



Meeting:	DCI Projects Meeting
Date and Time:	Monday 24 Oct 2011 - 10:00-16:00
Venue:	Face-to-face
Agenda:	https://www.egi.eu/indico/event/669

PARTICIPANTS	2
AGENDA BASHING	3
ITEMS OF BUSINESS	3
INTRODUCTION – STEVEN NEWHOUSE	3
EMI – ALBERTO DI MEGLIO	3
IGE – ALEXANDER PAPASPYROU	4
STRATUSLAB – CHARLES LOOMIS	4
EDGI – AD EMMEN	4
VENUS-C – ANDREA MANIERI	5
SIENA – STEPHANIE PARKER	5
DISCUSSION	6
ACTIONS	9



Participants

Name and Surname	Abbr.	Representing
Steven Newhouse	SN	EGI.eu
Michel Drescher	TF	EGI.eu
Steve Brewer	SB	EGI.eu
Sergio Andreozzi	SA	EGI.eu
Alberto Di Meglio	AD	EMI
Andrea Manieri	AM	Venus-C
Nadia Nardi	NN	Venus-C
Charles Loomis	CL	StratusLab
Ad Emmen	AE	EDGI
Alexander Papaspyrou	AP	IGE
Stephanie Parker	SP	SIENA



AGENDA BASHING

No changes requested.

ITEMS OF BUSINESS

Introduction – Steven Newhouse

(See slides) SN set out the context recalling the last year meeting (end of May 2010) and related output document and plan (<https://documents.egi.eu/document/207>); he presented the architecture of the technology stack for EGI in terms of layers and responsibilities; about the platform layer, the idea is to put mechanisms enabling communities to deploy the components that they need; SP asked if this architecture and responsibilities is the result of a survey with communities; SN explained that this comes from informal discussions had in various contexts with the various communities;

EMI – Alberto Di Meglio

(See slides) The sustainability topic was discussed at the last week all-hands meeting; the current presented view is the official view from EMI, any other conflicting position is not relevant; AD started from the starting picture and the strategy to decompose the middleware stacks into components to be pushed into standard distributions (e.g., EPEL, Fedora); most of the components managed by EMI become part of OS distributions (e.g., Fedora, Debian); the idea is that most of the EMI services will be available from them; exploitable elements: software products, documentation, processes/procedures/organizational experience; essential pre-condition for sustainability is to identify what software components are of interest for user communities to be supported and what components have the chance to survive; SN asked how EMI is going to identify those software components; AD answered that at the moment, EMI relies on deployment statistics, but in the future the idea is to use other channels, e.g., specific surveys; a survey run within EMI collaboration board on what components will be retained for usage after EMI resulted in an overwhelming answer: being an NGI within HEP, they are going to use the EMI components as needed by WLCG;

Two main sustainability strategies foreseen as possible:

- **Strategy 1:** at the end of EMI, software is given back to user communities; they will take care of it; the cloud position is very unclear; existing communities needing current grid services do not need cloud services (although they need virtualization); new communities needing cloud do not need current grid services; the collaboration board asked to revise virtualization strategy, investigate what users need when talking about clouds; if EMI can satisfy with current services, OK otherwise it should be considered out of scope; skeptical about current requirements gathering process, not too much interaction with users
- **Strategy 2:** current EMI partners do not have organizational models to operate in any of the standards open business models expect the infrastructure one (see slides where the four models



are explained); the proposal is to start an incubator marketplace where such an organization and structure can be created; idea is to have a foundation dedicated to scientific research related software; it should be sustainable since the beginning, e.g., not through asking EC funding; it should be an independent entity where EMI is a partner; work as Apache or Eclipse; SN commented that this is similar to globus.org; SN said that globus.org did not succeed as much as expected; AD said that probably it was too much bound to globus context, this should not be EMI centered; AD mentioned a similar initiative like open science (<http://www.openscience.org/>); EMI would like to have a dedicated meeting on this topic by the beginning of February

AM mentioned a project submitted to the EC that passed the threshold but not funded; it was about open science book, a social network marketplace for science

IGE – Alexander Papaspyrou

(See slides) One of the activities performed by IGE when repackaging is make the components compliant with EU law (e.g., EU privacy proof); Another type of contribution is training materials, FAQs, workshops. Globus online supported by US, problems when used by EU, e.g. if US government decides to use the Patriot Act, then this is an issue; problem on how to move the service also to European data centers to work under EU law.

