The e-IRG White Paper 2011 public consultation
Comments summary and chapter champions responses
[June 2011]

This document contains complementary material to the e-IRG White Paper, published on 1 July 2011 at the e-IRG website (http://www.e-irg.eu/publications/white-papers.html).

Summary of e-IRG white Paper comments and e-IRG responses

The following table summarises the results of the public consultation of the e-IRG White Paper (between 7th of April 2011 and 10th of May 2011). This document includes the summaries of responses from the e-IRG White Paper section editors, along with the comments in their original form and is downloadable together with the final version of the White Paper under http://www.e-irg.eu/publications/white-papers.html.

Editor, organisation, e-mail, project, date	Sections commented/ Comment (paraphrased)	Summary response
1. Peter Wittenburg, Max Planck Institute for Psycholinguistics, Peter.Wittenburg@mpi.nl CLARIN (ESFRI) 24.04.2011	i. Foreword/ Take into account changing research requirements and link with emerging RIs	i. Amended the foreword, did not include the two topics in the list as they are not treated in the topics of this White Paper
	ii. Introduction and summary/Refer to the cost aspect of infrastructures and their services	ii. Agree that the cost aspect is important but no changes done as the cost is not in detail dealt with in the paper
	iii. e-Infrastructure governance/ "Constitutional" user involvement and participation in costs	iii. Agree with this comment, but this is more seen as a follow up remark and not as a suggestion to change the existing text. The comment is in accordance with the current text which does not refer to a mere advisory role of users. The text speaks about: 'setting the strategy", "influence the decision making process" (p6, a4) and about "getting a key role in setting the strategy" (p7, a3). Peter is right that this is an invitation to translate these principles in practical solutions, which might be different according to the application area involved. The current text covers Peter's remarks and intentions.
	iv. Future of Research Networking/Not clear what "integrated services" means	iv. An adjustment of the wording is done.
	v. Authentication, Authorisation and Accounting/Imminent establishment of VO framework; Concrete measures instead of roadmap; e-Infrastructures	v. Bullet 1 and 3 of the recommendations deal with specific issues, which are related to the attributes of VOs and AAIs. In the actual AAI-Chapter of the White Paper this issue is considered implicitly in bullet 2 ("Interoperation

	to imminently work out common solutions, so that RIs are not forced to reinvent the wheel (referring to Shibboleth interworking issues).	between existing national AAI") and particularly in bullet 3 ("Full support for the management of distributed dynamic Virtual Organizations") of the section "Proposed approach". Concerning bullet 3 the recommendations ask to look for common solutions.
	vi. e-Infrastructure services/ Stimulate competition with commodity providers; Their licences often not acceptable.	vi. Point taken up in the text.
	vii. Data infrastructures (main comments)/ Not clear if one single data infrastructure is implied as a result of a managed approach; Examples that show that a single data infrastructure should not be the target.	vii. Described the term interoperability Update from 'infrastructure' into 'infrastructures' Some of the comments which were too technical have not been incorporated in the White Paper
2. Matti Heikkurinen, Emergence Tech Ltd., matti@emergence-tech.com 28.04.2011	i. General comments/ Recommendations to be highlighted per stakeholder and build on previous ones	i. Point taken, this will be done in the separate recommendations document
	ii. Foreword/ The case for e-Infrastructure might have been better supported than referring to overwhelmed scientists and service orientation	ii. Slight change in foreword
	iii. Introduction and summary/ Supercomputing could be better referred to as enabling technology of spearheading innovation	iii. Point integrated
	iv. e-Infrastructure Governance/ Reflect on models that that would allow users to choose the best services in cases switching would mean moving from centrally funded service to one that is	iv. Funding models: Agree, added a sentence at the end of the change in the governance paragraph to indicate the need for further work on this issue and a new bullet nr 6;

