Inspired

news from the EGI community

Antoine Taveneaux / wikicommons)

TOP STORIES

Adding pay-per-use models within EGI

page 2

GPGPUs: accelerated computing on EGI?

page 3

What is new in GLUE 2?

page 4

VisIVO: Large-Scale Visualisation on DCIs

page 7

MORE

EGi

European Grid Infrastructure

www.egi.eu

This Issue

Welcome to Issue 10 of *Inspired*!

We changed a few things to celebrate our tenth newsletter. The front page was redesigned, our layout was updated. And those of you who receive the newsletter by email may have noticed a (we hope) more user friendly interface.

As for the content, this issue brings you:

> Sy Holsinger announces a 'proof of concept' experiment on **pay-per-use models within EGI**

> Neasan O'Neill profiles the **UK NGI**

> Steve Brewer presents our first **EGI Champions**

> Ugo Becciani introduces **VisIVO** - a visualisation tool for e-Science

> John Walsh writes about on the **potential of GPUs in EGI**

> Michael Schiffers explains how **MAPPER** allows scientists to look at things from all angles

> Stephen Burke tells us **what's new in GLUE 2**, the language used to describe different grid resources

Your thoughts and comments are always welcome!

Sara Coelho
sara.coelho@egi.eu

This Issue's photo is a close up of Alan Turing statue at Bletchley Park.
His life and legacy was celebrated during the 2012 Turing year.
In 2013, EGI will travel to Manchester, where Turing developed some of his work, for the Community Forum in the Spring.
(Illustration: Antoine Taveneaux / wikicommons)

Registration for the Community Forum is now open

The European Grid Infrastructure invites everyone with an interest in distributed computing technologies and their applications to science to register for the EGI Community Forum.

The Community Forum will be held in Manchester, UK (8-12 April 2013), at the very heart of the University of Manchester's main campus, Oxford Road, only minutes from the centre of Manchester City.

Online registration for the conference is now open and early-bird discounted rates will be available until 22 February 2013.

The meeting will be hosted by EGI.eu and UK NGI, a partnership between GridPP and the National e-Infrastructure Service (NES).

The main theme of this year's Community Forum is International Collaboration. The event will bring together researchers, engineers and IT people from around the world to discuss the use of distributed computing and storage infrastructures.

The Community Forum will be held in conjunction with the 3rd EMI Technical Conference and co-located with the 3rd European Globus Community Forum (8 April).

Confirmed keynote speakers include:

> **Eleni Katragkou**, climate scientist based at the Aristotle University in Thessaloniki, Greece and EGI Champion

> **Giovanni Frisoni**, neurolologist and deputy director of the National Centre for Alzheimer's and Mental Diseases in Brescia, Italy.

More information

EGI Community Forum website - *http://cf2013.egi.eu*

Online registration - *http://go.egi.eu/reg-cf13*

Adding pay-per-use models within EGI

Sy Holsinger announces a 'proof of concept' experiment

The EGI Council has recently endorsed a thought experiment on how pay-per-use models could be implemented and users billed for their usage of EGI resources.

In a distributed environment such as EGI, services are provided by a variety of different organisations spread across Europe. EGI.eu – as federator of resources – can, in this context, provide different degrees of shared services according to a range of models. The possibilities are:

> **'Independent Advisor'**, facilitating users in finding the best resource provider that can fulfil their needs.

> **'Matchmaker'**, where the federator also takes care of performing the resource allocation as a central service.

> **'One Stop Shop'**, where the federator supports users to find the best resource providers, allocates resources and handles financial transactions as well.

To explore these options, EGI.eu will run a 'proof of concept'-style experiment with NGIs, resource centres and user communities. This will help us to define areas such as accounting, pricing, invoicing, payments and service level management.

The pay-per-use experiment is a response to an evolving landscape that offers access to the flexibility and agility of on-demand ICT infrastructures through the use of cloud computing. These benefits are attracting different parts of the research community to EGI.

