, Jaroslava	57, 57	Vogiatzis, Nikos	77
di Girolamo, Alessandro	57, 59	Kenyon, Michael	56	Louridas, Panos	32	Paganelli, Florido	50	Schwichtenberg, Horst	15	<b>W</b>	
Girone, Maria	64	Kewley, John	73	Luca, Tomassetti	24	Paladin, Massimo	64	Sciaba, Andrea	57	Walker, Christopher	21
Gordon, John	58	Kinnunen, Tapani	31	Lueck, Rupert	72	Paolini, Alessandro	66	Sgaravatto, Massimo	53	Walsh, John	39
Graciani, Ricardo	61	Kitowski, Jacek	74	Luke, Howard	47	Petitdidier, Monique	15	Simón García, Álvaro	52, 58	Wolak, Tomasz	43
Grein, Guenter	56	Kocot, Joanna	31	<b>M</b>		Petzold, Andreas	18	Sinclair, Gillian	77	Wynter-Vincent, Naomi	78
Guijarro, Manuel	59	Kokoszkiwicz, Lukasz	56, 57	Maffioletti, Sergio	75	Philippon, Guillaume	63	Sipos, Gergely	19, 19, 20, 28, 78	<b>Y</b>	
<b>H</b>		Kondov, Ivan	26	Magini, Nicolo	64	Poul, Marcel	53	Sitera, Jiri	53	Yen, Eric	17, 34
Hansper, Vera	72	Konstantinov, Aleksandr	46, 46	Mancinelli, Valentina	64	Prelz, Francesco	53	Smirnova, Oxana	50	You, Jinya	34
Harald Barreiro Megino, Fernando	59, 64	Konya, Balazs	50	Manzi, Andrea	48	Prieto-Castrillo, Francisco	36	Spiga, Daniele	35, 64	<b>Z</b>	
Herres-Pawlis, Sonja	29	Korkhov, Vladimir	32	Matteo, Manzali	24	<b>R</b>		van der Ster, Daniel Colin	36, 65	Zangrando, Lisa	53
de la Hoz Santiago, Gonzalez	14	Koukis, Evangelos	32	Matyska, Ludek	53	Raciazek, Jerome	15	Streit, Achim	70	Zappi, Riccardo	25
<b>I</b>		Kouril, Daniel	46, 47, 53	McLennan, Richard	70	Rebato, David	53	Suárez-Ortega, César	36		
Imamagic, Emir	61	Kreuzer, Peter	57	Memon, Mohammad Shahbaz	48	Riedel, Morris	48, 51, 51, 76	Sustr, Zdenek	46, 53		
Ionides, John	29	Krieger, Michael	74	Memon, Ahmed Shiraz	49	Rogge, Carsten	56	<b>T</b>			
Italiano, Alessandro	66	Krinkele, Karina	16	Mezzadri, Massimo	53	Ronchieri, Elisabetta	52, 66	Toth, Simon	53		
<b>J</b>		<b>L</b>		Michal, Procházka	47	Rosende Dopazo, Roberto	52	Trentadue, Raffaello	37		
Jaekel, René	30, 30	Lagana, Antonio	75	Michel, Franck	18	Rotondo, Riccardo	27, 28	Triantafyllidis, Christos	61		
Javier, Sanchez	14	Lamanna, Giovanni	16	Michelotto, Diego	25	Ruda, Miroslav	53	Trompert, Ron	65		
Jones, Bob	72	Lammerts, Evert	33	Miguel, Villaplana	14			Tuckett, David	56, 57, 59		
		Lanciotti, Elisa	60	Millar, Paul	49			Tugores, Antonia	37		





# Users and communities

This track brings together a series of sessions that focus on activity within specific discipline-based research communities as well as communication and coordination activity across all of the communities as whole. Representatives from the various communities will again present overviews of their current work during the double session on Thursday afternoon. In addition to a presentation on EGI's services specifically developed for users and communities, there will be varied selection of success stories presented within the Community-tailored Services sessions on Tuesday.

The track commences on Tuesday morning with an overview of the various community-focused Virtual Team projects that have been established by EGI with participation from the NGIs. Finally, the most far-reaching of the track's events should be the ESFRI Cluster projects workshop on Tuesday afternoon where EGI's service offerings will be evaluated against the emerging needs of the four cluster projects representing the infrastructure needs of the ESFRI projects.

*Steve Brewer* Chief Community Officer, [EGI.eu](http://EGI.eu)



## ScalaLife – providing scalable software services and expertise to Life Sciences

This submission presents a new EU project, ScalaLife. One of the main goals of the project is to provide a long-term support for the computational life science communities.

**Rossen Apostolov**  
rossen@kth.se

## Added value of new features of the ATLAS computing model and a shared Tier-2 and Tier-3 facilities from the community point of view

Originally the ATLAS computing model assumed that the Tier-2s of each of the ten 'clouds' would keep on disk collectively at least one copy of all 'active' AOD and DPD datasets. Evolution of ATLAS computing and data models requires changes in ATLAS Tier-2s policy for data replication, dynamic data caching and remote data access.

In this way Tier-1s and Tier-2s are becoming more equivalent from a networking perspective and the hierarchy of Tier-1 and 2 is no longer so important.

This talk will present the usage of Tier-2 resources in different grid activities, caching of data at Tier-2s, and their role in the analysis in the new ATLAS computing model. The Tier-3s in the US and the Tier-3s in Europe are rather different because in Europe we have facilities, which are Tier-2s with a Tier-3 component (Tier-3 with a co-located Tier-2). In our infrastructure, this model has in addition adopted sharing ATLAS Tier-2 and co-located Tier-3 facilities, which provides a different approach from for example the Tier-3 model in the US.

**Gonzalez de la Hoz Santiago**  
santiago.gonzalez@ific.uv.es

**Salt Jose**  
jose.salt@ific.uv.es

**Fernandez Alvaro**  
alvaro.fernandez@ific.uv.es

**Sanchez Javier**  
javier.sanchez@ific.uv.es

**Villaplana Miguel**  
miguel.villaplana@ific.uv.es

**Oliver Elena**  
elena.oliver@ific.uv.es

**Kaci Mohammed**  
mohammed.kaci@ific.uv.es

**Sanchez Victoria**  
victoria.sanchez@ific.uv.es

**Amoros Gabriel**  
gabriel.amoros@ific.uv.es

## SuperB evaluation of Dirac distributed infrastructure

The recently founded Nicola Cabibbo Lab will host the SuperB experiment: an asymmetric energy e+e- collider and detector that will provide a uniquely sensitive probe of New Physics in the flavour sector of the Standard Model. SuperB distributed computing group performed a detailed evaluation of DIRAC Distributed Infrastructure in terms of service capabilities, efficiency and reliability for two main use cases: end-user analysis and Monte Carlo simulation production. The new Dirac release 6 has

been configured to respond to SuperB requirements overall the majority of Dirac functionalities. Data management different setup has been considered: the native Dirac File Catalogue system against LFC (LHC File Catalogue) in terms of features, reliability and performance. Testbed and configuration descriptions have been reported including test and evaluation results, using sites that belong to the EGI distributed infrastructure and sites that exploit the OSG middleware.

**Giacinto Donvito**  
giacinto.donvito@ba.infn.it

**Armando Fella**  
armando.fella@pi.infn.it

**Bruno Santeramo**  
bruno.santeramo@ba.infn.it

## Earth Science VRC activities

This session will present the activities of the Earth Science Grid Community on and around the European Grid Infrastructure. It will introduce the activities and future goals as well as current results. Examples of such developments are tools for web- and console-based data discovery and download, as well as the preparation of a Grid workflow that uses climate model data from Earth System Grid.

**Andre Gemuend**  
andre.gemuend@scai.fraunhofer.de

**Horst Schwichtenberg**  
horst.schwichtenberg@scai.fraunhofer.de

**Monique Petitdidier**  
monique.petitdidier@latmos.ipsl.fr

**Sebastien Denvil**  
sebastien.denvil@ipsl.jussieu.fr

**Jerome Raciazek**  
jerome.raciazek@ipsl.jussieu.fr



## Development and deployment of distributed e-VLBI components

Engineering Research Institute Ventspils International Radio Astronomy Centre of Ventspils University College (VIRAC) is a participant in several Very Large Baseline Interferometry (VLBI) networks, notably European VLBI Network (EVN). Data processing is of high significance for whole VLBI measurement and it requires a considerable amount of computational power.

Accordingly, a data processing system was developed, which computes observations results on several High Performance Computers. Recent trends are towards real time data processing, so called e-VLBI, with several advantages when compared to traditional delayed data processing, for example, allowing dynamic detection of transient events in skies.

**Karina Krinkele**  
*karina.krinkele@venta.lv*

## e-Science on weather simulation using the WRF model

Climate change has become a world-wide topic today. The Weather Research and Forecasting Model (WRF), a mesoscale numerical weather prediction model, is today's most advanced and most widely used atmospheric model. By taking advantage of worldwide distributed computing resources, WRF could run the weather simulation model to support the common methodology in a more accurate and efficient manner. We engage in the e-Science collaboration and activity on the global meteorology virtual research environment.

**Simon Lin**  
*simon.lin@twgrid.org*  
**Eric Yen**  
*eric.yen@twgrid.org*

## CTACG – Cherenkov Telescope Array Computing Grid

The Cherenkov Telescope Array (CTA) is an array of many tens of Cherenkov telescopes deployed on an unprecedented scale. One of the challenges in designing the CTA observatory is to handle the large amount of data generated by the instrument and to provide simple and efficient user access at any level and according to astrophysical standards in order to serve the data and

the software for data analysis to the physics community. This project is driven by the CTA consortium comprising 25 countries. The CTA Computing Grid (CTACG) project uses grid technology to perform heavy Monte Carlo simulations and investigates the potential of grid computing for future CTA data management aiming to serve a large worldwide scientific virtual research community.

**Giovanni Lamanna**  
*lamanna@lapp.in2p3.fr*

## High-throughput interactive in-silico drug discovery platform: GVSS-2

Molecular docking simulation is a time consuming process to search exhaustively all conformations of a compound. However, the massive in silico process has benefitted from high-throughput computing technology. To provide intensive computing power and effective management, we developed an in silico drug discovery platform – GVSS (GAP Virtual Screening Service) by

taking advantage of global e-science infrastructures such as EUAsiaGrid and EGI. To improve the accessibility of the distributing computing resources, service grid and desktop grid, ASGC developed GVSS-2 empowered by GAP-2 with new functions of an automation of the pre-requisite process and adding a new protein structure database.

**Simon Lin**  
*simon.lin@twgrid.org*



## Achievements and perspectives of the biomed technical team

The Life-Science Grid Community (LSGC) gathers five Virtual Organisations (VOs) related to the Life-Science field. Among those, the biomed VO has set up a technical support team to be the technical interface between VO users and the NGI sites providing computing, storage and infrastructure services.

The support team is the front-line for handling requests from sites and users; it monitors the resources, and enforces a set of pro-active measures to improve the service quality. It liaises with EGI-InSPIRE instances (UCB and UCST) to report needs,

share experience, and learn from other communities' experience and best practices. This abstract presents those goals in more detail, the achievements of the past year, and highlights current actions and challenges. Overall, the handling of technical issues in an international VO still requires substantial manpower, to the expense of domain-specific activities. Exciting challenges remain to bring the full potential of grids to the end-users.

**Franck Michel**  
*fmichel@i3s.unice.fr*

## The evolution of the EGI user support platform

One of the main objectives of the EGI-InSPIRE project is to increase the use of the European Grid Infrastructure by lowering the barriers to adoption for new communities. The intensive and efficient reuse of EGI assets is a key factor for reaching this goal. The project selected a set of tools that can facilitate the reuse of tangible assets within the community. These tools—namely the Applications Database, the Training Marketplace, VO Services and Requirement Tracker—help the community to monitor and keep track of relevant applications, developer tools, developers, training materials, sites and

services and to reuse these for custom purposes. The talk presents the latest features of these services, and how they contribute to the EGI User Support Platform, a framework that enables support teams to tailor EGI services to the needs of their clients.

