

EGI Feedback on D3 - Final Study Report (draft) 'Developing impact measures for e-Infrastructures'

Sergio Andreozzi

Strategy and Policy Manager, EGI.eu

sergio.andreozzi@egi.eu

- EGI
 - Key Info
 - Contribution to the EU policies
- Feedback on survey
- Feedback on final report

- **European Grid Infrastructure (EGI)**
 - Federation of National Grid Infrastructures
 - Represented by EGI.eu (Dutch foundation)
- **National Grid Infrastructures (NGIs)**
 - Sole national point of contact for ‘Grids‘
 - Integration of individual resource centres
- **Current Status:**
 - 35+ NGIs & 350+ resource centres
 - 300,000 CPUs & 200PB+ storage



- Consolidation of 10 years of investments in a pan-European e-Infrastructure for distributed computing
- Built and educated a community to run a professional service
 - Availability & reliability reports, Helpdesk, etc.
- Transition from project-based to a sustainable network of organisations supported by the EGI-InSPIRE project
- Integrates national and EU funding
- An asset for the EU and the community that is available for the RIs to build upon

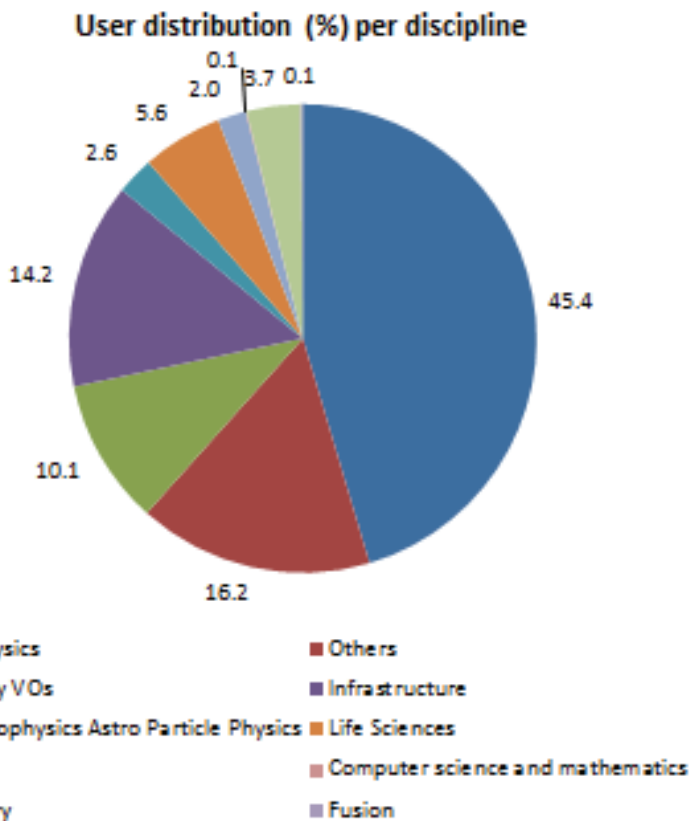
EGI Usage (April 2011)

Average usage 2010-2011 vs 2009-2010

- 27.8M jobs/month, 91,000 jobs/month (+82%)
- 74.8M CPU wall clock hours/month (+35%)
- 2.8M jobs/month for non-HEP users (+47%)

