

Inspired

ISSUE 17
OCTOBER 2014

news from the EGI community



(Mat / Wikimedia Commons)

TOP STORIES

What is the Open Science Commons?

page 2

Report on the EGI-GÉANT symposium

page 3

User focus: Climate Modelling

page 5

How to do business with SMEs

page 6



**European Grid
Infrastructure**

www.egi.eu

MORE

- 01 EGI - SURFnet collaboration on federated attributes
- 07 Interview: SCI-BUS project
- 08 Outcomes of the DCH-RP project

This Issue

In this issue...

- > Sergio Andreato introduces the Open Science Commons - EGI's new vision for 2020
- > Andres Steijaert reports on the two-day EGI-GÉANT symposium
- > Nuno Ferreira investigates opportunities for collaboration with the Climate Modelling user community
- > Javier Jiménez uncovers an example of how to do business with SMEs

We also have an interview with Peter Kacsuk, coordinator of the SCI-BUS project, an overview of the DCH-RP project's achievements and an introduction to the collaboration between EGI and SURFnet on identity management.

Send your feedback and suggestions to:
sara.coelho@egi.eu Thanks!



With a funding cycle coming to and end, this issue has a special focus on the achievements of EGI's partner projects.
(Matl / Wikimedia Commons)

EGI and SURFnet collaboration on federated attribute management to access e-infrastructures

Peter Solagna introduces the ideas behind the pilot

To authorise users on their sites, service providers in EGI currently use X509 certificates to get information about users' identity and the VOMS extension to get additional information (attributes) about access rights within the given site.

Often, new users approaching the e-Infrastructure already own credentials provided by their home institutions that can be used to authenticate the services they need to access. Usually, these credentials provide a high-level of assurance, which makes them a valid alternative to X509 certificates. For the users who do not own X509 certificates, it would be an advantage to re-use the credentials that they already own, without acquiring yet-another credential to access EGI services.

With this in mind, EGI kicked started a pilot project in

partnership with SURFnet, the Dutch NREN, and the NGIs from the Czech Republic, Greece, Italy, Portugal and Slovenia. The goal is to investigate if SAML federated identity credentials, a widely used, open-source product from the OASIS standards organisation, can be used to bridge EGI services, without using X509 credentials. The main objective is to demonstrate that user communities can manage user membership independently, with a similar workflow as it is now done with the VOMS services.

SURFnet has extensive experience in the management of SAML credentials and attributes. The pilot will leverage on this know how and SURFnet's OpenConext software stack that provides the building blocks to set up a collaboration infrastructure in which identity providers,



attribute providers and service providers can exchange standardised attributes. OpenConext aggregates the information about the user and provides a single entry point for the service providers. In addition, OpenConext also makes sure that users authorise the release of personal information to the service provider they want to use.

EGI is committed to improving users' experience of its services, increase their productivity and lower the barriers to access all e-Infrastructures. Exploring new services to enable the use of federated identities is one of the activities towards these goals. Enabling federated identities, though, is just the first step.

What is the Open Science Commons?

Sergio Andreozzi explains the idea behind EGI's vision for 2020

The new EGI vision statement for an 'Open Science Commons' reads:

Researchers from all disciplines have easy and open access to the innovative digital services, data, knowledge and expertise they need for performing collaborative excellent research leading to a smarter economy, a wiser society and a sustainable environment.

What does this mean in practice?

Open Science refers to the opening of knowledge creation towards all stakeholders, including society in general. Historically, open science started in the 17th Century with the institution of science as a profession and with the birth of a publishing system that ensured recognition for original ideas and motivated scientists to disclose research results. Now, open science, sometimes called Science 2.0, is driven by the digitalisation of the research process and by the globalisation of the scientific communities.

EGI proposes to recognise that e-Infrastructures are a key component of the scientific process that should be managed as a commons.

A **Commons** is a resource management principle by which a resource is shared within a community in a way that allows non-discriminatory access, while ensuring adequate controls to

avoid congestion or depletion when the capacity is limited.

An example is the Wikimedia Commons, an online database created to manage images and media files for Wikipedia in all languages. In the e-Infrastructures sector, the term has been initially associated to the communication infrastructure and later expanded as the e-Infrastructures Commons, which aims to harmonise network, high-performance computing, grid, cloud, and data services. EGI proposes to recognise that e-Infrastructures are a component of the whole scientific process that should be managed as a commons.