	partially funded from the user budgets	
	- Who decides who are the "leading edge users"?	- There should be not such a thing as a formalised process to select who is a 'leading edge user'. This would certainly kill the innovation process in its earliest stage. The first sentence of the related paragraph is changed accordingly;
	- User-centric approach: it is not clear here why this new funding model would be attractive for the users.	- *Last bullet under "User-centric approach":* a warning has been inserted
	v. Authentication, Authorisation and Accounting/ Realising the recommendations from the governance section would require quite robust and open accounting solution for the e-Infrastructure	v. Accounting has been inserted.
3. Steven Newhouse, EGI.eu, Steven.Newhouse@egi.eu 6.05.2011	i. e-Infrastructure Governance / Definition of e-Infrastructure governance	i. The text makes it now clear that the Wikipedia definition has been used.
	- User representation in EGI	- Text changed
	- ERIC reference not clear	- Changes have been made in this section after discussions
	ii. Authentication, Authorisation and Accounting/No consideration of non-proliferation issues	ii. The non-proliferation issue is important but is too specific to be part of this White Paper.
	iii. Energy and Green IT/Facilitate discussions to develop Green IT best-practices for the e-Infrastructure providers	iii. Recommendation – added
	iv. Data Infrastructures/ The data infrastructure should take into account the experience of existing infrastructures, e.g. EGI and PRACE/DEISA –	iv. Included references to EGI

	updated	
	v. Annex II- Terminology / Comments on EGI related and VO/VRC definitions	v. Terminology adapted
4. Matti Heikkurinen, Emergence Tech Ltd., matti@emergence-tech.com gSLM project 6.05.2011	i. e-Infrastructure Governance / Highlight cross-organisational service level management frameworks as an important issue	i. Service level management: an extra bullet under Openness is added.
	ii. e-Infrastructure Services / Similar comment for the services section.	ii. Adapted in the text
	iii. Authentication, Authorisation and Accounting/ Making the necessary information available for monitoring the quality of services would be most welcome	iii. Introduced in the chapter
5. Philippe Gavilet, CERN Philippe.Gavillet@cern.ch GISELA project 7.05.2011	i. General comments/ Stressed the nature and background of comments (Latin-American context)	i. Note taken
	ii. Foreword / Comments on data deluge (not new), ecosystem (definition required), digital divide and desktop grids (not covered)	ii. Notes taken, no text added
	iii. e-Infrastructure Governance / Very supportive of user-oriented approach. Supported the realisation of an e-Infrastructure survey to gather requirements from DCI end-users.	iii. This remark seems primarily targeted at EGI.eu and goes well beyond the scope of this Chapter and its level of granularity
	iv. Future of Research Networking/Integration of networks and Grids need to be discussed as a whole. This is already promoted in Latin America.	iv. The need for integration of networks and grids is already well covered in the White Paper.
	v. Energy and Green IT/Water cooling solutions	v. The chapter doesn't exclude water cooling as a solution for future data

	and best practices referenced	centres; therefore I treat it rather as a use case than a real comment
6. Kees Neggers, Surfnet, Kees.Neggers@surfnet.nl 8.05.2011	i. Terminology/ Agreement with proposed changes in definitions of VO versus VRCs	i. Change made
7. Alfons Hoekstra, University of Amsterdam, A.G.Hoekstra@uva.nl, MAPPER project 9.05.2011	i. Future of Research Networking/ Emphasise the need to integrate different wide area networks, to provide optimal interoperability performance and a uniform networking interface. An EU-wide network reservation system would be ideal	i. A sentence is added explaining that it is important that network services are also made available via a common interface to facilitate integrated access to other e-infrastructure services. It is felt that the second comment is already covered in the text.
	ii. Authentication, Authorisation and Accounting/ Largely agree with the recommendations	ii. -
	iii. Exascale computing and related software/ Requirements of new programming models, algorithms, tools and libraries should be also taken into account, including backwards compatibility (if possible) Training courses and workshops required besides a cook book.	iii. No changes made
	iv. e-Infrastructure services/Conclusions are geared towards SaaS and Cloud computing that is incompatible with peta/exascale. Recommendations should aim at user requirements; not top down approach. The unconditional addition of more service-based software will make e-infrastructures more	iv. It was certainly not the idea to add services without considering integration and ease of use. No major changes are made.