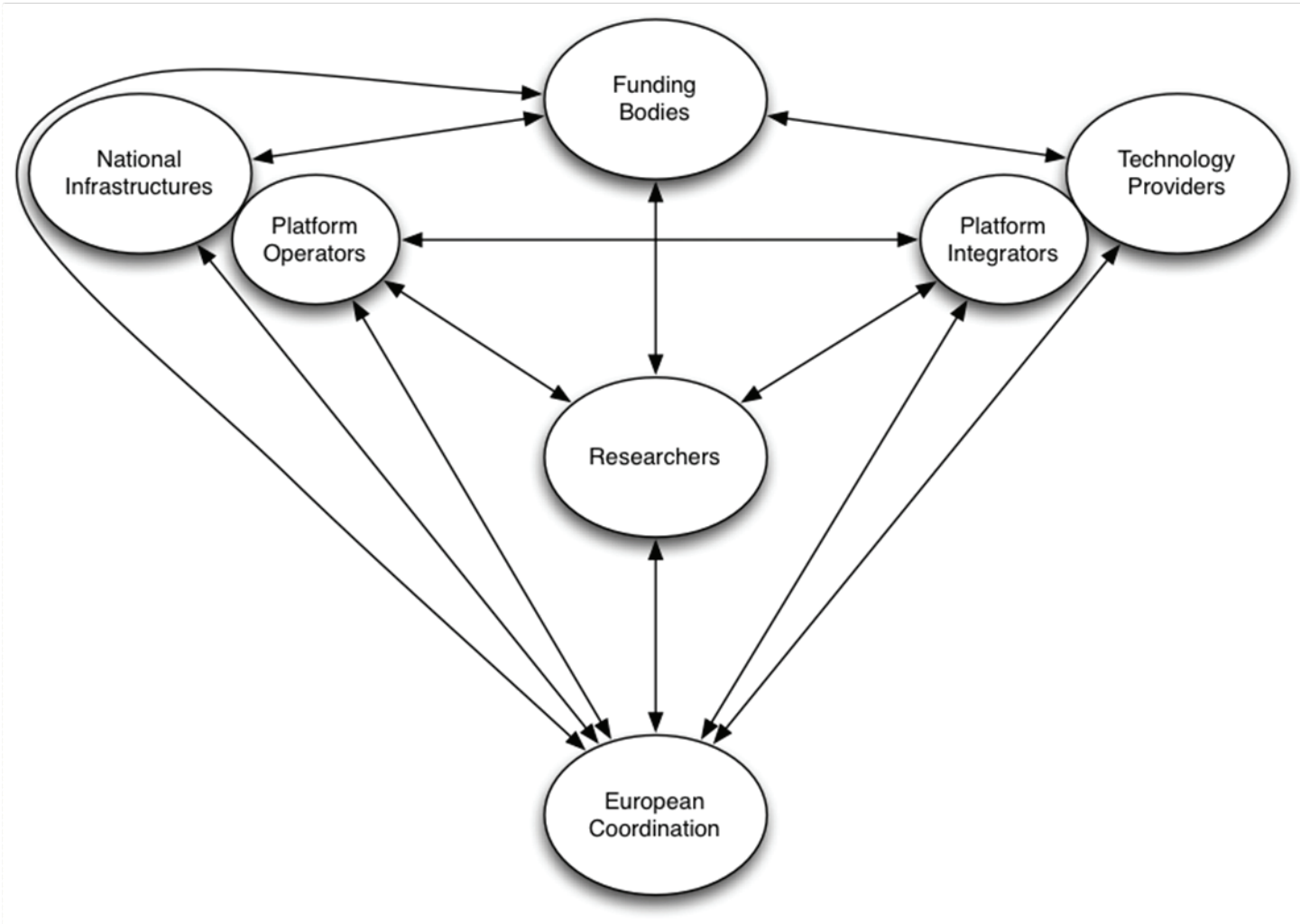
Year to Year Increase

- 18,271 End-users (+47%)
- 219 VOs (+17.7%)
- ~30 high activity VOs (no change)



User Communities

- | | |
|------------------|---------------------|
| Archeology | Fusion |
| Astronomy | Geophysics |
| Astrophysics | High Energy Physics |
| Civil Protection | Life Sciences |
| Comp. Chemistry | Multimedia |
| Earth Sciences | Material Sciences |
| Finance | ... |



- Horizon 2020 (Jul '11):
 - EGI Position Paper
 - <http://go.egi.eu/csf-egi-position-paper>
- Cloud Computing Consultation (Aug '11)
 - EGI answers to survey: <https://documents.egi.eu/document/1011>
 - EU Report: <http://go.egi.eu/eu-cc-consultation>
- ERA Consultation (Nov '11)
 - EGI Position Paper
 - http://go.egi.eu/EGI_ERA

- EU 2020:
 - EGI Role towards Europe 2020 (Jan '11)
 - <https://documents.egi.eu/document/317>
 - Monitoring DAE and IU actions relevant to EGI:
 - https://wiki.egi.eu/wiki/Europe_2020_actions
- European Interoperability Strategy 2.0 (EIS 2.0)
 - Considered in the EGI Standards Roadmap
 - <https://documents.egi.eu/document/721>