The three pillars of the Open Science Commons

The Open Science Commons is built on :

1) e-Infrastructure Commons, an idea developed by the e-IRG to encompass integrated services and interoperable infrastructures across Europe and the world.

2) Open Data Commons, where observations, results and applications can be made available for scientific research and for anyone to use and reuse.

3) Knowledge Commons, in which communities have shared ownership of knowledge and co-create solutions with access to expertise and the technical support they need to exploit state-of-the-art digital services.

In the context of the Open Science Commons, research results become a shared community resource (i.e., a commons). For this to happen,

researchers need to change their own behaviours and support services that simplify the sharing of research results, their discovery and reuse. This is because open access to innovative services for large scale computing and scientific data will only have maximum impact if combined with open access to knowledge and expertise.

The new Open Science Commons vision will be enabled by the richness of the EGI ecosystem, where different actors deliver value to the processes involved into research and innovation.

Resource centres will provide reliable, secure and high-quality computing and data services that are the engines of digital science.

Technology providers will build software solutions to expand what is possible.

NGIs will support e-Infrastructures at national level and engage with the national researchers.

And **EGI.eu** will coordinate the integration of e-Infrastructures at EU level and beyond. Put together, this vision will allow researchers to create the knowledge needed for growth in the environment, in the economy and in society.

What do you think?

How do you relate with this new vision statement? Your opinion is important as this would lead the development of a new strategy for the coming years.

policy@egi.eu

EGI and GÉANT symposium on federated community cloud services for research

On the 25th and 26th of September, EGI and GÉANT hosted a joint symposium on federated community clouds for research, as part of the EGI Autumn Conference. Over 150 participants gathered at the Science Park in Amsterdam to discuss community cloud services for the research community. Andres Steijaert reports on the event.

The two day EGI-GÉANT symposium included presentations from cloud providers and integrators, cloud specialists and users, from both the NGI and NREN domains. The speakers addressed the topics of cloud platforms, security, authentication and authorisation, technology, standards and user experiences and support. This gave a good overview of the state of play, the current capabilities and the elements where the NGI and NREN community can link their efforts and collaborate. EGI and GÉANT presented their cloud activities and emphasised that by working together, we are able to span the research community and bring a cohesive experience across the

cloud stack: network, middleware, authentication, services, delivery and adoption. The opportunities and challenges of the cloud distribution model impact us all. In discussion and panel sessions, the attendees talked about cloud delivery from a providers' perspective and cloud adoption from the users' point of view. The central theme was how to improve the user experience and service uptake.



The symposium made clear that users within research have high expectations on how cloud services are being delivered and used. These expectations are shaped by services offered by big cloud players. On the other hand, users emphasised how their needs differ from commodity and commercial cloud services and that federated community clouds and joint efforts in EGI and GÉANT are important in order to put the right solutions in place - solutions that meet the specific requirements of the community and can be used in a trusted environment. Not only from a technical point of view, but also by providing a governance structure, policies, support model and a funding, metering and billing scheme that fit the community. Several elements emerged as linking pins between NGIs and NRENs, which can be explored in more detail as they help to drive cloud adoption:

The view from EGI.eu

EGI and GÉANT can join forces to help address many of the challenges that research communities are facing when it comes to cloud services, such as identity management, authentication and authorization, trust, technical and legal issues in data sharing and stewardship and scalability of data analytics, just to name a few. I see a growing demand of expertise, technologies and federated services; EGI and GÉANT have a role to play in serving the European Research Area by supporting a wide spectrum of user groups: the long-tail of science, international collaborations and ESFRIs and the private sector.

Tiziana Ferrari, Technical Director, EGI.eu

> **Standards and**

interoperability: to be able to federate community cloud services - standards are a key ingredient. Pinpointing the most relevant standards and stimulating their implementation across NGIs and NRENs will help to achieve interoperability. During the conference, several of those standards and open platforms were identified, such as CDMI, OCCI (rOCCI), SAML2 and OpenStack.

> **Security:** a safe and secure use of cloud services is paramount and the community acknowledges security expertise to safeguard cloud platforms is scarce. It is beneficial to look at how to share resources in this domain and collaborate on security policies, monitoring and vulnerability handling.

> **Identity management and authentication:** providing users with the ability to log-in with their trusted institutional account across a range of cloud services is an essential component to establish a cloud federation and a wider use of community cloud services. There were presentations about eduGAIN, HEXAA, OpenConext

and Perun and the attendants expressed an interest in collaborative follow up activities.