**Gergely Sipos**  
*gergely.sipos@egi.eu*

## Experiment representation at the WLCG Tier-1 centre GridKa

The GridKa Computing Centre at the Karlsruhe Institute of Technology is one of the biggest Tier-1 centres for the Worldwide LHC Computing Grid and one of the major resource providers in NGI-DE.

In 2010, three local VO representative positions were established. The representatives' duties are to represent both their respective LHC experiment at GridKa and GridKa within their respective LHC experiment. This presentation will focus on the representatives' technical and communication tasks and their experiences gained.

**Andreas Petzold**  
*andreas.petzold@kit.edu*

## Outreach to new EGI communities

In 2011 the EGI-InSPIRE project revised its user engagement activities and defined a new structure for the NA2 work package within an updated Description of Work. The purpose of the restructuring is to improve the efficiency and flexibility of the interaction between the NGIs, EGI.eu and other organisations to achieve common goals in the field of 'Engaging with New User Communities'. The new user community engagement activity consists of two key elements: Virtual Teams and NGI International Liaisons.

Virtual Teams (VTs) bring together individuals from NGIs and EGI.eu to carry out

focused, short term activities that directly or indirectly relate to engagement with new users. NGI International Liaisons (NILs) are the interfaces within the NGIs who integrate national resources into multi-national outreach activities. The talk provides an overview of the VT-NIL outreach model and presents the outcome obtained and lessons learnt from the first set of Virtual Teams.

**Gergely Sipos**  
*gergely.sipos@egi.eu*



## EGI user requirements – processes, tools and achievements

The evolution of the European Grid Infrastructure is driven by the users. Therefore, capturing and communicating feedback from users to the infrastructure as well as technology operators and providers is a key goal for the EGI-InSPIRE project and for the EGI community as a whole. A requirement engineering and tracking process was put in place by the EGI.eu User Community Support Team and Operation teams during the first year of the project. The process integrates the User Community Board, Operations Management Board and Technology Community Board with user, operation and technology provider

communities, projects and teams. Software tools, such as the EGI Requirement Tracking system (RT), RT web gadgets, wiki pages are in place to support the process stakeholders. The presentation will provide an overview of the EGI requirement engineering process and the achievements in user requirement gathering, analysis and solutions during the last 12 months of the project.

**Gergely Sipos**  
*gergely.sipos@egi.eu*

## Support for non-LHC VOs by GridPP in the UK

GridPP provides grid accessible resources across 19 institutes. While the primary task of GridPP is to support the Large Hadron Collider experiment Virtual Organisations (VOs), it allows much wider access with dedicated fairshares of 1% to non-LHC VOs. In this paper we present an overview of the

non-LHC VOs using GridPP resources and highlight the work done, by GridPP and the VOs, to improve their use of the infrastructure. The paper explores common issues encountered by these communities and describes what they have done, and are currently doing, to overcome challenges faced.

**Christopher Walker**  
*c.j.walker@qmul.ac.uk*

## Extending WeNMR e-infrastructure outside Europe

WeNMR brings together research teams in the Structural Biology area into a Virtual Research Community at a worldwide level, focusing on biomolecular Nuclear Magnetic Resonance (NMR) and Small Angle X-ray Scattering (SAXS). This has been achieved through the implementation of a grid based e-infrastructure, now fully integrated into EGI, aimed at providing the user community with a platform integrating and streamlining the computational approaches necessary for NMR and SAXS data analysis and structural modelling. Nowadays more than 20% of the users come from outside Europe. Therefore,

in the last year, the WeNMR team has worked to extend the e-infrastructure with new resource centres from South Africa, Latin America, North America, Mediterranean and Asia-Pacific areas, with the goal of involving their NGIs to support their local NMR/SAXS community. In particular, a program to interoperate with the OSG grid in US in collaboration with their SBGrid Virtual Organisation has started.

**Marco Verlato**  
*marco.verlato@pd.infn.it*



# Software services for users and communities

Users increasingly expect software applications to be delivered as customised, online services with high availability and simple interfaces. This track brings together a series of sessions that focus on such services coming from discipline-based research communities as well as from multi-disciplinary software developer teams.

The track commences on Tuesday with a double session of various community-specific solutions to support data intensive applications, followed by a third, discussion session about the same topic. Three sessions on Wednesday will study the federated cloud architecture proposal of EGI from different perspectives, while a parallel session goes into details of workflow solutions. On Thursday the track will feature three sessions of presentations on portals and gateways, an increasingly popular and rich topic to simplify access to e-infrastructures. The track is concluded on Friday with a session of presentations on community-specific EGI services.

*Gergely Sipos* Senior User Community Support Officer, EGI.eu



## MILU, the three middleware user interfaces

The Miramare Interoperable Lightweight User Interface (MILU) is a unified command-line client software providing access to: gLite, ARC, GT4 Grids aimed at scientist end-users. MILU is a Plug and Play software supported on a variety of Linux distributions. It comes with a complete configuration for several VOs; a single command allows switching among them. MILU allows concurrent usage of different grid infrastructures from a single Linux host; and thus is an ideal foundation for more complex scientific gateways.

In this talk we shall briefly survey the architecture of MILU, the main technical challenges in implementing it, and report on its usage experiences gathered in the EUIndia and e-NMR communities and as a production user interface.

**Tyanko Aleksiev**  
*tyanko.aleksiev@oci.uzh.ch*

## Exploiting grid resources for data simulation by using a general purpose framework

We present a general purpose software framework, which allows different multi-disciplinary communities to take advantage of a distributed computational infrastructure. It has been designed specifically for organisations that cannot afford the adoption costs of more specialised and complex frameworks, but that still require a user-friendly, standard and highly customisable access to the grid.

Our framework heavily relies on a bookkeeping database, storing both application-specific and infrastructure metadata, which is tightly coupled with a web portal. The first makes available to the users information on the execution status of jobs and their specific meaning and parameters, and contributes in orchestrating the submission mechanism. The latter provides job submission management, bookkeeping database interactions and monitoring functionality.

**Fella Armando**  
*armando.fella@pi.infn.it*  
**Luppi Eleonora**  
*eleonora.luppi@fe.infn.it*  
**Manzali Matteo**  
*matteo.manzali@fe.infn.it*  
**Tomassetti Luca**  
*luca.tomassetti@fe.infn.it*  
**Vianello Enrico**  
*enrico.vianello@fe.infn.it*

## A general purpose grid and cloud portal to simplify scientific communities' integration into distributed computing infrastructures

In the framework of the Italian Grid Infrastructure, we have designed a web portal which aims at providing a powerful and easy to use gateway to distributed computing resources and cloud services provisioning.

We feel that one of the key points that this kind of application must be able to address is the possibility to hide the complexity of the X509 certificates management.

Through the portal, users can securely run their applications without facing the complexity of the authentication infrastructure. It also interfaces with external Infrastructure as a Service (IaaS) frameworks for the dynamic provisioning of computing resources.

**Marco Bencivenni**  
*marco.bencivenni@cnaf.infn.it*  
**Paolo Veronesi**  
*paolo.veronesi@cnaf.infn.it*  
**Luciano Gaido**  
*luciano.gaido@to.infn.it*  
**Valerio Venturi**  
*valerio.venturi@cnaf.infn.it*  
**Diego Michelotto**  
*diego.michelotto@cnaf.infn.it*  
**Giuseppe Misurelli**  
*giuseppe.misurelli@cnaf.infn.it*  
**Riccardo Zappi**  
*riccardo.zappi@cnaf.infn.it*  
**Riccardo Brunetti**  
*riccardo.brunetti@to.infn.it*  
**Andrea Ceccanti**  
*andrea.ceccanti@cnaf.infn.it*

## Cloud, ready for bioinformatics?

The overall strategy for increasing the users community of the StratusLab infrastructure is to create a 'virtuous cycle', starting with porting selected use cases to demonstrate the added value, and two of them are related to bioinformatics. The first study, 'Bioinformatics Services', is focused on creating bioinformatics appliances providing tools and data related to sequence analysis that scientists and engineers can deploy on demand. The second usecase, 'TOSCANI: Towards

StruCTural AssignmeNt Improvement' is a project with a collaboration between CNRS and Institut Pasteur to improve the determination of protein structures based on Nuclear Magnetic Resonance (NMR) information. These use cases demonstrate the ease of use of a cloud infrastructure that stays connected with the existing public bioinformatics resources. First releases of the required bioinformatics appliances have been developed and made available to the scientists via the StratusLab Marketplace.

**Christophe Blanchet**  
*christophe.blanchet@ibcp.fr*  
**Clément Gauthey**  
*clement.gauthey@ibcp.fr*  
**Charles Loomis**  
*loomis@lal.in2p3.fr*





## Workflows for the simulation of organic light-emitting diodes

Computational materials science is essential for development of products with novel properties. With the accelerating development cycles of materials, the development of simulation approaches for predictive, de-novo characterisation and optimisation of materials and device properties emerges as a grand challenge to European Research and Development. A unified multi-disciplinary approach towards the deployment of models, tools, algorithms and simulation and visualisation techniques is required to transform isolated solutions for specific problems into comprehensive, industry-ready platforms, which

are capable of predicting the properties of complex materials on the basis of their constitutive elements. The project MMM@HPC (Multiscale Materials Modeling on High Performance Computer Architectures) will demonstrate a direct impact towards solving this challenge in the areas of energy storage, e.g. Li ion batteries, and energy conversion, e.g. organic light emitting diodes (OLEDs).

**Stefan Bozic**  
*stefan.bozic@kit.edu*

**Ivan Kondov**  
*ivan.kondov@kit.edu*

## A grid execution model for computational chemistry applications using GC3Pie and AppPot

The increasing availability of computer power on grid platforms has prompted the implementation of complex computational chemistry suites of codes on distributed systems and, at the same time, the development of appropriate frameworks allowing to define event-related dependencies between applications (which may involve different computer environments) in a workflow-like fashion. The work carried out to implement complex computational chemistry usecases on distributed systems making use of the high throughput execution framework GC3Pie and the AppPot cloud/grid virtual

machines (both developed by GC3, University of Zurich) is here presented and discussed. The work has been carried out in collaboration with the User Support Unit of the Italian Grid Initiative (IGI) which tackled the part related to the porting of the single applications.

**Alessandro Costantini**  
*alex.costantini@gmail.com*

## The Biovel Project: robust phylogenetic workflows running on the grid

Altered species distributions, the changing nature of ecosystems and increased risks of extinction all have an impact in important areas of societal concern. Biologists and environmental scientists are asked to provide decision support for managing the biodiversity component of our environment at multiple scales (genomic, organism, habitat, ecosystem, landscape) to prevent and mitigate such losses.

Biovel want to address this need, offering a series of robust and reliable web

services that could be managed with the suite of tools of the myGRID project. The project will propose workflows for these services that will ensure best practice and efficiency of usage.

Within the first round of services produced by the project there are phylogenetic inference workflows.

These workflows will provide to end-users the capabilities to execute applications that could easily exploit several kind of resources, like the EGI grid infrastructure, local batch farms or dedicated servers.

**Giacinto Donvito**  
*giacinto.donvito@ba.infn.it*

**Saverio Vicario**  
*saverio.vicario@gmail.com*

**Pasquale Notarangelo**  
*pasquale.notarangelo@ba.infn.it*

## A data engine for grid science gateways enabling easy transfer and data sharing

Grid infrastructures allow users to access and use computational and storage facilities distributed in different locations all around the world.