- Providing a single market for accessing distributed computing resources and connected data in Europe through EGI's federation of national resource providers
- Promoting competitiveness and interoperability through open standards within the European Interoperability Framework (EIF)
- Reducing inefficient research spending and stimulating innovation across Europe by pooling financial resources and knowledge between different countries within a common infrastructure
- Offering large scale computing facilities that enable the exploration of new computing models and address scientific grand challenges facing society

- Position itself as a key enabler for the online ERA for the free circulation of researchers, knowledge and technology
- Promote excellence in education and skill development by simplifying multi-disciplinary cooperation
- Bridge geographical boundaries beyond Europe thanks to the many collaborations and integration with worldwide e-Infrastructures

Feedback on the Survey

- Effort to answer the survey:
 - Around 60 hours
- Some questions difficult to understand
 - Too generic
 - Not properly defined
- Some questions difficult to answer
 - Missing raw data

- **Question 1.1.3: Queued jobs vs. Real-time accessibility**
 - Not fully clear about the definition of real time accessibility (vs. queued job)
- **Question 2.1.5: Publications per year**
 - Not clear definition of publications (e.g., consider only peer-reviewed scientific publications or include also considered articles in magazines, newsletter, etc.)
- **Question 2.1.8: How many tools for problem solving**
 - At the application level, we develop/provide a web application (AppDB) to create a marketplace for users to share problem solving tools to run in EGI; we do not develop problem solving tools ourselves (network effect)
 - Metrics on 1 Nov '11, AppDB contains 358 applications and 36 tools
- **Section 3/4: Sustainability and Innovation**
 - The assessment should cover the aspect of continuous innovation of the e-Infrastructures for which EGI plans to apply for EC funding

- **Question 4.1.2:** Do private firms include only commercial or also private non-profit organisations? (both)
- **Question 4.1.3** Are user charged for using the infrastructure?
 - Users are not charged for using the infrastructures, but they buy the resources and provision them in the EGI resource centres; EGI takes care of the management/operations; funding for buying capacity may come from national or EU grants
- **Question 4.1.10** Have users reported innovations or patents?
 - We have reported innovations (according to the OSLO definition) but not patents; you may want to consider splitting the question
- **Question 5.1.4** Collaborations with other e-Infrastructures
 - Our collaboration span is captured in this webpage: <http://www.egi.eu/collaboration>
- **Question 6.1.4** Training offering
 - The EGI-InSPIRE project develops/provisions a web site called the ‘Training Marketplace’ to enable the community to share training and learning resources; the project does not develop material, but facilitates the exchange of it (network effect)

Feedback on the Report

- *“It aims at setting up a sustainable European Grid Infrastructure by fusing together the new Distributed Computing Infrastructures (DCIs) which include clouds, supercomputing and desk-top grids”*
- **Suggests to replace with**
 - “It aims at setting up a sustainable pan-European federation of shared computing, storage and data resources from national and intergovernmental resource providers that delivers integrated and secure distributed computing services to European researchers and their international partners.”

- *“According to the survey, there are no additional capital investments needed to keep the infrastructure running or to deliver the same capabilities after the end date of the project”*
- EGI community invests from national sources per year:
 - 10M+ on capital, 9M+ operation, 25M+ staff
- The infrastructure at the end of the project can be continued, nevertheless it will decay without further innovation to meet demands from its users and to remain competitive as new technologies become available
- EGI will compete for future EC funding to bring innovations into a number of areas being identified in the EGI Strategy

- *“It offers the use of the infrastructure free of charge...”*
- The current resource allocation model envisions that
 - User communities acquire funding from national/EU grants for resources
 - User communities interact with NGIs to provision their computing and storage resources into resource centers federated into EGI
 - NGIs will operate the resources on behalf of users and make them remotely and securely accessible and sharable
- Therefore, user communities pay for the resources during the allocation phase and are not charged for the usage
- Some user communities may have opportunistic access to unused resources

- p.67 note 3:
 - Change “For E-Science Grids, we are unable to include an appropriate measure”
 - To “For E-Science Grids, the figure about new users refers to users registered using the global identity system” as presented in p.164
- p.67 table 1:
 - Increase of user base:
 - Should there be a measure for capturing the increase of registered users?
 - Increase of available resources
 - Why no measure for EGI?
- p. 163-164
 - *“Innovative: ... The project has also developed new research standards ~~and patents.~~”*
 - EGI-InSPIRE contributed open standards, but not patents.