	complex and likely slower and less reliable. Integration of the existing software framework with an interoperable, user-oriented modular environment. v. Data infrastructures/ We largely agree with the recommendations.	v.-
8. Thomas Brunner, SWITCH, thomas.brunner@switch.ch 10.5.2011	i. Future of Research Networking / GÉANT is the flagship EC project and deserves particular consideration, including from an organisation and political point of view. GLIF is more about technology and innovation.	i. It is felt that the current wording of the document, being also in accordance with earlier e-IRG documents, is a fair representation of the state of affairs.
	ii. Authentication, Authorisation and Accounting / A more simple but striking explanation of the basic concept of AAI, namely the separation of the responsibilities between Identity-Provider and Resource-Provider should be made as White Paper are more targeting policy makers.	ii.
9. Neil Geddes, STFC, neil.geddes@stfc.ac.uk (on behalf of STFC) 10.5.2011	i. General comments / Large number of draft recommendations at different levels (from low level to high-level).	i. Points noted, recommendations have been updated.
	ii. Exascale computing/ Concerned about the recommendation on encouraging the development of a European hardware technology. E-Infrastructure providers should be free to buy the best hardware.	ii. It is felt that the recommendation is misunderstood. It is not the idea to limit the purchasing possibilities. Delegates chose to leave the recommendations as is.

Comments in the original form

1. Peter Wittenburg

Comment on e-IRG White Paper (v2.6) - Part 1 & 2

Part 1

Part 1 is addressing chapter 7 on Data Infrastructures, where I see the need for corrections, despite the many valuable statements.

- At several places the text gives the impression of speaking about "one data infrastructure" as a result of a "managed process". As the recent discussion in Budapest has shown and as the HLEG report pointed out, there needs to be an organic and agile discussion and development process with multiple actors and activities in a competitive scenario. The White Paper needs to make clearer what its position is.
- Also the cited GRDI document is ambivalent in this respect and asks for a "managed" approach when it speaks about an IT driven roadmap to a cost-effective, efficient collaborative data research environment. Here the term "collaborative data research environment (CDI)" as defined by the HLEG is used in a context which was not intended. CDI was meant as an interactive scenario where users, community-oriented research infrastructures and common data service providers will come together always assuming that a number of solutions will emerge and this not only as a result of a top-down discussion process.
- The document states that "yet there is no such existing data infrastructure". I think that this statement is one of the occasions where the document gives the impression to argue for "one data infrastructure". Indeed such an infrastructure does not exist, but the question is whether we see it as a goal to be achieved. I am not the only one who has severe doubts. What we see is that there are many initiatives of different sort:
    - Some of the big organizations (CERN, EBI, ESA, etc) are working on their own solution for data management.
    - Some disciplines are working on establishing principles for improved data lifecycle management.
    - Research organizations such as the Max Planck Society have made far going decisions about data management to help their researchers.
    - A variety of initiatives have been dealing with questions such as quality assessment, policy-rule based replication, data federations etc.
    - Commercial services are around that partly are being used by many researchers. Knowledge which for years was mainly exchanged via publications has economic values which can be seen in the success of publishers. Data/information as one of the major sources for scientific progress and innovation will have even a higher economic value, which means that new commercial players will try hard to get control of information and to offer services, i.e. commercial players will stay in this sector.
- The document correctly states that we have a fragmented situation. Here are some of the major deficits:
    - Many researchers don't have proper and trusted solutions for depositing (sensitive) data. With respect to opportunities there are enormous differences between communities, countries and even research organizations.
    - We lack a systematic approach to essential data-related services which would be accessible for community-oriented research infrastructures.
    - Until recently there was no European "flagship" project that could act as a catalyzer between the different initiatives and foster harmonization. EUDAT is willing to fulfill this role, but it has to be shown whether EUDAT will succeed.
    - Commercial service providers come along with license conditions which are not acceptable for most researchers.

Which are the most promising solutions to overcome this situation knowing that there is no ONE single solution? We cannot answer this question of course. However we know that we cannot "design" such an infrastructure.

- It is not of help if we speak about common data infrastructures without mentioning the types of services it should give and the types of services that are requested by researchers and research infrastructures. Without identifying these services we risk to create another organization associated with lots of overhead and hunting for users. It will be the relevance of certain services, their cost effectiveness and the level of trust that will determine acceptance and persistence. Services must be lean and all cost calculations need to separate cost components[1] (this has been worked out very well in the White Paper at other places). Infrastructures may not become simply other opportunities to apply for funds for all sorts of activities including software development, basic IT research etc.
- The term "interoperability" is used in the document without clarifying what is exactly meant and whether this has any realism. The e-IRG Data Management Task Force report devoted a whole chapter on interoperability which actually is a summarization of the many different facets this term has. Interoperability at semantic level for example between different communities and even different projects in the same community can be impossible to be achieved, i.e. we can only facilitate the attempt of researchers to carry out cross-walks by various means.