However, a real exploitation of such platforms by large communities is still not happening mainly because of the complex architecture of the Grid Security Infrastructure and of the command line based interface that turn out to be quite cumbersome for the vast majority non IT-expert users.

In this work we show a new SG component providing users with the ability to move data to storage elements and share them in an easy and clever way.

**Marco Fargetta**  
*marco.fargetta@ct.infn.it*

**Riccardo Rotondo**  
*riccardo.rotondo@garr.it*

**Roberto Barbera**  
*roberto.barbera@ct.infn.it*





## The grid goes 'social': extending authentication to members of social network

Grid infrastructures are being built in several areas of the world but, despite the huge investments made by the European Commission and by other funding agencies, both at national and international level, the total number of users is in the order of magnitude 104. This is lower than the number of users of the international research and education networks (e. g., GÉANT in Europe) that have been funded more or less at the same level of grids (about 107), and definitely less than the order of magnitude of users accessing the social networks (e.g. facebook and Google).

Therefore, to increase the use of grid and involve new users it is important link grid services with other research and education, or social services, creating a unified working environment.

The first step towards this new environment is a unified authentication system.

We report the experience of developing an authentication infrastructure allowing access to grid SGs to social users.

**Marco Fargetta**  
*marco.fargetta@ct.infn.it*

**Riccardo Rotondo**  
*riccardo.rotondo@garr.it*

**Roberto Barbera**  
*roberto.barbera@ct.infn.it*

## A generic gateway and its customisation in different scientific areas

Science gateways (or portals) enable user communities to exploit distributed computing infrastructures (DCI) through simple, user-friendly interfaces. Different functionalities may be used through different portlets implementing the desired feature. A number of portal technologies are available for use by scientists or for deployment by system administrators of a user community. Usually, these portal technologies or installation serve the needs of a specific user community. Within our talk we present the SCI-BUS project, which aims to bring together a big number of scientific areas under a generic core portal technology

called WS-PGRADE/gUSE, that provides access to a number of Distributed Computing Infrastructures through the DCI Bridge and offers exploitation of SaaS cloud services through the CloudBroker platform. The project offers a customisation methodology of WS-PGRADE, that enables creating application-specific portals for the community's needs.

**Zoltan Farkas**  
*zfarkas@sztaki.hu*

**Gergely Sipos**  
*gergely.sipos@egi.eu*

## The MoSGrid community—from national to international scale

The project MoSGrid (Molecular Simulation Grid) offers a web-based science gateway supporting the community with various services for molecular modelling, docking, and quantum chemistry. Users gain access to distributed computing infrastructures (DCIs) via intuitive user interfaces for sophisticated tools, specialised workflows, and distributed repositories. Currently, the MoSGrid community consists of over 50 research groups from a number of fields related to chemistry and bioinformatics located in Germany. However, the underlying security infrastructure is not limited

to national requirements. It can be easily configured for international users.

MoSGrid intends to address the international community by participating in the EU project SCI-BUS (Scientific gateway Based User Support), collaborating with the EU projects SHIWA (SHaring Interoperable Workflows for large-scale scientific simulations on Available DCIs), and EDGI (European Desktop Grid Initiative).

**Sandra Gesing**  
*sandra.gesing@uni-tuebingen.de*

**Sonja Herres-Pawlis**  
*sonja.herres-pawlis@cup.uni-muenchen.de*

## Workflow and data management for nuclear magnetic resonance.

We are developing a web-based workflow and data management system targeted at scientists within the field of nuclear magnetic resonance (NMR). NMR is a key method for investigating biological macromolecules, and analysis of NMR data currently involves both manual and computational steps. For the computational stages there is a strong tradition of using distributed computing methods, as the popularity of the WeNMR VRC testifies. However, presenting these

methods so that they can be easily built into 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 workflows that can address important scientific problems—particularly in the area of improving automated NMR data analysis—as well as improved accessibility to WeNMR services and the potential for improved lab data management.

**John Ionides**  
*john.ionides@gmail.com*



## Interactive information extraction based on distributed data management for D-Grid projects

The current D-Grid infrastructure primarily covers resource management and exchange at the data level supporting mainly technical resources such as computational capacity, data transport networks, storage resources, and management software.

The WisNetGrid project ([www.wisnetgrid.org](http://www.wisnetgrid.org)) aims to broaden the focus of resources exchange towards the actual content, such as research and production data, to enable interdisciplinary usage. To achieve this goal, resource sharing is supported on different abstraction layers. First, we create an information

layer by providing a universal interface to access data on the grid independent of the underlying grid storage system. Second, at the knowledge layer, we offer highly interactive knowledge extraction and management tools that can also take advantage of a community's grid resources.

**René Jaekel**  
[rene.jaekel@tu-dresden.de](mailto:rene.jaekel@tu-dresden.de)

## WisNetGrid service layer – enabling creative research workflows

Many projects in academia and industry nowadays use and provide services electronically in order to be able to select from or target to a larger community across regional boundaries, respectively. The demand for flexibility and adaptability arises with altering user requirements and evolving amount of functionalities within the field of research computing. While flexibility and

adaptability are achieved by service-oriented architectures, the heterogeneity between resources of different domains is addressed by formally modelling services and workflows as well as the attached information resources. The degree of automation and flexibility is further increased by a set of tools for searching, ranking, and composing services and workflow even at run-time.

**Martin Junghans**  
[junghans@kit.edu](mailto:junghans@kit.edu)

**Sudhir Agarwal**  
[agarwal@kit.edu](mailto:agarwal@kit.edu)

**René Jaekel**  
[rene.jaekel@tu-dresden.de](mailto:rene.jaekel@tu-dresden.de)

## SOMA2 – application oriented molecular modelling workflows in www-browser

SOMA2 gateway is a www-browser operated molecular modelling workflow environment developed by CSC – IT Centre for Science Ltd. The SOMA2 environment allows users to combine scientific applications into unique application workflows, which are automatically executed in the underlying computing system. SOMA2 offers a flexible framework for integrating and executing

molecular modelling applications, and facilitates automated molecular data exchange. SOMA2 source code is distributed under the GPL open source license. Since version 1.4 SOMA2 has support to make use of grid resources with suitable middleware software. Grid-related development work has been part of EGI-InSPIRE project's WP6-SA3.

**Tapani Kinnunen**  
[takinnun@csc.fi](mailto:takinnun@csc.fi)

## InSilicoLab: a user workspace implementing e-science principles

E-science is a challenging vision that still has not been fully achieved in contemporary research. Among its principles, collaboration between globally dispersed groups of scientists and use of large-scale computing resources and large data collections are usually pointed out as most important. Modern infrastructures, like grids, already meet many of these requirements. However, for full implementation of the e-science vision, tools that go further – towards support of collaboration environments capable of extracting knowledge from data obtained with use of the infrastructures – are needed.

InSilicoLab is an application portal that supports these aspects of research by facilitating access to computational software deployed on grids and management of data and processes involved in scientific computations. The portal provides mechanisms to track the processes that lead to valuable data, as well as for sharing these data and processes with fellow scientists enabling their collaborative work.

**Joanna Kocot**  
[ymkocot@cyfronet.pl](mailto:ymkocot@cyfronet.pl)



## Using SHIWA workflow interoperability tools for neuroimaging data analysis applications

Neuroimaging is a field that benefits from Distributed Computing Infrastructures (DCIs) to perform data processing and analysis, which is often achieved using grid workflow systems. Collaborative research in neuroimaging requires ways to facilitate exchange between different groups, in particular to enable sharing, re-use and interoperability of applications implemented as workflows. The SHIWA project provides solutions to facilitate sharing and exchange of workflows between workflow systems and DCI resources. In this work we analyse how the SHIWA platform is used to implement various

usage scenarios in which workflow exchange supports collaboration in neuroscience. The SHIWA platform and the implemented solutions are described and analysed from a 'user' perspective of workflow developers and neuroscientists. We conclude that the platform is valuable for the foreseen usage scenarios, and we identify remaining challenges to be addressed in future SHIWA platform releases.

**Vladimir Korkhov**  
*v.korkhov@amc.uva.nl*

## Okeanos IaaS

Okeanos is an IaaS offering virtualised computing resources. It is developed by GRNET (the Greek Research and Technology Network), to be offered to the whole Greek research and academic community. The software powering Okeanos is available via an open source licence. Okeanos offers to its users, through a simple web-

based GUI, access to Virtual Machines, Virtual Ethernets, Virtual Disks, and Virtual Firewalls. Okeanos was conceived to offer its users easy and secure access to GRNET's data centres, focusing on user friendliness and simplicity, while being able to scale up to the thousands (of Virtual Machines, users, terabytes of storage).

**Evangelos Koukis**  
*vkoukis@grnet.gr*  
**Panos Louridas**  
*louridas@grnet.gr*

## Data-intensive processing with Hadoop and friends at BiG Grid

Hadoop is a framework built to support large-scale data-intensive processing. In spite of its age—it is only four years old—it has gained popularity amazingly fast and is currently deployed in both science and industry at unprecedented scales. Hadoop is an open source implementation of a system designed by Google, including the architecture of the distributed file system in use at the company, and the parallel processing model used by its in-house scientists.

In this talk we present the current Hadoop pilot service of SARA in The Netherlands, and the way we will make it operational within the Dutch

NGI, BiGGrid. We will talk about typical work being done on a Hadoop system, sciences that benefit from such a system, the types of tools it provides, and how it enables e-science. The talk is intended for both community representatives and NGI representatives. Our goal is to show communities what can be achieved, and to inspire NGI's in their approach to diversity of infrastructure.

**Evert Lammerts**  
*evert.lammerts@sara.nl*

## AppPot: bridging the grid and cloud worlds

AppPot is a system for creating Linux virtual machines that can run with most virtualisation software (KVM, Xen, VirtualBox and UML are supported) plus a set of scripts for controlling the execution of non-interactive computational programmes in the VM.

Indeed, thanks to the use of 'User-Mode Linux' virtualisation, AppPot VMs can run on grid clusters as regular grid jobs, without the need for sysadmin support or root access. In particular, AppPot is a way to implement generic application deployment on a computational grid, and especially to enable users to

provide their own software to the computing cluster: users can use an AppPot VM on their computer for coding, and then run the same VM as a grid job or in a Cloud IaaS infrastructure for production runs.

In this talk we give an overview on AppPot and its features, the technology that makes it possible, and report on experiences running it in production use within the NGI-CH SMSCG infrastructure.

**Riccardo Murri**  
*riccardo.murri@uzh.ch*



## Cloud federation over the world wide grid

e-science is able to be realised by distributed resource sharing and collaboration over the worldwide grid (WWG), such as EGEE and EGI. With the advent of virtualisation and service-oriented systems such as the cloud, integration of WWG and cloud technology (World-Wide Cloud, WWC) could provide a better service granularity for variant user requirements, especially the scientific community.

Cloud federation, both to provide cloud services across sites and to support federated services among cloud service providers, is achieving not just resource-level elasticity but also service level reconfiguration and repurposing. With the institution or community-based cloud federation model, the WWC could minimise the resource access barrier and shorten the time to finish jobs without limitations in scalability.

**Wayne San**  
*waynesan@twgrid.org*

**Jinya You**  
*jingya.you@twgrid.org*

**Felix Lee**  
*felix@twgrid.org*

**Eric Yen**  
*eric.yen@twgrid.org*

**Jinny Chien**  
*chen.yi.chien@cern.ch*

**Tim Chen**  
*yuting.chen@twgrid.org*

## Challenges in data management and running DNA sequencing experiments on grid

Modern DNA sequencing machines produce data in the range of 1–100 GBytes per sample and with ongoing technological developments this amount is rapidly increasing. The majority of experiments involve re-sequencing of human genomes and exomes to find genomic regions that can be associated with disease. In the bioinformatics field the development of analysis software for DNA sequencing experiments progresses rapidly. There are many analysis tools freely available, e.g. for sequence alignment, quality control and variant detection, and frequently new tools are developed.