- Check acronym spelling
 - EGI-InSPIRE is correct (e.g. p. 47)
- p.83
 - European Interoperability Strategy and Framework (EISF) is two different documents
 1. European Interoperability Strategy (EIS)
 2. European Interoperability Framework (EIF)
- p.32: Add that EGI-InSPIRE project is enabling the transformation from a project-based Grid into a sustainable European Grid

- Framework for a monitoring system
 - We welcome the set up of a framework for monitoring the socio-economic impact assessment of e-Infrastructures in connection with funding programs and EU strategic policies
 - This is a complex matter requiring dedicated expertise
 - Output can help e-Infrastructures refine their governance with appropriate strategic indicators
 - Users should be included in the evaluation process
 - A possible working group to coordinate activities in this area should ensure representation of e-Infrastructures
- Development of a Toolbox
 - We agree on the usefulness of complementary tools
 - Always estimate the burden on e-Infrastructures and envision how to stimulate engagement (e.g., through reimbursing dedicated efforts?)

- Solid results can be obtained if:
 - Socio-economic impact assessment becomes a regular process at the stakeholders' governance level
 - Adding indicators and setting up the measurement process may require considerable effort with results available in a future time
 - Indicators are validated together with the stakeholders
 - A working group with key stakeholders is established
 - Better engagement from participants into the impact assessment process should be stimulated through funding or reimbursement of the required effort

Thank you for your attention!

Questions?

Sergio Andreozzi
sergio.andreozzi@egi.eu

EGI.eu Strategy and Policy Team
policy@egi.eu

EXTRA SLIDES

- **Researchers:** The *individual researcher*, who is interested in exploiting the available e-Infrastructure for their research - ‘me-science’ – by using whatever they can rapidly adapt or integrate to meet their needs in order to publish first and gain the recognition of their peers. At a larger scale, the *research collaboration*, while composed of individual ‘me-science’ researchers, and inheriting many of their self-centred goals, may have sufficient critical mass and coordination that they are able to contribute to and to a limited extent sustain their own community around shared resource goals needed to tackle societal challenges. *Virtual Research Communities* (VRCs) are composed of research groups that span different disciplines in many organisations across different countries that have structured themselves to tackle a ‘grand challenge’ within their own scientific community.
- **Technology Providers:** The technology area within the EGI ecosystem is built upon open-source or commercial software coming from *community* and *generic technology providers* that is put together by *platform integrators* to meet the needs of particular user groups. For instance, the EMI project integrates a platform for high-throughput computing from software that is developed within the project primarily for the EGI community (i.e. community technology providers such as EMI) with software developed outside the EGI community (i.e. generic technology providers such as Apache) to meet particular use cases coming from their target research community (e.g. WLCG).

- **National Infrastructures:** These include the *NGIs* that represent national activities within EGI and undertake national coordination duties through the *resource infrastructure provider*. They have the responsibility to manage and deliver the operational infrastructure coming from the individual *resource centres* within the country. They may also have the responsibility to act as *platform operators* ensuring that any community specific services provided by the resource centres are operating effectively or this responsibility may be undertaken centrally by that community.
- **European Coordination:** Within the EGI ecosystem, the community has recently established an independent legal entity (EGI.eu) to provide *European Coordination* through a defined governance structure and to coordinate on behalf of the community the community's activities.
- **Funding:** The primary source of EGI funding has come from the public sector through *national funding bodies* with additional investment from *European funding bodies*, such as the EC, to support European level integration and structuring. *Commercial organisations* remain a largely untapped source of funding due to legal and logistical concerns and are more likely to be delivering services to the EGI ecosystem in the future rather than purchasing services from the EGI ecosystem. Within EGI a *community funding* scheme has been established for EGI.eu where organisations with the community that benefit from EGI contribute to the coordination costs.