This approach also allows researchers another way of accessing new or additional resources, or to make the most of funding opportunities that allocate money directly to the purchase of resources and/or services for their work. Therefore, EGI is keen to explore pay-for-use options to increase the sustainability of a more market-driven e-Infrastructure to support the digital European Research Area.

Any pay-per-use model will complement the current practice of researchers accessing resources free at point of delivery ICT services, which were paid upfront by funding agencies. Resource providers would be given the flexibility to 'opt-in' for those who are able and willing to participate in the pay-per-use model.

More information

Background reading
Exploring how researchers can pay for EGI Resources - *http://go.egi.eu/1391*

Pay-per-use survey
NGIs and Resource Centres can participate by completing the following survey - *http://www.surveymonkey.com /s/HMZ2686*

User communities interested can also participate by contacting policy@egi.eu.

Evolving EGI Workshop - 28-30 January 2013, Amsterdam

The workshop will give members of the EGI community an excellent opportunity to discuss and actively contribute to shaping the future of EGI. Apart from the pay-per-use experiment, the workshop will also focus on:

> the evolution of the EGI global services and their funding streams after the end of EGI-InSPIRE. We will discuss which services – both old and new – are critical for the success of EGI
> developing a process that allows user communities to undertake excellent science on EGI resources through peer reviewed access to pool resources coming from the cooperating NGIs and their resource centres

The meeting will be co-located with the final e-FISCAL workshop, summing up the project's findings on cost model, figures, comparison with commercial cloud providers and possible pricing models.

Registration for the event is now open at: *http://go.egi.eu/Evolving-EGI-WS-2013-Registration*. The draft programme and timetable are also available online at *http://go.egi.eu/Evolving-EGI-WS-2013*.

GPGPUs: accelerated computing on EGI?

John Walsh writes about the potential of General Purpose GPUs

The GPGPU Virtual Team (GPGPU-VT) brought together 22 volunteers from across the EGI community, as well as from the MAPPER project, to look at the impact of GPGPUs and investigate if they are an added value to EGI.

The team gathered use-cases to discover how GPGPUs are currently being exploited and asked resource centre admins and user communities about current and future use of GPGPUs.

Do GPGPUs allow users to run bigger and/or faster applications than would otherwise be possible?

Indeed, the latest 'Top 500' Supercomputers shows that 62 systems use accelerators or GPGPU co-processors – an increase from the previous year.

The potential has not gone unnoticed by EGI users and resource centres. Over the past few years, user communities have started to develop GPGPU applications, while resource centres have started to deploy and integrate GPGPUs into their facilities.

Are GPGPUs making an impact with the grid-user community and resource centres?

The results from the survey indicate that they are!

The main findings of the survey have shown that over eighty percent of resource centres that currently offer GPGPU facilities

> ### GPUs & GPGPUs
>
> A **Graphics Processing Unit (GPU)** is a component 'add-on' device for desktop computers and servers that employs specialised parallel processing techniques to render complex graphical images.
>
> **General Purpose Graphics Processing Unit (GPGPUs)** are recent generations of GPUs that can be used to perform more general computing tasks efficiently by taking advantage of high levels of data parallelism.

intend to increase their capacity over the next two years.

Furthermore, fifty percent of the respondents said that they intend to offer GPGPUs resources to their users within the next two years. Most of the GPGPU devices deployed so far are high-end Nvidia-based units.

From the user's perspective, sixty-five percent said that they already used GPGPU applications (but not necessarily via grid) and over ninety percent would be interested in using GPGPU resources if they were available through the grid.

The most popular application development framework is CUDA, and a large number of users are also developing applications using OpenCL. Interest in the newer OpenACC framework is also growing.

So why are GPGPUs not used more widely at the moment on the grid?

Unlike storage and CPU resources, GPGPUs have no formal grid description or support, and as a result, they are not a discoverable resource. There are also difficulties in sharing and scheduling GPGPU access between grid-user jobs. It is difficult, for example, to

determine how many GPGPUs are available; how many are in current use; what access policies apply; how fast they are.