We use workflow technology to allow easy incorporation of such software in our data analysis pipelines, and analysing multiple data sets at once. Since the end of 2008 we have been using our framework, called e-BioInfra, for various DNA experiments on the Dutch grid. Here we will present our current procedure for analysing DNA experiments, comment on the experiences and focus on improvements that were needed for analysing genomics data.

**Barbera van Schaik**  
*b.d.vanschaik@amc.uva.nl*

## The VRC-driven GISELA Science Gateway

GISELA (Grid Initiatives for e-Science virtual communities in Europe and Latin America), is a project co-funded by the European Union under its Seventh Framework Programme. GISELA aims at: (1) Implementing the Latin American Grid Initiative (LGI) sustainability model rooted on National Grid Initiatives (NGI) or Equivalent Domestic Grid Structures (EDGS), in association with CLARA, the Latin American NREns and collaborating with the European Grid Initiative (EGI); (2) Providing Virtual Research Communities (VRCs) with the e-infrastructure and application-related services required to improve

the effectiveness of their research. In order to extend its user base and make grid access and use easy for non IT-experts, GISELA has developed a VRC-driven Science Gateway. A Science Gateway is a community-developed set of tools, applications, and data collections that are integrated via a portal or a suite of applications.

**Diego Scardaci**  
*diego.scardaci@ct.infn.it*

**Roberto Barbera**  
*roberto.barbera@ct.infn.it*

## CRAB: a user friendly application for distributed data processing for the Compact Muon Solenoid experiment at the LHC

The CMS Remote Analysis Builder (CRAB) application addresses the needs of the CMS community, allowing the users to easily access the grid resources. CRAB interacts with the local user environment, the Data Management services and with the grid middleware, limiting the knowledge of the technical details required of the end user. CRAB has progressed from a limited initial prototype nearly five years ago, to a system heavily employed by the whole CMS collaboration to prepare over 100 analysis papers. CMS observes more than 400 unique users submitting CRAB jobs per week, with close to

1000 individuals per month. Up to 200,000 CRAB jobs per day run on the grid.

The CRAB team has an ambitious programme planned in 2012: to release a new generation of CRAB that aims to make a step towards an SaaS architecture. This work will present the joint CMS experiment and CERN IT-ES effort to realize such a project, highlighting the impact on the service maintenance and first experiences dealing with beta users.

**Daniele Spiga**  
*daniele.spiga@cern.ch*



## Building a 'grid of clouds' – how one HEP experiment is evaluating strategies to incorporate 'The Cloud' into existing grid infrastructures

Emerging standards and software often marketed as 'Cloud Computing' bring attractive features to improve the operations and elasticity of scientific distributed computing. At the same time, the existing European Grid Infrastructure and Worldwide LHC Computing Grid (WLCG) have been highly customised over the past decade or more to the needs of the VOs and are operating with remarkable success. It is therefore interesting not to

replace 'The Grid' with 'The Cloud', but rather to consider strategies to integrate cloud resources, both commercial and academic, into the existing grid infrastructures, thereby forming a 'grid of clouds'. This work will present the efforts underway in the CERN IT experiment support group along with the ATLAS experiment to adapt existing grid workload and storage management services to cloud computing technologies.

**Daniel van der Ster**  
*daniel.colin.vanderster@cern.ch*

**Fernando Barreiro Megino**  
*fernando.harald.barreiro.megino@cern.ch*

## gridCake and gridCamp: making the grid easier

The grid is a very useful technology for researchers, but its front-ends aren't very user-friendly because they are usually designed for users with computational skills. This fact could cause problems to end-users because they, generally, want to focus their efforts on their research and avoid learning concepts not related with their studies.

Furthermore, the existing tools for software development that interacts with grid services are complex, generally poorly documented and

don't support the latest grid middleware versions. This makes it very difficult to develop user interfaces for grid services.

This work deals with the limitations of existing grid user interfaces and development tools. It describes the development of a very easy to use Java grid library called gridCake. Finally, a grid web portal, called gridCamp, is presented. GridCamp uses gridCake, and mixes grid resources with social network features.

**César Suárez-Ortega**  
*cesar.suarez@ciemat.es*

**Joé Miguel Franco-Valiente**  
*josemiguel.franco@ciemat.es*

**Francisco Prieto-Castrillo**  
*francisco.prieto@ciemat.es*

## Handling of network and database instabilities in CORAL

The Large Hadron Collider (LHC), the world's largest and highest energy particle accelerator, started its operations in September 2008 at CERN, Switzerland. Huge amounts of data are generated by the four experiments installed at different collision points along the LHC ring. The largest data volumes come from the 'event data' that records the signals left in the detectors by the particles generated in the LHC beam collisions and are generally stored on files. Relational database systems are commonly used instead to store the 'conditions data' that records the geometry, configuration and other

working parameters of the detectors at the time the event data were collected.

The Common Relational Abstraction Layer (CORAL) software is widely used by the LHC experiments for storing and accessing conditions data using relational database technologies. This talk looks at handling network and database instabilities in CORAL.

**Raffaello Trentadue**  
*raffaello.trentadue@cern.ch*

## Web4Grid, web interface for grid jobs

Grid has the potential to be a key instrument for a wide variety of scientific topics which require many calculations to be performed. The reduced size of the research groups or the diversity of problems makes it unsuitable to develop specific interfaces in most of the instances and the grid shell commands are a barrier that requires a significant amount of determination to cross.

To make grid usage as common as computational clusters, user-friendly interfaces are needed, and Web4Grid has been developed in the framework of Ibergrid to allow easy access to the grid infrastructure for all the Portuguese and Spanish grid users.

**Antonia Tugores**  
*antonia@ifisc.uib-csic.es*



## Operating a public cloud with StratusLab

The StratusLab project provides an open-source cloud distribution that allows data centres to expose computing resources as an 'Infrastructure as a Service' (IaaS) type cloud. Administrators can run services over the cloud to improve availability, scalability, and maintainability. The StratusLab project also maintains a reference cloud infrastructure, used by several scientific user communities, and a complete grid site hosted on the cloud.

This presentation describes deployment and operational issues related to provision of clouds for e-Science, based

on the experiences of the StratusLab project. This will include issues with cloud provision within educational and research institutes, as well as with deploying grid services on the cloud. In particular the session will include a discussion on the impact of virtualisation on operating costs for e-Science and grid infrastructures, informed by an analysis of the cost of running the StratusLab reference infrastructure and grid site.

**Floros Vangelis**  
*efloros@grnet.gr*

## Supporting grid-enabled GPU workloads using rCUDA and StratusLab

The increasing capabilities of general purpose graphics processing units (GPGPUs) over the past few years have resulted in a huge increase in their exploitation by all the major scientific disciplines where massively parallel processing capabilities are desired. However, there are two major problems in supporting grid access to such resources: Firstly, there is currently no standardised way for resource centres to advertise/publish availability of these resources. Secondly, there are deficiencies in current batch scheduling systems that ensure exclusive access to those resources.

We present the results of an initial investigation into grid-enabling access to many general purpose graphics processing units (GPGPUs) distributed over a local cluster. We exploit two distinct virtualising technologies—rCUDA and StratusLab. The hybrid approach is used to achieve greater levels of parallelism, and to provide the necessary GPGPU resource isolation.

This is currently a work in progress.

**John Walsh**  
*john.walsh@scss.tcd.ie*

## Beehub, facilitating scientific collaboration

In this presentation we will be explaining and demonstrating a recently launched file storage and sharing service for (Dutch) scientists called Beehub.

**Tom Visser**  
*tom.visser@sara.nl*





# Middleware services

This year's Community Forum is packed with in-depth information about various middleware services that provide access to the underlying distributed computing infrastructure.

The European Globus Community Forum on Monday opens the stage with a full day covering the available Globus services and tools such as Globus online, and discussing the consequences of European Data privacy laws on Grid Middleware development.

Continuing every day right up to Thursday, each day is packed with three dedicated sessions on middleware services provided by EMI, covering Information Services, Data Access and Transfer on Tuesday, batch job management and middleware service operation quality assurance on Wednesday, wrapping up on Thursday with general Software Quality Assurance, Standards in action and practical application development using EMI components.

*Michel Drescher* Technical Manager, [EGI.eu](http://EGI.eu)



## Improvements to the EMI build and test tools

One of the major goals of the EMI is the integration of the components of the pre-existing middleware into a single consistent set of packages with uniform distributions and repositories. The EMI build and release is based on ETICS. During the last year, a lot of changes have been implemented on the development tools side and to ETICS in particular: such as new platforms, infrastructure changes or new QA reports, in order to provide the features needed for the EMI-2 and EMI-3 releases.

**Andres Abad Rodriguez**  
*andres.abad.rodriguez@cern.ch*

## Advanced data staging in the ARC Computing Element

The Advanced Resource Connector's Computing Element (ARC CE) is responsible for staging input and output data for tasks running on the computational resources it manages. This paper presents the newly redesigned staging framework within the ARC CE, which addresses several issues encountered as the data demands of tasks have increased.

**David Cameron**  
*david.cameron@cern.ch*

## EMI Release Management Highlights – EMI 1 Updates and EMI 2 Status

The European Middleware Initiative (EMI) is a close collaboration of four major European technology providers: ARC, gLite, UNICORE and dCache. Its main objective is the delivery of a consolidated and harmonised set of components for deployment in EGI, PRACE and other DCIs production environments.

This presentation will give an overview of the status of EMI products with a special focus on the support and maintenance activities performance during the first year after the EMI-1 Kebnekaise release.

The status of the EMI-2 Matterhorn, scheduled for release at the end of April, 2012, will also be presented, highlighting new features provided in EMI components and the updates of the EMI certification and release processes.

**Andrea Ceccanti**  
*andrea.ceccanti@cnaf.infn.it*

**Cristina Aiftimiei**  
*cristina.aiftimiei@pd.infn.it*

## Advancements in EMI Testing Infrastructure implementation

The EMI Project has succeeded in merging into a single release (EMI-1 Kebnekaise) more than fifty software products from four major European technology providers (ARC, gLite, UNICORE and dCache). To satisfy end-user expectation in terms of functionality and performance, the release process implements several steps of certification and verification. The final phases of certification is aimed at harmonising the strongly inter-

dependent products coming from various development teams through parallel certification paths. The role of the EMI Testing Infrastructure is to provide operational and infrastructural resources to implement inter-component certification phases and involve EMI end-users in early testing or preview activity.

The presentation will describe the EMI Testing Infrastructure and the lessons learned after EMI-1/2 release experience.

**Danilo Dongiovanni**  
*danilo.dongiovanni@cnaf.infn.it*

**Tomasz Wolak**  
*twolak@cern.sh*





## SIENA: grid and cloud standards for e-Science and beyond

The assets and capacities of Europe's Distributed Computing Infrastructures (DCI) can bring value-added services within the context of the Digital Agenda for Europe and empower the broadest possible set of researchers.

This workshop presents the recommendations of the SIENA Roadmap on grid and cloud standards for e-Science and beyond. These discussions follow on from Cloudscape IV, which looked to continue the dialogue between public

sector officials, industry and SDOs over the deployment of cloud based services for the scientific community and public procurement.

The workshop will focus on the work of the DCI projects in advancing the standards landscape and in particular with regard to the adoption of cloud standards such as OCCI, CDMI and OVF.

The workshop will include presentations from members of the Roadmap Editorial Board, DCI projects and SDO representatives.

**Nicholas Ferguson**  
*n.ferguson@trust-itservices.com*

## Performance testing of distributed computational resources in the software development phase

Software harmonisation between different software vendors is possible through adoption of standards (common protocols and interfaces). We have observed that reliable and comparable performance parameters of grid subsystems or large-scale grid infrastructures are not usually known, neither for developers which implement the standards in the existing solutions nor for potential users which are looking for high performance distributed computational solutions.