The views from the community

At the end, I talked to Simon Leinen from SWITCH. He said that "it would be great if a platform were established to continue the cooperation that was started in Amsterdam. Because the two communities share an overall goal: to turn the promise of scalable and cost-effective cloud technology into something that scientists can use simply and effectively." There seems to be a consensus within the community that a strong collaboration between EGI and GÉANT, and NGIs and NRENs, will be an important tool to advance our goals. Kostas Koumantaros, who works for the GRNET, both the Greek NREN and NGI, told me that "NRENs AND NGIs are siblings separated at birth, therefore this symposium was similar to an attempt to reconnect in their adult life, which is a necessary process to improve synergies and cooperation between the two

domains, especially in the area of security and trust." And Enol Fernández, from the Spanish NGI thinks that "the NRENs provide the solid foundation for the NGIs and the EGI-GEANT symposium thus allowed to further improve this collaboration by bringing in the common areas where both can cooperate to provide a better cloud service for the scientific community."

More information

The **EGI-GÉANT Symposium on Federated Community Cloud Services for e-Science** took place from 25 to 26 September 2014 and was part of the EGI Conference on Big Data Processing on Cloud. The programme and presentation slides are available online at: <http://indico.egi.eu/indico/conferenceDisplay.py?confId=2160>

Andres Steijaert works at SURFnet and is the Activity Leader for cloud and mobile services in the GÉANT GN3plus project.

TERENA and DANTE joined forces to form the GÉANT Association

In October 2014, TERENA and DANTE agreed to re-organise under a unified governance structure to create the GÉANT Association.

The decision to restructure came from the community of European NRENs, which both DANTE and TERENA serve.

There will be no service disruption while the process of integration is completed and the merger will improve the

relationships between all stakeholders.

The new Interim CEO of the GÉANT Association is Bob Day, Executive Director of Janet, the United Kingdom's NREN, and former chair of DANTE.

The first event of the GÉANT Association will be the TNC 2015 in Porto, Portugal.



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

<https://tnc15.terena.org/>

User Community focus: Climate Limited-area Modelling

Nuno Ferreira attended the CLM Community Assembly in Frankfurt to better understand the challenges from climate researchers.

Following a suggestion made by Eleni Katragkou, EGI Champion in the area of Climate Modelling, I travelled to Frankfurt for the CLM Community Assembly (2-5 September 2014) to find out more about opportunities of cooperation with this user community.

The challenge of climate change projections relies on the estimation of the occurrence of future extreme events, but also on the implications that such events might cause in many sectors of social and economic life.

The Climate Limited-area Modelling (CLM) community is a vibrant international network of scientists, who joined efforts to develop and apply the COSMO model (COSMO-CLM), rooting back to the initial developments in 1999 by the German Meteorological Service, the DWD. This model is relevant for climate research because it is one of the few limited area numerical model systems designed for high spatial resolutions. Even for the most optimist climate change scenarios some level of adaptation is necessary, which is where regional climate systems, such as COSMO-CLM can inform stakeholders of the impacts to expect at the regional to local level.

The CLM community wants to improve the COSMO-CLM model in many aspects, as described in their Science Plan for the next years. Andrew Ferrone, based at the Public Research Centre - Gabriel Lippmann in Luxembourg, explained me their bottom up approach to foster collaboration.

“CLM community members participate actively in several working and project groups, which aim to develop the COSMO-CLM regional climate system model, evaluate the official released versions and offer users with technical support,” he said. “One example was the coordination of the community to perform several climate runs for the latest report of the IPCC.”

EGI preliminary assessment

Jennifer Brauch (DWD) and Barbara Früh (DWD; CLM Community coordinator) introduced me to how this works. All CLM members have to accept the agreement that regulates the cooperation among them. This is of the utmost importance with regards to the development of the COSMO-CLM source code, enforcing developers to follow rules of good scientific practice, while complying with agreed standards for source code development. The end-product

is highly optimised code to run in an HPC environment.

Adaptation of the code to emerging computer technologies was mentioned as an important goal of CLM Science Plan, namely to be able to run the code on hybrid CPU/GPU architectures, such as EGI. Other relevant requirements include the need of access to resources to support European organisations in their endeavours to model climate change and provide support to customers from developing countries. Accounting, monitoring and access policies were also mentioned as critical to achieve these goals.