The beneficiaries of IGE are: first order of importance, the end-users; second order of importance, middleware providers; third order of importance, infrastructure providers. To transfer the output of IGE back to the community, IGE is establishing 3 things:

1. European Globus Coordination Forum (EGCF): hope to give independent governance; SN said that coordination is not zero-cost; training part to be moved here
2. European Globus Hub (EGH): IGE website
3. European Globus Liaison Office (EGLO): formal communication point to the US

Data privacy and legality issues: people are concerned about putting data into the cloud, especially if being located outside Europe; costs can come from data privacy (e.g., different member states of EU may have different requirements/regulation on data ownership); security breaches and notification to users, US clouds have little support to let users know that a data was accessed (issue from both technical and contractual viewpoint).

StratusLab – Charles Loomis

Project to end May 2012. There is already a first draft sustainability plan (D3.3); in D4.4, more description of assets; (see slides)

EDGI – Ad Emmen

(See slides) Applied value chain analysis to identify exploitable components; identification of activities, three types: cost, revenue generating, external 3rd parties. As part of EDGI activities:

- software infrastructure: connectors, bridges, application repositories



- 'infrastructure' infrastructure

After the project end, expect that connectors will go in the related software projects, so do not expect to need to maintain. Market analysis of possible people/organizations willing to enter the IDGF as paying members; there are universities who are used to pay money for computing resources and have already proved cost savings by using desktop grids. Idea is to have different. The foundation was created a year ago, at the moment no paying members. Expect to generate revenue for 50K-100K E/year from memberships. There are partners who generate revenues from doing training courses, they would benefit from being members not for using the services, but for expanding their customers segment.

Venus-C – Andrea Manieri

(See slides) Venus-C started in a similar period when Microsoft proposed a similar idea to NSF; focus at the platform level, hybrid cloud; Data confidentiality, contract and jurisdiction are the main issues of scientific clouds.

Three foreseen usage models:

1. The research institution asks a service provider; this service provider can use a third-party cloud provider for the infrastructure; 4 scenarios will run in this model
2. The research institution offers local resources where to run the Venus-C platform; 3 scenarios will run in this model
3. An individual scientist will install the Venus-C software on his laptop and he will access directly the cloud infrastructure; 2 scenarios will run in this model

There are 27 cases for experimenting with users. During the review, it was suggested to also add HPC-related use cases.

All the software components will be released under open source license, possibly the same (work in progress to have one). The open source license for the Generic Worker from Microsoft will be defined at the end of the development of this tool.

Market analysis and socio-economic impact will be performed. Evaluating with MS about how to establish an open source community around the developed platform.

SIENA – Stephanie Parker

International Cloud Symposium, last week. SIENA and NIST roadmap are important and will also try to broaden collaboration with Asia. 10 take-aways. Importance on standards from EC perspective; interoperability package; there are different levels of interoperability, policy makers will work specifically on individual layers; economic drivers: openness to benefit more from cloud scale. Technology should be seen more as commodity; focus should be on the provided services where the value is. A new skill set is emerging, e.g., US governments are hiring less technical people and more experts on data architecture and cloud contracts (technology becoming commodity). Next iteration of roadmap, 21 November 2011.



Discussion

Sustainability issues

- Software
 - Money
 - Developers
 - Users
- infrastructure
 - Connecting to users
 - Mechanics
 - Test environment (AD, testing across sites is not available)
- Coordination: people, resources (to keep everything together, e.g., StratusLab)

SN: not all current software within EGI is sustainable; furthermore spending money on a specific segment of the community is difficult as the others would not be happy;

AM: FAO as customer of D4Science services; they are interested in the services in production but they do not see CNR or the rest of the consortium stable enough to support the services; for sure, they do not want to pay for an infrastructure nor for a software; they want to pay for the service

AD view on EGI.eu is that customers are NGIs; NGIs should look for users; EGI.eu needs to provide services for NGIs to make them more efficient; discussion around if users would pay EGI.eu more than NGIs for the services; Services that EGI can provide at EU level, IMO, can only be paid by EC; NGIs should face directly the users as the users are most likely to pay NGIs and to EGI

SN said that large VRCs (e.g., EMBL) have a EU-span and then would pay; geographic dimension vs. scientific dimension

AD said that EC could pay to build capacity in developing areas should be sustained by the EC

AM suggested that we all should do the value chain analysis as done by EDGI

AD: the path should be moving out from EC/centrally funded to national funding; they have more direct visibility on needs and benefits



e.g.: .org is IGI (Italian Grid) while the .com is the national funding agency

VOMS is funded by Italian ministry through IGI; if German NGI needs it, will use as it uses Apache HTTPD

Would VRCs be willing to pay for a service that requires working software delivered on their resources?