**Cernak Jozef**  
*jcernak@upjs.sk*

## Quality Control in EMI

The purpose of this presentation is to give an overview of the Quality Control activity in EMI, which is concerned with the verification of released software products, and the monitoring of internal procedures. The Quality Control (QC) pertains to the monitoring of project outcomes to see whether they comply with quality standards set out in the EMI Quality Model or within

internal procedures, such as those concerning the release and packaging of software components. Operating throughout the project, its aim is to identify unacceptable or non-conformable results and inform project executive boards about their existence so that corrective actions can be undertaken to eliminate, or mitigate, future negative impacts on project's results.

**Giuseppe Fiameni**  
*g.fiameni@cineca.it*

## Common framework for extracting information and metrics from multiple change trackers

An important aspect of EMI is the delivery of 'quality software'. For this reason the quality assurance (QA) group was introduced. There are a key number of beneficiaries of this work in a number of work packages. These EMI development and support activities are required to produce Key Performance Indicators (KPIs) and metrics for milestone, quarterly and yearly deliverables based on the information provided by the QA group. However, EMI has a large number of different sized products, different middlewares and various bug/feature requests for change (RfC) trackers used by each product team. The

only way to reliably produce KPIs and metrics related to change management in such a varied project is to introduce simplifying, common environments that are readily accessible by varying numbers of customers of the project. For this reason an extensible XML-based framework was generated for storing, plotting, querying and tabulating change tracker information for all its customers.

**Eamonn Kenny**  
*eamonn.kenny@cs.tcd.ie*



## EMI Execution Service implementation in ARC

Recent development of EMI Execution Service interface (EMI ES) specifications produced rich production-ready definition. To prove its usability and to test its completeness and interoperability capabilities it is important to provide few implementations. This paper represents the implementation of the EMI ES interface in the Advanced Resource Connector's Computing Element (ARC CE) service (A-REX) and client.

**Aleksandr Konstantinov**  
aleksandr.konstantinov@fys.uio.no

## EMI common authentication library

The EMI components use Public Key Infrastructures (PKI) with the SSL/TLS protocol suite for authentication of users and other services they interact with. For historic reasons, most components implement the authentication functions independently on each other, which yield codes with similar functionality being implemented at various places again and again. For the same reasons there is no 'profile' defining details of SSL handshake or verification of certificates. Also the grid specifics, such as processing of X.509 certificate and handling of CA policies, are not implemented in a common way. Such an arrangement

makes it harder to add new features and also maintain the code basis.

EMI has decided to design and implement a common authentication library that implements the main functions necessary to establish an authenticated connection and exchange messages that are sufficiently protected.

**Daniel Kouřil**  
kouril@ics.muni.cz

**Zdenek Šustr**  
sustr4@cesnet.cz

**Krzysztof Benedyczak**  
golbi@icm.edu.pl

**Aleksandr Konstantinov**  
aleksandr.konstantinov@fys.uio.no

**Poul Marcel**  
poul@mail.muni.cz

## Extended grid data management for UNICORE – accessing files in SRM and LFC

Currently, most data in the grid is stored on resources managed by the Storage Resource Manager (SRM). Additionally, many of them are indexed in file catalogues, like the gLite File Catalogue (LFC), as well. One handicap of the UNICORE data management was the impossibility to access files stored on remote grid resources outside of other UNICORE instances like on SRM storages. Thus, one task for EMI is to improve this interoperability issue.

We will present how the UNICORE model of storage handling has been used to implement client access to SRM storage using the LFC. Typical file storage operations like directory browsing and file transfers for the stage-in or stage-out processes in UNICORE jobs are possible now. Files indexed by the LFC can be accessed through an automatic resolution to its physical copies on storage elements, and are transferred transparently to the user. However, there exist some limitations in the implementation due to technical hurdles, that will be pointed out.

**Christian Loeschen**  
christian.loeschen@tu-dresden.de

## Federated access to data storage with Moonshot

Manipulating data of different forms is something we do every day. Since data files often contain sensitive information, users require that they are secured properly. In grids PKI is traditionally used for authentication, however, it does not work with a common filesystem like NFSv4 or CIFS. PKI has also a bad reputation in terms of users experience and people try to leverage other authentication mechanisms, like identity federations. Unfortunately, existing identity federations cannot be used smoothly to access remote filesystems.

In this contribution we will demonstrate how

the Moonshot federated authentication infrastructure can be used to provide easier access to storage systems, where users use their home identities to access file systems provided by other institutions. Unlike other solutions, we do not use a web portal or similar 'translation' service but allow the users to directly mount remote volumes to their computers.

**Howard Luke**  
lukeh@padl.com

**Daniel Kouřil**  
kouril@ics.muni.cz

**Procházka Michal**  
michalp@ics.muni.cz



## The gCube data transfer facilities

The iMarine EU project, started on the 1st November 2011 as a follow up of the D4Science II project, aims to establish and operate a data infrastructure supporting the principles of the Ecosystem Approach to Fisheries Management and Conservation of Marine Living Resources.

The data-infrastructure enabling technology, gCube, developed throughout three different EU projects (DILIGENT, D4Science I and II) is a software framework which enables the declarative and interactive creation of transient Virtual Research Environments that aggregate and deploy

on-demand content resources and application services by exploiting computational and storage resources of grid and cloud infrastructures.

We report on the results of an evaluation of data management solutions from the EMI software portfolio and the experience of collaborating with EMI on the adaptation of existing services to implement data transfer mechanism within the gCube framework, a key element in sustainability for the middleware.

**Andrea Manzi**  
*andrea.manzi@cern.ch*

## Lessons learned from UNICORE EMI-ES Adoption towards improved open standards

The EMI project unites a set of production grid middleware technologies providing scientific communities a secure access to distributed and heterogeneous, compute and data resources. Within the compute area, job management and monitoring are considered to be the most significant areas of work. The EMI compute team embarked upon an effort to review the existing standards and their adoption in the domain of job management and to explore advanced execution service concepts that have been provided in the EMI-ES specification. The goal of this paper is to present the concepts of the EMI-ES

interface and its information model that is required to manage, monitor, and model activities in production Grids. In this paper, we will delineate the architectural details of EMI-ES, and one of its 'proof of concept' realizations in UNICORE. While we intend to contribute this effort to open standards, we will also shed light on the existing grid standards by comparing them with EMI-ES.

**Mohammad Shahbaz Memon**  
*m.memon@fz-juelich.de*  
**Morris Riedel**  
*m.riedel@fz-juelich.de*

## EMIR: An EMI service registry for federated grid infrastructures

The European Middleware Initiative (EMI) is a European project that represents a collaboration of four middleware stacks namely ARC, dCache, gLite, and UNICORE. All these middleware services should be easily deployable in a DCI. However, the immediate challenge is the discovery of those services in a particular DCI that is typically done via so-called registries. This is a major requirement for operational systems, and the middleware itself. Existing registries such as ARC Information Index or UNICORE registry are designed to index middleware specific services. Given the centralised nature,

the scope of these registries can become limited when considering a federated infrastructure that relies on service of different technology providers. DCIs such as EGI are federated in nature, thus a unified registry should reflect this requirement. In this contribution, a new unified registry EMIR is proposed, which attempts to overcome the challenges of federation, robustness and performance implications of today's DCIs.

**Ahmed Shiraz Memon**  
*a.memon@fz-juelich.de*

## dCache, agile adoption of storage technology

For over a decade, dCache has been synonymous with large-capacity, fault-tolerant storage using commodity hardware that supports seamless data migration to and from tape. It has satisfied the requirements of various demanding scientific user communities to store their data, transfer it between sites and fast, site-local access.

Technology changes over time. When the dCache project started, the focus was on managing a relatively small disk cache in front of large tape archives. Over the project's lifetime, technology changes have driven down the cost-per-GB of hard disks.

This triggered a shift towards systems where the majority of data is stored on disk. More recently, the availability of Solid State Disks, while not yet a replacement for magnetic disks, offers an intriguing opportunity for significant performance improvement if they can be used intelligently within an existing system.

**Paul Millar**  
*paul.millar@desy.de*



## EMI Data, preparing for the finish line

The European Middleware Initiative is completing its second year. While during the first year, the focus has been on the successful integration of the various components of the four middlewares into a single distribution, the second year was dominated by finishing important developments in order to make the EMI bundle compliant to standards and more convenient to install, configure and to operate.

**Patrick Fuhrmann**  
*patrick.fuhrmann@desy.de*

## Software for distributed systems – the EMI product portfolio

The European Middleware Initiative (EMI) brings together ARC, dCache, gLite, and UNICORE to provide a harmonised set of products and streamlined releases to the DCI community. While there are many technical solutions around, EMI is one of the key players in providing software for large-scale distributed systems that are operated around the world today. Having products and solutions in various technical

areas such as compute, data, information, and security, it is interesting to understand that these products also implement many of the principles and paradigms of distributed systems. This contribution will provide an overview of the whole EMI product portfolio focusing on its key features and their role in distributed systems based on comparisons with known literature such as books offered by Tanenbaum.

**Morris Riedel**  
*m.riedel@fz-juelich.de*

## The ARC Information System: overview of a GLUE2 compliant production system

Information systems are the backbone of any kind of distributed grid systems. NorduGrid ARC was one of the first middlewares that came up with a reliable and comprehensive information system architecture. This solution, initially inspired by an LDAP-based approach of Globus, has been used as a distributed dynamic database for grid resource discovery and monitoring in research infrastructures for many years. This paper gives an overview of the system architecture, the underlying

technology and features. Although the original ARC information system came with a custom information schema, the standard GLUE2 model is being broadly endorsed by middleware developers now. ARC contributed to the GLUE2 development, and is among the first to implement it, as described in this paper. Ongoing new developments targeting further convergence and harmonisation of otherwise different grid information systems are also discussed.

**Florido Paganelli**  
*florido.paganelli@hep.lu.se*

**Balazs Konya**  
*balazs.konya@hep.lu.se*

**Oxana Smirnova**  
*oxana.smirnova@hep.lu.se*

**Mattias Ellert**  
*mattias.ellert@fysast.uu.se*

## Standardisation strategies of the European Middleware Initiative

For several years there have been many issues surrounding the integration of Open Grid Services Architecture (OGSA) concepts in Distributed Computing Infrastructures (DCIs) and standardisation of grid services within grid middleware systems provided by EMI. OGSA represents a rather massive architecture in the context of distributed systems based on the concept of a plethora of grid services that cover functionality of many technical areas. The issues can be partly explained

by the fact that working interactions among numerous grid services as envisaged by OGSA are non-trivial when we consider their implementations based on web services message exchanges. In contrast, the standardisation strategy of EMI follows a more lightweight and practical approach. This paper will contrast OGSA versus standards adopted by EMI following an approach that is based on the success of TCP/IP in contrast to ISO/OSI.

**Morris Riedel**  
*m.riedel@fz-juelich.de*



## T-StoRM: a StoRM testing framework

StoRM, an implementation of the SRM interface, is a multi-service software subject to intense testing, validation and verification activities in order to guarantee high-quality services. Its characteristics of being usable on different file systems and of supporting several transfer protocols raise the need for StoRM to be validated on a variety of deployment scenarios with multiple machines. With this in mind, T-StoRM is a StoRM testing framework that aims at improving and automating service evaluation. It provides several abstract support classes that can simplify writing test suites, which are a logical group of similar test

cases amongst installation, configuration, conformance, system, regression and stress categories. Such a solution addresses the need for improvement the software development life cycle and optimisation of the deployment of a new software release. In this paper, we describe the design and development of T-StoRM, and we present its usage during the StoRM development life cycle.