Current access to GPGPUs through the grid requires local knowledge. Extending grid to support GPGPUs is non-trivial - the lack of any formal standardised GLUE description, the lack of information providers to aid publishing the state and access policies of the resources, and the need to create a batch system integration knowledge-base and best-practice documentation all indicate how much technical effort is required to promote GPGPUs as first-class discoverable and exploitable grid resources.

As a follow-up to the GPGPU-VT, we are interested in establishing a technical group of users and resource administrator who can bring their experience and determination to solve these problems.

> ### More information
>
> The EGI discussion forum portal (https://forum.egi.eu/) also has a section dedicated to the topic of GPGPUs, and participation is very welcome.

What is new in GLUE 2?

Stephen Burke tells us what to expect of the new version

> ### What is GLUE ?
>
> **GLUE** – Grid Laboratory for a Uniform Environment – is a schema, or framework, used to describe the properties and state of all grid services in a uniform way. GLUE 2 is its second version.

The GLUE story so far

The original GLUE schema dates from 2002, when it was set up as an interoperability project between US and European grids. GLUE was updated in 2005 and then again in 2006, but we always required updates to be backward-compatible. This made migration to a new version relatively easy, but was also a major constraint on what we could achieve, essentially meaning that we could add new attributes but not change or restructure existing things.

By 2007 it was decided to design a new schema from scratch, building on five years of real-world experience and with no requirement for backward compatibility. This allowed us to create a uniform structure for information about any grid service, to incorporate old features in a more logical way, to create new ones, and to build in more flexibility for future changes.

This work was completed in the context of an OGF working group, making GLUE a genuine grid standard for the first time.

GLUE 2 deployment

The draft GLUE 2 specification was open for comments during the Summer of 2008, and was finally approved as an OGF standard in March 2009. Since then we have progressively implemented and deployed GLUE 2 in the e-infrastructure. The BDII has been GLUE 2-enabled since 2010, all gLite services publish themselves in the GLUE 2 format since the EMI 2 middleware release, and it is expected that ARC, UNICORE and Globus services will be added by the EMI 3 release.

The recent campaign to upgrade all EGI sites to EMI versions means that GLUE 2 information is now widely available and ready to be used.

What is new?

People used to GLUE 1 should find GLUE 2 fairly familiar and all the usual information should be present, together with some additional things which have been requested over the years

The structure is now uniform rather than having completely different formats to describe computing, storage and other services.

It is also much easier to add new information in GLUE 2. For example, for computing systems the schema allows arbitrary benchmarks to be published, rather than the obsolete SpecInt/SpecFloat 2000 benchmarks which are the only supported values in GLUE 1. For storage systems there is more scope to publish detailed information about how much space is used and available. As the grid evolves it is therefore likely that new information, and perhaps new services, will only be available in GLUE 2.

Future developments

For the time being GLUE 1 will remain the default while we gain experience with GLUE 2. The latest version of the Workload Management System (WMS) can match against both GLUE 1 and GLUE 2 attributes in the Job Description Language (JDL), and the data management middleware is also being upgraded to use both formats, as are other tools which use the information.