As we all know, EGI has experience in all of the above mentioned topics and there is scope to strengthen the collaboration with CLM community. We are working together to explore joint work that could lead to a project proposal for H2020.



The CLM Community has 224 scientific members from 53 climate research institutions all over the world, being the largest obliging cooperation in the field of regional climate modelling.

An example of how to do business with SMEs

Javier Jiménez talks to Jesús Marco about the IFCA experience in engaging with SMEs

A key element in EGI's strategy is to develop a working relationship with industry, in particular SMEs. While looking for examples within the community, I was very happy to learn from Jesús Marco from the Institute of Physics of Cantabria (IFCA) that their projects with SMEs and R&D units of large enterprises are a tangible reality. I was interested to know more about how they accomplished this and the results achieved. Here is what I found.

IFCA is a Joint Research Centre of the University of Cantabria and the National Research Council in Spain (CSIC). They offer their data processing and consultancy services to academia, and since 2012 they have also been serving customers from the private sector. These services are provided on a pay-for-use basis and are selected to support research and innovation. Their activity has been regarded as successful in terms of new customer acquisition and knowledge. They have several cases that include local and multinational SMEs, and the R&D units of large international enterprises in projects ranging from creating a dynamic model for a water reservoir, studying the efficiency of industrial gas burners or the diffusion of gas particles in the atmosphere. The cases presented many challenges from the technical point of view, which were solved with effort and a touch of creativity.

In one of those examples, Company ABC (identity undisclosed

to ensure confidentiality) had a project that required running high-broadband consuming graphics. The company was severely limited by poor network connection. IFCA solved the problem by contacting a local SME company (COSMIKAL) who brought a solution that allowed ABC to run their graphics as if they were almost at their own site.

Win-win

By involving a third party, IFCA created a win-win situation with added value for all involved:

- > **ABC** was able to finish the project and improve their products, without large investments in new hardware or broadband connection. They also accelerated the process by relying on the expertise of IFCA, a data computational expert, resulting in a shorter time-to-market.

- > **COSMIKAL** expanded their business by acquiring a new client, which would have not been otherwise possible if the data centre had not made available their resources to industry.



- > And **IFCA** fulfilled its mandate to support the industrial and business activity and increased their technical expertise with a new challenging case.

The success of this story lies on IFCA's role in creating value with the integration of another partner in the value chain. This could also be a good start point for the inception of the so-called innovation spaces, where the research and academic sector, large companies and SMEs, the public sector and society can interact for addressing societal challenges.

I think this is an excellent example of value creation and the return of investment that citizens expect from the taxes they pay. It is also what the European and national funding agencies are targeting with their policies, and what EGI will need to demonstrate as an output of their activities. This is why EGI is working on articulating a formal business engagement programme to support more successful cases such as this one.

SCI-BUS project: achievements and lessons learned

As the project is coming to an end, Elisa Cauhé sat down with Peter Kacsuk to learn more about what SCI-BUS's accomplishments

What was the main achievement of the SCI-BUS project?

SCI-BUS created WS-PGRADE/gUSE - the now most widely-used gateway technology in Europe, by more than 30 user communities and three commercial companies.

Were the scientific requirements easy to deploy?

We learnt that most scientific user communities want to run their computing tasks on many different DCIs (e.g. clusters, grids) and some need several simultaneously. To support this, we developed a DCI Bridge component, which enables access to different types of DCIs through an added plugin. We have developed plugins for all major grid middleware, clusters, clouds and super-computers. This service of the SCI-BUS gateway seemed to be the most popular feature and many communities selected it because of this functionality.

The other major requirement was to access different data storages with different access protocols. We developed the Data Avenue service to enable upload/download/transfer files to/from all the major storage types having one of the fol-

lowing access protocols: HTTP, HTTPS, SFTP, GSIFTP, SRM, iRODs providing increased flexibility for users to access and share research data.

The consortium of SCI-BUS had computer scientists and researchers. How did this work?

I think every project should work like this. Computer scientists and researchers working together led to fast progress in the project. The researchers provided requirements, the computer people developed the gateways and the users gave the feedback needed for improvements. This strong collaboration resulted in 18 releases of the SCI-BUS gateway framework.

How will user communities be able to use the infrastructure after the end of the project?