**Elisabetta Ronchieri**  
*elisabetta.ronchieri@  
cnaf.infn.it*

## Grid Engine batch system integration in the EMI era

The Grid Engine (GE) family of batch systems, successors of SGE, is a modern open source batch system which is now supported in EMI under a new CREAMCE version, and fully compliant with EMI policies. This integration is aimed at increasing the robustness of the delivered functionalities with a major refurbish of the BUPdaterSGE, the daemon responsible for tracking submitted tasks. BUPdater was changed

to get jobs statuses more efficiently and is now able to collect the information status for a large number of jobs in a few seconds. The new implementation also includes default MPI support while GLUE2 schema compliance is still in development, which will be included in future versions. To improve the interaction with external services, DRMAA has been positively evaluated and will be included in future development.

**Roberto Rosende Dopazo**  
*rosende@cesga.es*  
**Álvaro Simón García**  
*asimon@cesga.es*  
**Gonçalo Borges**  
*goncalo@lip.pt*  
**Álvaro López García**  
*alvaro.lopez.garcia@  
cern.ch*

## CREAM Computing Element: status and new developments

The CREAM (Computing Resource Execution And Management) Service, a service for job management operation at the Computing Element (CE) level, is one of the software products part of the EMI middleware distribution. It implements a grid job management service which allows the submission, management and monitoring of computational jobs to local resource management systems.

We present some new functionalities introduced in this service with the first EMI major release, and we discuss about some other new features being implemented.

**Massimo Sgaravatto**  
*massimo.sgaravatto@  
pd.infn.it*

**Alvise Dorigo**  
*alvise.dorigo@pd.infn.it*

**David Rebatto**  
*david.rebatto@mi.infn.it*

**Massimo Mezzadri**  
*massimo.mezzadri@  
mi.infn.it*

**Francesco Prelz**  
*francesco.prelz@  
mi.infn.it*

**Lisa Zangrando**  
*lisa.zangrando@  
pd.infn.it*

**Eric Frizziero**  
*eric.frizziero@pd.infn.it*

**Paolo Andreetto**  
*paolo.andreetto@  
pd.infn.it*

**Alessio Gianelle**  
*alessio.gianelle@  
pd.infn.it*

**Sara Bertocco**  
*sara.bertocco@pd.infn.it*

**Marco Cecchi**  
*marco.cecchi@cnaf.  
infn.it*

**Fabio Capannini**  
*fabio.capannini@cnaf.  
infn.it*

**Salvatore Monforte**  
*salvatore.monforte@  
ct.infn.it*

## Using L&B to monitor Torque jobs across a national grid

gLite's Logging and Bookkeeping service (L&B) is designed to support other types of jobs other than native gLite. Support for PBS jobs was first introduced at HPDC 2008, and a pilot deployment was performed in MetaCentrum, the Czech national grid that was relied on PBS as its central batch system. Later, MetaCentrum commenced transition to Torque, an open-source alternative to PBS. To suit

MetaCentrum's needs better, Torque was extended to allow for running several independent instances, which could nevertheless forward jobs to each other if one instance could not satisfy current resource requirements. This lead to a new task: providing a tool to monitor grid jobs potentially migrating over several instances of Torque. L&B was invited to provide that functionality.

**Michal Vocu**  
*michal@cuni.cz*

**Zdenek Sustr**  
*sustr4@cesnet.cz*

**Daniel Kouril**  
*kouril@ics.muni.cz*

**Frantisek Dvorak**  
*valtri@civ.zcu.cz*

**Miroslav Ruda**  
*ruda@ics.muni.cz*

**Simon Toth**  
*toth@fi.muni.cz*

**Jiri Sitera**  
*sitera@civ.zcu.cz*

**Jiri Filipovic**  
*fila@ics.muni.cz*

**Marcel Poul**  
*poul@mail.muni.cz*

**Ludek Matyska**  
*ludek@ics.muni.cz*



# Operational services and infrastructure

This track includes presentations on the advancements and needs of systems supporting infrastructure and user operations: data management, accounting, the messaging service, the EGI helpdesk, approaches to site monitoring, IPv6 and tools for network monitoring. This will be of special interest to site managers and users with expertise in data management and operational systems for monitoring, accounting and support.

The accounting session will provide information on the current status and plans of the EGI accounting infrastructure, and attendees will be invited to discuss requirements, and to learn about how information on usage of new resources can be embedded into application frameworks. Communities that are interested in an advanced usage of GGUS are welcome to attend the EGI helpdesk session, which delves into the features of the GGUS system that are used by advanced communities to track progress of incidents and to notify site administrators in case of critical problems.

*Tiziana Ferrari* Chief Operations Officer, EGI.eu





## User-centric monitoring of the analysis and production activities within ATLAS and CMS VOs using the Experiment Dashboard system

The Experiment Dashboard is a monitoring system developed for the LHC experiments in order to provide the view of the grid infrastructure from the perspective of the Virtual Organisation (VO). It enables a transparent view of the experiment activities across different middleware implementations and combines the grid monitoring data with information that is specific to the VO. Job processing is the core part of the VO computing activities. The scientists must be able to

monitor the execution status, application and grid-level messages of their tasks that may run at any site within the VO. The Dashboard Task Monitoring applications collect and expose a user-centric set of information to the user regarding submitted tasks. They provide a clear and precise view of the task status evolution and reason for failure as a function of time or site. Advanced graphical plots are also available which give a more usable and attractive interface to the analysis and production user.

**Julia Andreeva**  
*julia.andreeva@cern.ch*

**Mattia Cinquilli**  
*mattia.cinquilli@cern.ch*

**Ivan Dzhunov**  
*ivan.dzhunov@cern.ch*

**Edward Karavakis**  
*edward.karavakis@cern.ch*

**Michael Kenyon**  
*mkenyon@cern.ch*

**Lukasz Kokoszkiewicz**  
*lukasz.kokoszkiewicz@cern.ch*

**Pablo Saiz**  
*pablo.saiz@cern.ch*

**Laura Sargsyan**  
*lsargsya@mail.cern.ch*

**David Tuckett**  
*david.tuckett@cern.ch*

## The role of GGUS in delivering highly reliable WLCG operations

The Worldwide LHC Computing Grid (WLCG) strategy is to use the Global Grid User Support (GGUS) system for tracking grid-related incidents. Tickets are created every day by all four LHC experiments and their progress is discussed at 3pm CET every working day. For the WLCG Tier-1s—the grid sites providing the biggest fraction of computing capacity to the Virtual Organisations (VOs) of the LHC and connected with high bandwidth with CERN (mentioned here as the WLCG Tier-0)—if an incident is not recorded in a GGUS ticket, then it is not considered for further investigation. This talk

will describe the ways we used to extend and promote the rich GGUS functionality in order to get the very requiring WLCG user community on board and in order to deal with incidents that need immediate attention and rapid resolution.

**Maria Dimou**  
*maria.dimou@cern.ch*

## The GGUS report generator

This presentation will give an overview of the latest developments of the GGUS report generator. To deal with a number of new requirements a project plan has been set up for the implementation of the new version of the GGUS report generator. Major features are a flexible combination of different report

parameters for customised reports, pre-defined standard reports executed according to a time schedule, the calculation of key performance indicators, the calculation of statistic values, SLA monitoring and flexible data aggregation. The results of the first implementation phase will be presented here.

**Torsten Antoni**  
*torsten.antoni@kit.edu*

**Helmut Dres**  
*helmut.dres@kit.edu*

**Guenther Grein**  
*guenther.grein@kit.edu*

**Carsten Rogge**  
*carsten.rogge@kit.edu*

**Oleg Dulov**  
*oleg.dulov@kit.edu*

## Site Status Board: a flexible monitoring system developed in close collaboration with user communities.

Development of highly customisable and flexible solutions requires close collaborative work between developers and the user community. Gathering user requirements and understanding user needs helps developers to provide common, but highly customisable solutions that will fit the needs of different groups of users. One example of successful collaborative development is the Dashboard Site Status Board (SSB) framework, which allows Virtual Organisations (VO) to monitor their computing activities at distributed sites and to evaluate site performance from the VO perspective.

**Ivan Dzhunov**  
*ivan.antoni.dzhunov@cern.ch*

**Julia Andreeva**  
*julia.andreeva@cern.ch*

**Josep Flix**  
*jflix@cern.ch*

**Peter Kreuzer**  
*pkreuzer@cern.ch*

**Pablo Saiz**  
*pablo.saiz@cern.ch*

**Edward Karavakis**  
*edward.karavakis@cern.ch*

**Lukasz Kokoszkiewicz**  
*lkokosz@cern.ch*

**David Tuckett**  
*david.tuckett@cern.ch*

**Michal Nowotka**  
*micah.nowotka@cern.ch*

**Alessandro Di Girolamo**  
*ale@diggi.org*

**Andrea Sciaba**  
*andrea.sciaba@cern.ch*

**Jaroslava Schovancova**  
*jschovan@cern.ch*

**Jaroslava Schovancova**  
*jschovan@cern.ch*



## Hybrid cloud-based grid infrastructure: experience and future

This work will expose our experience and plans regarding our OpenNebula-based hybrid cloud infrastructure, which offers many grid services like computing infrastructure, regional coordination and monitoring, accounting and metrics portals and software provisioning, allowing easy partitioning and allocation of resources.

**Alejandro Feijoo**  
*alfeijoo@cesga.es*

**Carlos Fernandez**  
*carlosf@cesga.es*

**Ivan Diaz Alvarez**  
*idiaz@cesga.es*

**Álvaro Simón**  
*asimon@cesga.es*

**Esteban Freire**  
*esfreire@cesga.es*

**Rubén Díez**  
*rdiez@cesga.es*

## Improving resilience of T0 grid services

This presentation will cover the operational aspects of running grid services at the CERN Tier-0 grid site, with particular focus on Configuration Management, Service Monitoring and ways to increase service resilience. Currently, some of our grid services are provided by

virtual machines instantiated in the service consolidation service. This, together with an intensive use of DNS load balancing, highly increases their availability. The presentation also describes how grid services would need to evolve in order to be able to run in a computing cloud.

**Manuel Guijarro**  
*manuel.guijarro@cern.ch*

## User community input to new types of accounting

Come and find out about new types of accounting being proposed/developed in the EGI ecosystem. Do they meet your needs? Do you have other requirements for these or other types of accounting?

**John Gordon**  
*john.gordon@stfc.ac.uk*

## DDM Site Services: a solution for global replication of HEP data

The ATLAS Distributed Data Management (DDM) is the project built on top of middleware currently maintained by EMI and responsible for the organisation of the multi-Petabyte ATLAS data across more than 100 distributed grid sites. One particular component of the system—the DDM Site Services—is the set of agents responsible for the discovery and placement of ATLAS data between sites. DDM Site Services manage aggregated throughputs of over 6GB/s or one million file transfers a day and have to work with extremely high reliability and availability. This contribution

will build upon the production experience acquired during the last two years of LHC data taking and show the changes, adaptations and improvements that we implemented on the system to guarantee a flawless service. In a second part, we will give an update on the service and activity monitoring frameworks that publish the information needed by shifters and experts.

**Fernando Harald Barreiro Megino**  
*fernando.harald.barreiro.megino@cern.ch*

**David Tuckett**  
*david.tuckett@cern.ch*

**Alessandro di Girolamo**  
*alessandro.di.girolamo@cern.ch*





## Grid/cloud computing platform for medical imaging and signal processing in clinical decision system

Recent IT solutions and standards provide new opportunities for clinical decision systems (CDS), improving patient healthcare. Conceptually, CDS include algorithms for analysis and semantic interoperability of combined patient health information, as well as image and signal analysis algorithms and methodology. Such an approach gives doctors and patients a powerful tool to enhance diagnostics, to be more precise in drug prescription, to save a patient's time, and to enable special repositories for knowledge acquisition. Algorithms for CDS, developed in electrophysiology,

orthopaedics and radiology are based on a data platform, serving medical institutions. The other part of CDS is based on knowledge management services to support translational research content. As a consequence, an efficient cloud/grid computing platform is being designed, giving powerful analytical procedures to doctors and patients for medical data analysis and functionality.