During 2013 we will do extensive testing to ensure that everything works as expected, and users of the information system will be encouraged to verify that it satisfies their needs, and suggest improvements. The process has taken many years, but we now have an information schema which should be able to cope with new challenges such as the development of cloud services and the use of multicore CPUs and be useful for many years to come.
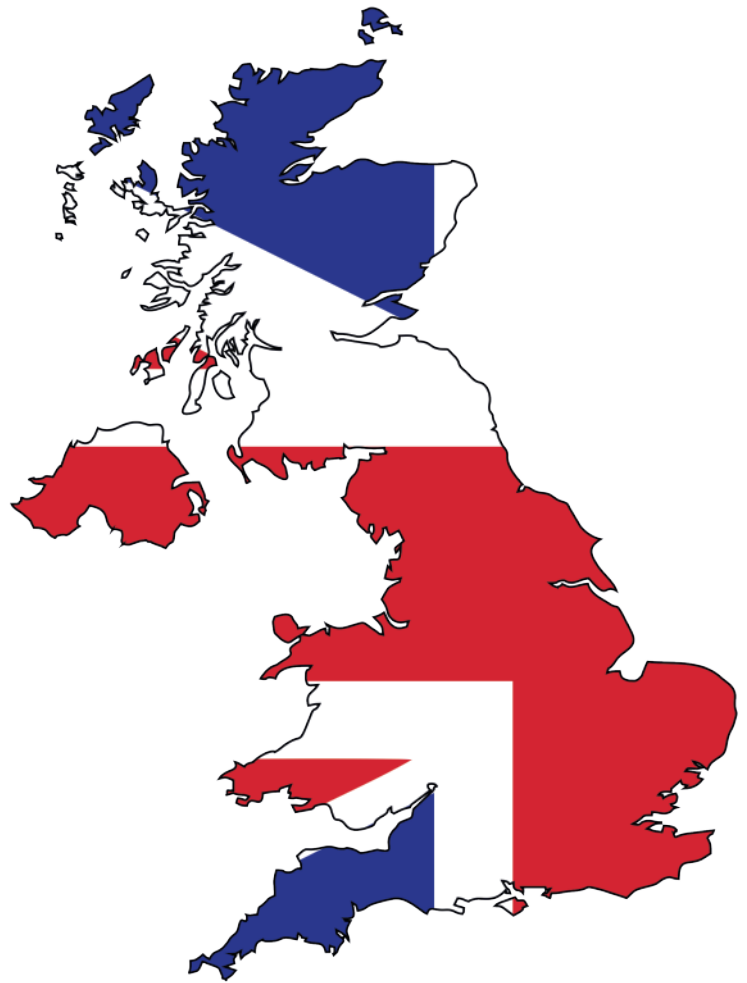
The UK NGI, a united community

Neasan O'Neill catches up with the story so far

The UK has been a member of the European grid since the very beginning; in fact two of the first four sites integrated in the first European grid back in 2001 were based in the UK. Ten years later and the UK is a major provider of grid resources to both EGI and wLCG.

The UK NGI is a partnership between two organisations, GridPP and the National E-infrastructure Service (NES). Their combined effort and expertise allow the organisation to provide cutting edge resources for the UK's research community and further afield.

Funded by the UK's Science and Technology Facilities Council (STFC), GridPP was created in 2001 to provide computing resources for the UK's high-energy physics community.

NES was founded in 2004, as National Grid Service, to complement GridPP's offer and a mission to build a grid community beyond physics.

Both projects are products of the UK's vision for a knowledge economy, founded on the Lisbon strategy laid out by the European Council in 2000. Since then, GridPP and NES have become important tools for many researchers in the UK, with both being involved in a wide range of research and contributing to the global community at all levels.

The research supported by the UK mirrors that of the wider European infrastructure, from the high-energy physics at CERN to how dinosaurs walked and modelling landslides in the Caribbean. But it is not all pure research: the UK is a major contributor to areas such as security policy, outreach, accounting and management. The country will also host the next Community Forum in April.

So for the second time in six years the European grid community will be travelling en masse to the UK. As with the previous forum in 2007 it is the city of Manchester playing host, with the dates set as the 8th to the 12th of April.

The meeting will focus on new technologies and supporting user communities, with topics ranging from virtual research communities and cloud platforms to operations and coordination. The registration for the event is now open at http://go.egi.eu/reg-cf13.

More information

UK NGI - *http://ukngi.ac.uk*

GridPP - *http://gridpp.ac.uk*

NES - *http://ngs.ac.uk*

EGI Champions – a network of ambassadors for the grid

Steve Brewer presents our first Champions

The Champions network will help EGI to broaden both the number and diversity of its users. Acting as ambassadors for the grid, the Champions will complement EGI's outreach efforts by bringing a personal touch to the process of talking to new users.

We will offer the financial support the Champions need to travel to conferences and make direct contact with user communities new to the grid. So how will the EGI Champions network help the rest of the grid user community?

The network will grow as more Champions are recruited and this will produce benefits in a variety of ways.