ESFRI projects to be established need governing body receiving funding and employing the money to build tools;

EGI.eu needs to plan for sustainability of software components that will be unsupported

AP: agree with AD that software is not the real problem; AD: if the rest is in place, then focus on the software is less relevant

AP: let's assume AD succeeds in putting in place EMI into the open source model, if there is a user community, this will work on maintaining/evolving this software; the successful bits of "open EMI" would be sustained

AD: EGI should have at least a solid base of open source software from which to pick up

SN: the VRC needs to take responsibility; they can put effort/money or just take the software; should EGI be running a EU Forge? Coordinating?

Should EGI take money to develop a new meta-scheduler system? Probably not. Should promote work in the community to do this work.

MD: the current middleware satisfies specific business needs; other communities have different business needs; by moving to virtualized infrastructure, then users can deploy they own software stack; the packaging of the services as VMs is a business service that, for instance, WLGCC will do for its users

AM: in this picture, you seem to be an IaaS, competitor of etc. Amazon;

SN: who should sustain the platform? (the software layer abstracting the complexity of the hardware to the user, e.g., CREAM); the VRCs are candidate

AP: so VRCs look like NGIs in some of their functions;

SN: NGIs aggregate resource centres; VRCs aggregated users and their needs for services; the question is where do these services come from?

Open Science Initiative (AD)

Try to mimic what apache has done without the restrictions that are suitable for science environment; By end of October, AD will release notes from all-hands meeting; AD then to start working on that; to organize a meeting beginning of Feb 2012; to have something in place by June 2012 (e.g., legal entity for non-profit; in France is possible, not very cheap; Netherlands is very affordable; Germany is very expensive); show portal, service prototype, channels for users;

SN: AM will you also build a community around the products? AM said it depends on the success on the different scenarios, Microsoft want to analyze the potential market; Eng wants to explore the option to install D4Science to see if there are customers who want professional services on top of that.



EGI Sustainability

AD: what about EGI sustainability?

SN: main focus now is on securing the bottom layer of the architecture by providing a roadmap, will provide more flexibility and open to grow up the user base; a big part is to define the role of VRCs; the ESFRI projects have structure and money to take their responsibilities; at the national level, NGIs have/will have skills to package VMs and manage for the local communities

Scientific Cloud Initiative

AM: what about the Scientific Cloud Initiative from Bob Jones?

SN will found out at the meeting this week

Open EGI to commercial customers

AM: what about opening the infrastructure to private companies?

SN, having companies contributing resources is viable; consuming resources is more difficult because research networks forbid exploitation of networks for commercial purposes; for pre-competitive research is OK;

AM: also, usage by private companies usually impose extra requirements

EGI contribution to StratusLab after the end of the project

CL: as a customer, would EGI provide in-kind or money to maintain StratusLab software stack after StratusLab?

SN: this is a business case that needs to be presented to the stakeholders

There being no further business, the meeting concluded at 16:00.



ACTIONS

ID	Resp.	Description	When
1	AD/EMI	Circulate Open Science Initiative proposal	When available
2	All	Contribute their sustainability plan EMI: May 2012 StratusLab: May 2012 IGE: no deliverable EDGI: PM24; DEGISCO: PM18 (Nov 2011) Venus-C: May 2012; white paper by the end of the year about legal issues for scientific cloud SIENA: Nov 2011, next iteration of roadmap EGI-InSPIRE: Apr 2012	When available
3	All	Contribute to DCI Collaborative Roadmap <ul style="list-style-type: none">• Add a chapter on how much of the plan was put together• Future perspectives, what collaborative assets/legacy will be left	Feb 2012
4	EGI	Investigate if to provide extra money to maintain the infrastructure layer and explore it there is interest from NGIs to contribute to an infrastructure behind an OSI for the building/testing	
5	All	Value Chain Analysis for each project	Feb 2012?



Minutes prepared by Sergio Andreozzi, 28.11.2011

Minutes Approved Chair Steven Newhouse

COPYRIGHT NOTICE

Copyright © EGI.eu. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. To view a copy of this license, 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 © EGI.eu (www.egi.eu). Using this document in a way and/or for purposes not foreseen in the license, requires the prior written permission of the copyright holders. The information contained in this document represents the views of the copyright holders as of the date such views are published.