**Algimantas Juozapavicius**  
algimantas.juozapavicius@mif.vu.lt

## Consistency between grid storage elements and file catalogue for the LHCb experiment

DIRAC is a framework developed to provide a complete solution for using distributed computing resources. DIRAC has been developed in a very generic way that has made it suitable for serving many VOs. The LHCbDIRAC framework is the DIRAC extension specific to LHCb, one of the four experiments operating at the Large Hadron Collider of CERN, and is where the particular features requested by the LHCb community are implemented. LHCbDIRAC is organised into several systems providing all the functionality needed for using a distributed computing infrastructure, including a data

management system, which implements an interface to the underlying grid middleware and performs all operations of data transfer, registration and removal. In this paper we shall describe a new component of the LHCbDIRAC data management system which checks the consistency between the data stored on the grid storage elements and the central file catalogue of the experiment.

**Elisa Lanciotti**  
elisa.lanciotti@cern.ch

## Service Availability Monitoring (SAM)

Overall status of the recent developments in Service Availability Monitoring (SAM).

**Wojciech Lapka**  
wojciech.lapka@cern.ch

**Emir Imamagic**  
emir.imamagic@cern.ch

**Christos Triantafyllidis**  
ctria@grid.auth.gr

**Pedro Andrade**  
pedro.andrade@cern.ch

**Paloma Fuente**  
paloma.fuente.fernandez@cern.ch

**Marian Babik**  
marian.babik@cern.ch

**David Collados**  
david.collados@cern.ch

## Accounting of storage resources usage for the LHCb VO

The DIRAC framework was developed in order to provide a complete solution for using the distributed computing resources of the LHCb experiment. DIRAC has been developed in a very generic way and with a modular architecture, that has made it suitable for serving other VOs as well, e.g. the Belle II experiment at KEK and the ILC project. The LHCbDIRAC system is the DIRAC extension specific to the LHCb experiment, where the particular features requested by the LHCb community are implemented. The LHCbDIRAC framework is split into several systems, inheriting from the corresponding

DIRAC systems, providing all the functionality needed for using a distributed computing infrastructure, including a workload management system, a data management system, and monitoring and accounting systems. In this paper we describe some new functionalities of the LHCbDirac accounting system, which provide a historical view of space usage at grid sites.

**Elisa Lanciotti**  
elisa.lanciotti@cern.ch

**Ricardo Graciani**  
graciani@ecm.ub.es

**Adria Casajus**  
adria@ecm.ub.es



## Gratia: new development in grid accounting

The Grid Accounting Service (Gratia) was developed as a collaboration between Fermilab, US CMS, US ATLAS, and the Open Science Grid (OSG) and is in use across OSG, US LHC, Fermilab and the University of Nebraska. Repositories are operated by several sites, including OSG Operations at IU and Fermigrid. The Gratia development group is in close collaboration with the OSG software, campus, operations and technology teams, who work together on requirements, distribution, and support.

This talk describes several aspects of our most recent work. Among other things, we have been working on a prototype that uses the Open Nebula API to extract information about Virtual Machine states and resource usage in a cloud environment.

Another project that is currently in development strives to accurately record Campus Grids usage of the OSG resources.

In addition, we are working on modification of the Gratia-APEL interface as well as researching a means for sharing some of the metrics with other grid infrastructures.

**Tanya Levshina**  
*tlevshin@fnal.gov*

## StratusLab: use cases, features and sustainability

Cloud computing has emerged in recent years as the new paradigm for the provision of on-demand distributed computing resources. The StratusLab project develops and provides an open-source cloud distribution that allows data centres to expose their computing resources as an 'Infrastructure as a Service' (IaaS) type cloud. Administrators can run their services over the cloud to improve availability, scalability, and maintainability.

StratusLab supports a number of grid-specific features, and aims to integrate easily into traditional grid sites and the existing EGI ecosystem. We present a number of existing use cases, and the current status of the StratusLab distribution. New and enhanced features of the StratusLab cloud distribution such as persistent storage and the Marketplace will be introduced. We conclude with a discussion of plans to ensure future evolution and sustainability of the StratusLab software.

**Charles Loomis**  
*loomis@lal.in2p3.fr*

## Monitoring in CORAL

The Large Hadron Collider (LHC), the world's largest and highest energy particle accelerator, started its operations in September 2008 at CERN, Switzerland. Huge amounts of data are generated by the four experiments installed at different collision points along the LHC ring. The largest data volumes come from the 'event data' that record the signals left in the detectors by the particles generated in the LHC beam

collisions and are generally stored on files. Relational database systems are commonly used to store the 'conditions data' that record the geometry, configuration and other working parameters of the detectors at the time of the event data. The CORAL software is widely used by the LHC experiments for storing and accessing conditions data using relational database technologies.

**Alexander Loth**  
*alexander.loth@cern.ch*

## The Green Computing Observatory

The first barrier to improved energy efficiency of IT systems is the lack of large-scale collections of experimental data. The Green Computing Observatory (GCO) monitors a large computing center (Laboratoire de l'Accélérateur Linéaire-LAL) within the EGI grid, and publishes the data through the Grid Observatory. These data include the detailed monitoring of the processors and motherboards, as well as the global site information, such as overall consumption and overall cooling, as optimising at the global level is a promising way of research. A second barrier is making the collected data usable.

The difficulty is to make the data readily consistent and complete, as well as understandable for further exploitation. For this purpose, GCO opts for an ontological approach in order to rigorously define the semantics of the data (what is measured) and the context of their production (how are they acquired and/or calculated).

**Julien Nauroy**  
*julien.nauroy@lri.fr*  
**Guillaume Philippon**  
*philippo@lal.in2p3.fr*  
**Cecile Germain-Renaud**  
*cecile.germain@lri.fr*  
**Frederic Furst**  
*frederic.furst@u-picardie.fr*  
**Michel Jouvin**  
*jouvin@lal.in2p3.fr*  
**Gilles Kassel**  
*gilles.kassel@u-picardie.fr*



## Messaging service and client software

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

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

It is not trivial to write robust applications which make use of messaging. In order to prevent code duplication and extra maintenance, the EMI Messaging Product Team provides a set of tools and libraries that 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.

**Massimo Paladin**  
*massimo.paladin@cern.ch*

## Optimising the usage of multi-Petabyte storage resources for LHC experiments

In the last two years of LHC operation, the experiments have made a considerable usage of grid resources for data storage and offline analysis. To achieve successful exploitation of these resources, a significant operational human effort has been put in place and it is now the moment to improve the usage of the available infrastructure.

In this respect, the CMS Popularity project aims to track the experiment's data access patterns (frequency of data access, access protocols, users, sites and CPU), providing the base for the automation of data

cleaning and data placement activity on grid sites. As well, the popularity-based Site Cleaning Agent has been developed to monitor the evolution in time of the used and pledged space and remove unused data replicas at full Tier-2s.

This presentation will give an insight into the development, validation and production process of these systems. We will analyse how the framework has influenced resource optimisation and daily operations in CMS.

**Daniele Spiga**  
*daniele.spiga@cern.ch*

**Maria Girone**  
*maria.girone@cern.ch*

**Nicolo Magini**  
*nicolo.magini@cern.ch*

**Valentina Mancinelli**  
*valentina.mancinelli@cern.ch*

**Mattia Cinquilli**  
*mattia.cinquilli@cern.ch*

**Edward Karavakis**  
*edward.karavakis@cern.ch*

**Fernando Harald Barreiro Megino**  
*fernando.harald\_barreiro.megino@cern.ch*

**Domenico Giordano**  
*domenico.giordano@cern.ch*

## Experience in grid site testing for High Energy Physics with HammerCloud

HammerCloud is a grid site testing service for the ATLAS, CMS and LHCb experiments centred at CERN in Geneva. This tool, which is provided as an online service for operation managers, site administrators and, in general, grid experts, allows them to perform on-demand tests of their computing facilities in order to validate and measure their performance. In addition, HammerCloud runs automated tests to check the availability and reliability of the sites under different circumstances. The tests consist of real analysis code provided by the physics community to ensure real-world use cases for the grid

sites. Indeed, HammerCloud has been employed in HEP for more than two years and has helped to increase the performance and reliability seen by the grid users. In this work we will present the lessons learnt while deploying, optimising and evolving the system for the three VOs and the development plans for the near and mid-term future.

**Daniel Colin van der Ster**  
*daniel.colin.vanderster@cern.ch*

## Experiences running a HPC Cloud

We have developed experience in running an HPC cloud at SARA over the last two years. Starting out with a proof of concept environment, we now have a production HPC cloud infrastructure to offer to our users. In this environment, users can start-up their own HPC compute clusters, having their own private high bandwidth 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 and discuss our experiences that we have gained so far.

**Ron Trompert**  
*ron@sara.nl*



## Providing resources through cloud and grid interfaces by means of the standard CREAM CE and the WNoDeS cloud solution

We will show the work done in order to enable a standard EGI grid farm with a cloud infrastructure and user interface, by means of the standard CREAM CE and the WNoDeS cloud solution.

The final goal of this work is to provide to the end-user a rich set of interfaces for access to computing resources, such as: standard grid job submission, grid job submission to specific virtual machine images, a web interface for allocating virtual machines or interactive usage of specific virtual images.

The whole infrastructure is based on open source software such as: CREAM CE, WNoDeS, Torque/Maui and Lustre.

It is intended to fulfil the computing requirements of different communities of researchers, starting from LHC and other communities that have great experience in grid computing, but also providing simple and user-friendly interfaces to users who do not have the knowledge to exploit complex grid infrastructures.

**Paolo Veronesi**  
*paolo.veronesi@cnafe.  
infn.it*

**Giacinto Donvito**  
*giacinto.donvito@  
ba.infn.it*

**Davide Salomoni**  
*davide.salomoni@cnafe.  
infn.it*

**Gianni Dalla Torre**  
*gianni.dallatorre@cnafe.  
infn.it*

**Alessandro Italiano**  
*alessandro.italiano@  
cnafe.infn.it*

**Alessandro Paolini**  
*alessandro.paolini@  
cnafe.infn.it*

**Elisabetta Ronchieri**  
*elisabetta.ronchieri@  
cnafe.infn.it*



# Coordination and communication

The EGI ecosystem is sustainable only when it is capable of ensuring its value creation in the interactions between the participating organisations. We need a common understanding and consensus on the types of benefits being generated, to identify the vision that we want to achieve together and to define a strategic plan to move forward in the right direction. Furthermore, we need to identify the optimal communication strategies to put in place in order to reach the right audiences with the most appropriate language and messages.

This track addresses these vital aspects for the long-term evolution of EGI. The EGI strategic plan will be discussed at the 'Sustaining the EGI ecosystem' workshop. The 'Commission Original Software' session will focus on how research teams can commission original software and what are the logistics of this process.

*Sergio Andreozzi Strategy and Policy Manager, EGI.eu*



## A business model approach for a sustainable grid infrastructure in Germany

This talk presents a business plan as suggested for the national German grid initiative (NGI-DE). It is based on quantitative calculations, making it possible to forecast profits and losses, according to a set of mandatory services and 'products'. The presentation also wants to solicit input from the relevant grid user communities and the other NGIs with the goal of creating a common basis for and common understanding of sustainability strategies.