Welcoming new communities to the grid will also be improved with the Champions personal touch. In the early days of the grid larger numbers of researchers attended summer and winter schools in order to gain a thorough understanding of all levels of the technology. As we move towards simpler, more adaptable interfaces, users will no longer need such deep emersion in all of the tiers of the infrastructure, merely an informed nudge towards the appropriate gateway to the grid.

More information

More about the Champions network and how to become an EGI Champion - *http://go.egi.eu/champs*

The first stars in our Champions network are:

Ashiq Anjum, a computer scientist based at the University of Derby in the UK. His work focuses on optimising distributed computing workflows, algorithms and applications for the biomedical sciences and security research fields.

Eleni Katragkou, a physicist specialised in Atmospheric Physics and lectures at the Aristotle University in Thessaloniki, Greece. She works in regional climate models and their applications to climate and climate change research across Europe. Eleni won this year's Greek L'Oréal-UNESCO For Women in Science award.

Mark Santcroos, based at the Academic Medical Centre in Amsterdam, where he develops science gateways, workflows and applications for the biomedical sciences.

Silvio Pardi, is a computer scientist working on networks, grid and cloud technologies and their application to e-Science. He has a passion for mathematics and after a success with the verification of Goldbach's Conjecture, Silvio is looking for other possibilities to apply grid technologies to mathematical challenges.

Stella Arnaouti, a civil engineer based at the Aristotle University in Thessaloniki, Greece, and studied seismology and soil mechanics. Currently, Stella is working on the stability of the slopes of a mine tailing pond in Poland.

Tomas Kulhanek, based at the Charles University in Prague, Czech Republic, where he is involved in an interdisciplinary research in human voice analysis. He focuses on grid middleware, virtualisation, cloud computing and their applications in the field of computational physiology.

VisIVO: Large-Scale Visualisation on DCIs

Ugo Becciani introduces a visualisation tool for e-Science

A screenshot from the VisIVO visualisation of a 500 MI particle simulation of the large scale structure of the Universe by V. Antonuccio et al (National Institute for Astrophysics Catania). Image source: http://www.oact.inaf.it/visivo/
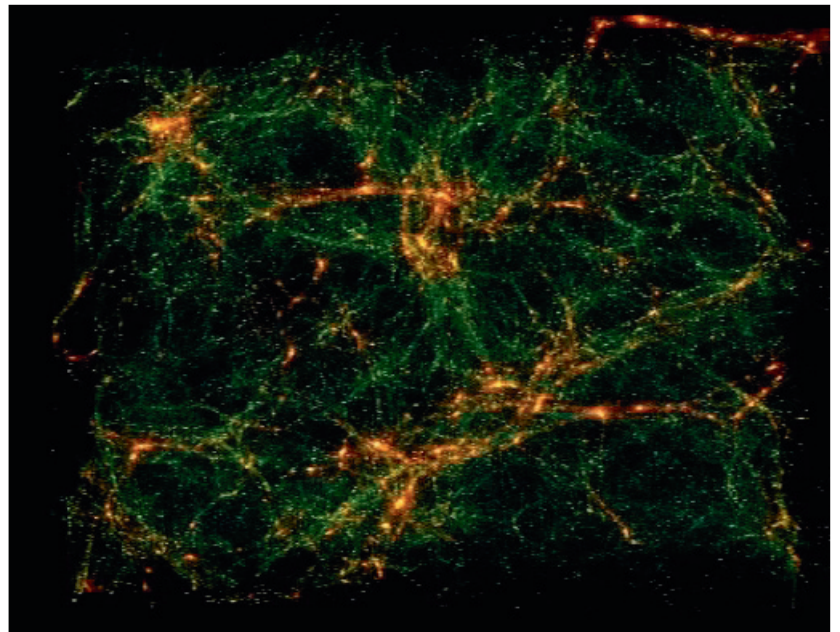
Scientific visualisation represents a critical step in understanding, interpreting, and verifying the outcomes of large-scale multidimensional datasets coming from measurements or simulation tools.

VisIVO is an integrated suite of tools and services for effective visual discovery in currently available (and next-generation) large-scale astrophysical datasets. It is based on the Visualisation ToolKit library (VTK) for multidimensional visualisation.