Summarizing I would like to stress that chapter 7 in my eyes needs re-writing including the recommendations which are a consequence of what is said beforehand.

Part 2

This part is addressing other aspects mentioned in the White Paper version.

- On page 3 a list of topics are mentioned. It would be good to also stress two points: (1) How to react quickly on the changing requirements being pushed forward from the research needs and (2) How to link with the quickly emerging research infrastructures? Both points are important as the Budapest discussions have shown.
- On page 4 the WP is stressing the point which services with what quality should be delivered. This is a very good point and could be made even more prominent by adding the cost aspect at the beginning and by asking all infrastructures to work towards an explicit list of services so that assessments can be done. I think that cost-explicitness is crucial for the establishment of a successful eco-system of infrastructures.
- On paged 6/7 the WP speaks about user involvement and a user centric approach. Also this is a very good point. The question is how this can be achieved in a satisfying way. If we add the normal "user advisory boards" to governance, we know that this does not have big consequences. User community influence should be more constitutional although I have to admit that I don't have a good solution. Important is the explicit participation in costs as the WP state. Models should be worked out based on good and bad practice. It must be ensured that the research communities can decide how much money they want to devote to infrastructure and which kind of services they find relevant.
- On page 8 the WP says that "IPR should not pose a barrier to such openness". This statement seems to be too strong, since whatever we do "IPR" will form a barrier. We cannot offer information about persons openly to all researchers.

---

[1] We would like to refer to two services for persistent identifiers which can be seen as good examples for lean services. DataCITE offers such a lean service for registering/resolving citable data collections; EPIC offers a similar service for all kinds of scientific data objects. Such services essential for data identity and integrity are offered at a very competitive costs.

- The notion of "integrated services" on page 11 is not quite clear to me. We have fast network services and we may have a data preservation service for example which can only exist based on fast networks. There is no separate integration required. So - it is not quite clear what is meant.
- I like the separate mentioning of AAA and almost all the WP is stating seems to be ok. There are two points of concern emerging from CLARIN needs:
  - The establishment of a VO framework seems to be a MUST, since a harmonization of attributes such as "role" can hardly be achieved. But we need to create it now and it would be a shame if all communities are starting to form VO frameworks. e-IRG can do stimulation, but this needs to be done NOW.
  - In the recommendation the WP states that we need a roadmap in this area. This is much too weak and not satisfying. We need to take concrete measures NOW, since otherwise the research communities will again start inventing own restricted solutions and spend time on issues GEANT was meant to solve. Again: e-IRG needs to stimulate an action NOW.
  - There are topics that have not been tackled yet and no one seems to invest reasonable funds to overcome the gaps. For all those that work on service oriented architectures on the one hand and rely on distributed AA with the help of Shibboleth type of approach on the other, there is no solution of passing user credentials over in service chains. The Grid world is not taking this up, since they argue that it is not their problem - they are using X509 certificates and solved that problem. The GEANT world does not seem to be aware of this problem or does not seem to feel responsible. How can we now prevent another situation where research infrastructures such as CLARIN need to invent their own solutions despite the fact that e-Infrastructures receive so much money to work out common solutions. K. Neggers used the term "agile" to indicate that something is severely wrong. e-IRG needs to make a statement.
- On page 22 there is a statement that e-Infrastructures should not "simply copying commodity services already on offer by commercial providers". It is often important to create competition in monopoly situations and more important the license conditions of commercial providers are often not acceptable. Both require to offer alternatives.

2. Matti Heikkurinen

Dear e-IRG White Paper team,

I read the consultation document with great interest, although I have to admit that due to timing constraints I needed to focus on just few of the chapters for in-depth study.

As a general comments, I thought that each chapter flowed very well, seemed consistent in its analysis and undoubtedly the document serves as a good "snapshot" of the current policy issues in the e-Infrastructure domain. The recommendations could perhaps be highlighted more and be more clearly targeted to stakeholders - perhaps along the lines of the e-IRG roadmap (half a page in the executive summary and two page section in the content)?