The SCI-BUS gateway framework is an open source project available on sourceforge (<http://sourceforge.net/project/s/guse/>) and every community has the opportunity to do minor or even larger modifications if it is required for their gateway.

For the next two years, other projects will provide funding for the further development of the core gateway framework.

Nevertheless, it is important to maintain active and live collaboration with all the user communities that applied the SCI-BUS gateway framework to set up their own science gateway and search for new funding possibilities together with them.

How do you envision the role of SCI-BUS technology in the future of EGI?

During the project, EGI successfully connected SCI-BUS with several user communities that have started to use our technology.

Another aspect of the collaboration with EGI is in the area of providing easy access to EGI Federated Cloud resources via the SCI-BUS gateways. Any science gateway based on this technology will also be able to exploit the resources of EGI Federated Cloud both for simple job submissions or workflow applications. In the latter cases, the SCI-BUS gateway technology enables the simultaneous use of several clouds of the EGI Federated Cloud. The SCI-BUS reference gateway operated by SZTAKI is connected to the BIFI, CESNET and SZTAKI clouds, which means EGI users who would like to try this can get an account on the SCI-BUS reference gateway and use it to run applications on the EGI Federated Cloud. The plan is that gradually all the clouds within the EGI Federated Cloud will be accessible by the SCI-BUS reference gateway and by every science gateway that is based on the SCI-BUS technology.

The SCI-BUS project

EC-funded SCI-BUS project aimed to ease the life of e-Scientists by creating a new science gateway customisation methodology based on the generic-purpose gUSE/WS-PGRADE portal family to enable scientists to focus on their work and exploit resources of main Distributed Computing Infrastructures (DCIs) without the need to deal with the underlying infrastructures' details.

<http://www.sci-bus.eu/>

Outcomes of the DCH-RP project

Michel Drescher writes about the key results of the two-year project

In 2012, EGI started working on developing new ways to support a number of its user communities. One of these emerging partners was the Digital Cultural Heritage community, or DCH for short.

Together, with a number of national cultural heritage institutes, public administrations (usually ministries or subordinate institutes) and partnering e-Infrastructure providers, EGI joined the consortium of the Digital Cultural Heritage Roadmap to Preservation project (DCH-RP) EC-funded project, which ran for two years until September 2014.

DCH-RP was part of the European Commission's programme to research preservation concepts and technologies to foster uptake and production use of research outcomes. The project's single most valuable output is a roadmap for the DCH community towards an e-Infrastructure providing preservation services and processes in production conditions.

EGI was responsible for coordinating and supporting memory institutes in their experimentation with a number of preservation technologies, which was organised in a number of proofs of concept. In the first proof of concept, the project looked at software relevant to digital preservation, for example simple website archiving tools (e.g. HTTrack, SWAT, WARC and Web Curator) or document conversion tools, such as AVS

Document Converter. This activity was an important milestone, as it demonstrated the consortium's commitment to change and willingness to take the initiative.

The second proof of concept followed up on the work of the first project year, published in the intermediate version of the DCH Roadmap for Preservation. The focus was on developing services or tools: EUDAT storage services, the SCIDIP-ES HAPPI toolkit for data authenticity and provenance, and an experiment that explored how to deploy a small but complete digital preservation platform on the EGI Federated Cloud infrastructure platform (conducted in collaboration with the APARSEN project).

The results of these experiments were described in two deliverables targeted at the stakeholders and maintainers of the roadmap itself. In an iterative approach, experiments would provide hands-on experience and recommendations to be integrated into a new revision of the roadmap that in turn allowed for the adjustment of technical experiments to fill some more

gaps and uncertainties in the roadmap.

A number of advocacy papers were produced summarising the major take away points of the final DCH-RP roadmap, of which one directed at e-Infrastructure providers. This advocacy paper highlights a few key topics for e-Infrastructures to engage with the Digital Cultural Heritage community as follows:

- 1) Benefits for engaging with the DCH sector
- 2) Services and issues to address
- 3) Recommendations for e-Infrastructure providers
- 4) Roadmap short-term (2015), mid-term (2017) and long-term (2018 and beyond) actions

In particular, the DCH roadmap introduced the concept of a 'Preservation as a Service' platform as a mid-term future activity. This demonstrates the community's serious interest in engaging with e-Infrastructure on a service delivery basis - a clear sign that EGI has taken up in a recent project proposal together with APARSEN to provide exactly such a platform available across all EGI's research communities.



<http://www.dch-rp.eu/>