VisIVO has been ported to gLite middleware by means of a library to allow a job running on a grid node to produce a set of images or movies with its internal data arrays without the need to produce intermediate files. This is particularly important when running on the grid, where the user wants to have a quick look of the results during the data production phase. The images in this way can be produced directly in the grid catalogue, while the user code is running in a worker node.

VisIVO Tools

VisIVO provides an integrated suite of tools and services that allows scientists to visualise highly-complex, large-scale datasets (without limits for the dataset) and create movies of these visualisations based on distributed infrastructures.

VisIVO consists of:

> *VisIVO Desktop* - a stand-alone application for interactive visualisation on standard PCs,

> *VisIVO Server* - a platform for high performance visualisation,

> *VisIVO Mobile* - a mobile application for astrophysical visualisation, connected to the VisIVO Science Gateway

> *VisIVO Library* allows a job running on a computational system (gLite, Desktop Grid, HPC, etc.)

VisIVO Science Gateway

VisIVO Science Gateway provides a web-based workflow-enabled framework for integrating large-scale multidimensional datasets and applications for visualisation and data filtering on DCIs. End-users can create, change, invoke, and monitor workflows provided with easy-to-use, specific, web-based interfaces hiding all the technical aspects of the VisIVO tools and DCIs configurations and settings.

VisIVO Science Gateway is developed on top of WSPGRADE, a highly flexible, cooperative, graphical user interface for the grid User Support Environment (gUSE).

The VisIVO Mobile application share the same accounts of the gateway and offers research groups and communities an exciting possibility to share results and experiences of analysis and exploration of astrophysical datasets within a collaborative visualisation environment.

More information

VisIVO website - *http://visivo.oact.inaf.it:8080/*

MAPPER: Looking at things from all angles

Michael Schiffers explains how

The MAPPER project is using EGI and PRACE resources to provide a better solution to complex scientific problems. Now EGI has included the software that provides core MAPPER services, into their standard software stack, improving researchers access to these tools.

The world, and the science we use to explain it, has a vast range of scales. There are teams who investigate things smaller than an atom, others who are interested in events that can take millennia. If an individual researcher needs to understand a process at more than one level, but when they do, there is multiscale modelling and simulation (MMS).

MMS allows researchers to investigate how small-scale features (for example, local cloud cover) combined with large scale (global air movements) contribute to an overall problem (weather predictions). With traditional computing it can be difficult to combine these effects into a single solution, especially if there needs to be feedback between the computations before the final result. Alongside climate modelling, there are examples from many scientific disciplines including the optimisation of networks of irrigation canals to guarantee adequate water supplies; understanding cardiovascular diseases or simulating nuclear fusion.

MMS is not trivial to implement. Firstly the applications working at different scales are not always able to exchange data and results easily. A second problem comes when the applications need different kinds of computing, sometimes the grid, sometimes supercomputers. So not only do you need applications to be able to communicate you need to do it across completely different infrastructures. MAPPER aims to solve both problems.

MAPPER provides an innovative, but generic, computing framework, a set of tools and services that enable researchers to run multiscale computing across multiple e-Infrastructures such as EGI and PRACE. Importantly they are reusing existing, well-proven software solutions, while extending them with standardised interfaces to high-level programming and execution tools. These are the core of the MAPPER system so that an application can coordinate heterogeneous resources and manage job queues. QosCosGrid, the software that provides core MAPPER middleware services, is now integrated into the standard EGI software stack, improving researchers access to these tools.

The collaboration of MAPPER, EGI and PRACE is a great example of how the European e-Infrastructures can work together to provide the right tools at the right time. Computational science-based research is ubiquitous across academia and most areas of manufacturing and extraction industries, finance, commerce, and government – with innovative, novel or large scale computing resources (including computational grids, clusters or high performance computing), and it is becoming increasingly prevalent.

With the relentless growth in computer power, ever more complex problems become solvable using advanced computer simulation, so the demand for the MAPPER methods, combined with a well-trained workforce with the skills to exploit such capabilities, will increase over the coming decades.

More information

The MAPPER website - *http://www.mapper-project.eu*