**Torsten Antoni**  
*torsten.antoni@kit.edu*

**Ruediger Berlich**  
*ruediger.berlich@kit.edu*

**Achim Streit**  
*achim.streit@kit.edu*

## How do research teams commission original software applications and ensure that they do what they want them to do?

The EGI ecosystem is built upon a complex construction of evolving software ranging from networking to middleware to user applications. As we move from the large project support model to smaller projects and then to domain-based support we need to re-evaluate the software development processes and their associated project management. The purpose of this workshop is to investigate

how research teams commission original software applications: what approaches are taken, how long does it take, how do you explain what you need? Most importantly, how do you ensure that it does what you want it to do? The workshop is targeted at research teams and their leaders to explain what they need, together with application developers to explain what is possible and the NGIs to explain what is available.

**Stephen Brewer**  
*steve.brewer@egi.eu*

**Richard McLennan**  
*richard.mclennan@egi.eu*

## The EGI Software Vulnerability Group and EMI

This talk provides a brief overview of the activities of the EGI Software Vulnerability Group (SVG), and progress made in collaboration with EMI in addressing software vulnerabilities.

This is followed by a short presentation by 2-3 members of SVG/EMI on how some specific problems occurred and how they were technically addressed.

**Linda Cornwall**  
*linda.cornwall@stfc.ac.uk*

## Communication: the key to future funding

Communicating your work effectively has never been more important for securing future funding. With Horizon 2020 beginning to come on-stream from 2014, funders are increasingly focused on getting the research they support noticed. The watchword is not to disseminate, but to communicate, and establish two-way engagement

with your peers and wider audiences. This interactive workshop gives the inside story on communicating your work through the press, social media and the web. The communications team from EGI will be joined by journalists and professional bloggers to give practical advice and a forum to share what works—and what doesn't.

**Catherine Gater**  
*catherine.gater@egi.eu*



## Setting up a Finnish NGI site—site perspective

In 2011, the Finnish Academy was granted a sizeable amount of funding to set up a national grid infrastructure based on grid computing. Nine Finnish Universities and CSC—IT Center for Science—form this national grid, and our university is one of the participants.

**Vera Hansper**  
*vera.hansper@csc.fi*

## Strategic plan for a scientific cloud computing infrastructure for Europe

Besides the core processes of R&D agencies and organisations for capturing, processing, analysing and archiving data, a new paradigm of having permanent on-line access to IT resources, information and collaboration tools has become a central aspect of scientific endeavour. Research infrastructures such as the members of the EIRO Forum are beginning to challenge the sustainability of an approach to ICT deployment that has predominated for a quarter of a century but is rapidly being overtaken by events.

**Bob Jones**  
*robert.jones@cern.ch*  
**Maryline Lengert**  
*maryline.lengert@esa.int*  
**Rupert Lueck**  
*rupert.lueck@embl.de*

## Linking authenticating and authorising infrastructures in the UK NGI (SARoNGS)

We aim to provide simple trusted access to digital services for the UK's research community including grid and cloud provision. To achieve this we have to satisfy conditions laid down by three types of entity: 1) Individuals; 2) Resources and 3) Identification and Attribute Authorities. Each of these is governed by policies and legal requirements placed upon them as well as human rights legislation, making it difficult if not impossible to fit one access mechanism to all stakeholders.

SARoNGS was a JISC funded technical project that was developed in the UK to apply a federated access model (The Shibboleth based UK access management federation) to the grid environment. It resulted in a production service supported by the UK NGI to issue grid credentials, obtain VO assertions and place them within reach of the user so as to provide access to these online services.

We will present the details of this service, the ways we joined the loose ends together, the remaining issues and future directions.

**Mike Jones**  
*mike.jones@manchester.ac.uk*

## Tweaking the certificate lifecycle for the UK e-Science CA

The UK e-Science Certification Authority (CA) is the second largest IGTF-accredited Grid CA. This paper considers changes to its normal certificate lifecycle (New, Renew, Expire/Revoke) to provide users with simpler ways to re-apply for the e-Science certificates it issues. If a user's email address changes, or his certificate has expired or is lost, he needs to re-apply to the UK e-Science Certification Authority for a new certificate rather than renew it (if the old certificate

hasn't expired it will also need revoking). This requires another face to face visit with the local Registration Authority (RA) Operator for photoID checking and might not be convenient if the user is at a different location.

It is this process of re-applying that we want to simplify, reducing involvement of the RA Operator and attempting to eliminate the need for face to face meetings beyond the initial one.

**John Kewley**  
*john.kewley@stfc.ac.uk*





## PLGrid PLUS: towards domain-specific infrastructure for supporting international research collaboration

The e-infrastructure requirements of the scientific community working on 'Big Science' problems are highly diversified. To harmoniously support, in terms of IT, the development of scientific research in the various problem areas, it is necessary to fit the characteristics of the IT infrastructure to problems being the subject of research. Preparation of specific computing environments, so called domain grids, i.e., solutions, services and extended infrastructure (including software), tailored to the needs of different groups of scientists, is the most important task implemented within the three-year project

'PLGrid PLUS: Polish Roadmap toward Domain-Specific Infrastructure for Supporting Computational Science in European Research Area'. The project, funded by the European Regional Development Fund under the Innovative Economy Program, started in October 2011 and is being carried out by the PL-Grid (Polish NGI) Consortium, coordinated by the ACC Cyfronet AGH.

**Jacek Kitowski**  
*yemosurs@cyf-kr.edu.pl*

## European Globus Community Forum (EGCF)

This talk will provide a brief overview of the European Globus Community Forum (EGCF). The EGCF is the organisational body of the Globus community in Europe.

Its goal is to boost an integrated approach to collaboration on Globus development and to provide an organisational platform to foster cooperation within Europe and beyond.

Its members are users, administrators, and developers, who are applying the Globus Toolkit as their middleware or are interested in doing so.

The presentation will focus on aims and services associated with EGCF. The EGCF will be one of the sustainabilities of the Initiative for Globus in Europe (IGE) project.

**Michael Krieger**  
*michael.krieger@risc-software.at*

## High Performance grid computing: getting HPC and HTC all together

GriF (Grid Framework) is a Workflow Management System (WfMS) designed to evaluate Service and User quality when helping its users (especially those of the COMPCHEM Virtual Organisation) in optimising the return of their parameter sweeping studies submitted to the Grid. GriF has now been extended to allow access to both HPC and HTC resources. The implementation of such a High Performance Grid (called HIPEG) has required a tight collaboration between experts of IGI (the Italian Grid Initiative) for the gLite middleware, of CINECA (the largest Supercomputing Center of Italy) for the

UNICORE software stack and of COMPCHEM the virtual community for both management and applications.

The extended GriF has been used to allow molecular scientists, unwilling (or unable) to deal with the complexity of the grid middleware, to carry out an extended high level ab initio calculation of the electronic structure of molecular systems when working within the grid environment.

**Antonio Lagana**  
*lagana05@gmail.com*

## GC3Pie: A Python framework for high-throughput computing

GC3Pie is a suite of Python classes (and command-line tools built upon them) to aid in submitting and controlling batch jobs to clusters and grid resources seamlessly. GC3Pie aims at providing the building blocks by which Python scripts that combine several applications in a dynamic workflow can be quickly developed. GC3Libs, the main component of the GC3Pie framework, provides services for submitting computational jobs to grids and batch

systems and controlling their execution, persisting job information, and retrieving the final output. GC3Libs takes an application-oriented approach to batch computing. A generic application class provides the basic operations for controlling remote computations, but different application subclasses can expose adapted interfaces, focusing on the most relevant aspects of the application being represented.

**Sergio Maffioletti**  
*sergio.maffioletti@gc3.uzh.ch*





## EGI sustainability

A long-term sustainability strategy needs to be developed and continuously refined to derive plans that will solve both the current and future issues facing e-Infrastructures. This second workshop dedicated to sustainability serves as a follow-up to the EGI Technical Forum. In Lyon, a two-session workshop covered

sustainability plans from the various entities of the EGI ecosystem, comprising resource providers, technology providers and user communities and explored business models in the context of EGI. In this workshop, we will present the progress made to date and the content of the EGI Strategy Plan.

**Steven Newhouse**  
*steven.newhouse@egi.eu*

**Sergio Andreozzi**  
*sergio.andreozzi@egi.eu*

## ScienceSoft – open software for open science

The EMI project brings together middleware services from ARC, dCache, gLite and UNICORE in order to provide a harmonised set of interfaces and a streamlined release for the DCI community. Nevertheless, EMI is a project funded by the EC and ends in 2013 leaving it unclear how harmonised developments will be continued and maintained. The main objective behind ScienceSoft, which will be introduced in this contribution, is to create the conditions for the continuing development,

support and use of the EMI software products after the end of the EMI project. EMI will establish ScienceSoft as an incubator for a broad open source community of developers and users of not only the EMI products, but also other software from projects of critical interest for the scientific research communities. This contribution aims to provide an overview of these activities including the value of ScienceSoft to the community, its goals, and its mandate.

**Morris Riedel**  
*m.riedel@fz-juelich.de*

## Supporting EGI strategic dissemination activities: the European e-infrastructures Observatory

One of the key challenges faced by e-infrastructure providers such as EGI.eu, which represents the European Grid Infrastructure of 350 resource centres organised through over 35 national resource infrastructure providers, is to report the diversity and scope of their resources and their usage to funders and decision makers at national and European levels.

The European e-infrastructures Observatory can provide to the e-infrastructures community a comprehensive monitoring and dissemination framework that will allow EGI's

stakeholders to keep track of the development and prospects in e-infrastructures, including a frame of reference for assessing progress over time towards global regions of interest and disseminating success stories of European-led initiatives.

Such a facility can be of great interest and value to EGI and its stakeholders, as a yardstick tool for progress monitoring, analysis and evaluation of Grid and other e-infrastructures across the European Union and beyond.

**Jorge-A. Sanchez**  
*j.sanchez@jnp.gr*

**Nikos Vogiatzis**  
*n.vogiatzis@jnp.gr*

## Championing e-infrastructure

The UK National Grid Service (NGS) has two Champion networks – Campus and Community who each reach out to a different area of the existing and potential user community in the UK. Campus champions promote e-infrastructure within their own institution to all researchers. Community champions are representatives

from distinct research communities who promote e-infrastructure to their peers and fellow researchers.

This presentation will describe the structure of these two groups, their methods of dissemination, the support provided to them by the NGS and the impact they have had in promoting e-infrastructure.

**Gillian Sinclair**  
*gillian.sinclair@manchester.ac.uk*



## How many users does EGI have?

To build up a comprehensive understanding about the success, uptake and impact of EGI, one needs to look at several metrics. The 'number of EGI users' is one of the most relevant metrics in this respect. While the importance of this metric is apparent, obtaining the number of EGI users is quite a challenge due to the federated, distributed

nature of the infrastructure and its user communities. The talk presents the developments that were recently undertaken by the EGI-InSPIRE project within the Operations Portal, on its underlying services and on reporting mechanisms to clarify the number of EGI users. The talk will also report on recent trends in the number of users.

**Gergely Sipos**  
*gergely.sipos@egi.eu*

## In case the 'happy accident' doesn't happen: training the next generation of senior managers within European e-Infrastructures

How do we train the next generation of senior managers within European e-infrastructures? Does the DCI community do enough to support the development of the next generation of senior managers, or do we rely on 'happy accidents' (coincidence and serendipity) to develop both our own and others' careers? What can the DCI community learn from Europe's established 'traditional' research infrastructures, and vice versa?

This presentation will discuss these questions in the context of a specific European project, RAMIRI (Realising and Managing International Research Infrastructures), that has tried to develop a basic, structured training programme for mid-career research infrastructure managers from both 'traditional' (single-site) research facilities and distributed/e-infrastructures. Whilst we remain a long way off developing an 'MBA for e-RI managers', the successes and challenges of the RAMIRI project over four years provide a number of insights, recommendations and future possibilities.

**Naomi Wynter-Vincent**  
*n.wynter-vincent@imperial.ac.uk*