It might also be beneficial to review earlier e-IRG recommendations and align white paper recommendations with them; as an example I noticed that the exascale section could build on the previous HPC-related recommendation ("...establishing organisational structures and processes that ensure that Eu- ropean know-how related to exascale computing can be rapidly accessed") and suggest more detailed model for the partnership between users, industry and computer scientists. As another example, the Data Infrastructures chapter could ideally report on progress on the appointment of "data representatives" by the major initiatives mentioned in the roadmap recommendation (e.g. if this function can be considered taken up by the EUDAT consortium).

Neither of the White Paper sections seemed to be in conflict with earlier recommendations, but quite different wording and emphasis made me thing that a reader might end up with a faint suspicion of an organisational amnesia. Even if this is just due to my earlier involvement with the roadmap, I would imagine that an explicit linking with the past recommendations would make the e-IRG message seem more consistent to policy makers.

While reading the governance section with some thought and doing a quick scan of the rest of the document, and ended up with following thoughts (unedited from my notes) for your consideration:

1. The governance section (page 6) hints at the different funding models, however it might be beneficial to reflect on models that would allow users to choose the best services in cases switching would mean moving from centrally funded service to one that is partially funded from the user budgets. To me this sounds like an acute challenge that is on a critical path when trying to convince new user communities to join in the e-Infrastructure.

2. Further on page 6: who decides who are the "leading edge users"? If they are selected to governance positions purely by the virtue of extraordinary demands, one might theoretically create a situation where initial leading edge users have a vested interest in not supporting experimentation that might cause one of the "mass user communities" adapting improved working practices that would make them leading edge user group. Could e.g. ERC's analysis related to advanced research grants be leveraged to make it more likely that emerging "leading edge" user communities are ideantified as early as possible?

3. Page 7, the last bullet point under "User-centric approach": it is not clear here why this new funding model would be attractive for the users. It sounds like in addition to existing e-Infrastructure funding (that in itself may already be seen as being competition for the research funding) the research budgets would be effectively cut by extra payments based on the usage of

the e-Infrastructure. Also, would this cover both CAPEX and OPEX of the e-Infrastructure? Perhaps more details on how e-IRG proposes to make sure that there are no sudden, unintentional funding cuts for the user communities due to the recommendation could be added?

4. It would probably be beneficial if the governance and e-Infrastructure services chapters were more closely aligned and would cross-refer each other, since they are obviously related (and occasionally overlapping). Ideally they would also follow each other in the document. (I should have read also the services chapter in detail, but unfortunately ran out of time)

5. The AAI section mentions accounting in the title but doesn't seem to discuss the issue in the text itself. I would imagine that realising the recommendations from the governance section would require quite robust and open accounting solution for the e-Infrastructure.

I also thought couple of relatively trivial presentation issues would be worth mentioning:

-Foreword issue (second paragraph): case for e-Infrastructure might have better supporting arguments than overwhelmed scientists and service orientation. Neither in itself brings socioeconomic benefits. One could of course mention that the first is a symptom and the latter an enabling technology for increased innovation potential.

-Page 4: to me "Supercomputing stays a hype" sounds like saying "Supercomputing will never live up to its promises". Perhaps it would be better to talk about something like "enabling technology of spearheading innovation"?

Hope these comments were of some value - and looking forward to reading the final version of the White Paper!

Best regards,

   Matti

--
Matti Heikkurinen - MD, Emergence Tech Ltd
matti@emergence-tech.com - http://www.emergence-tech.com/
+41 79 703 7347 - Geneva, Switzerland

3. EGI.eu, Steven Newhouse

Comments on the e-IRG White Paper 2011

N°	Page	§	Feedback
1	5	1	In order to define governance in the e-Infrastructure context, one should avoid using the definition of Corporate Governance taken from Wikipedia. The e-IRG SPG2 Final Legal Issue Report provided a definition that making a distinction between IT Governance and e-Infrastructure Governance. EGI.eu would suggest looking at this e-IRG deliverable and emphasizing the difference between these two terms stipulating why e-Infrastructure Governance is different.
2	5	5	The European Grid Initiative has resulted in the European Grid Infrastructure. Should this be reflected in the document text and definitions? See later comments on the definitions section.
3	6	2	The document is incorrect in stating that EGI only has indirect user representation. European organisations such as CERN and EMBL are members of EGI Council and this is open to other organisations willing to commit to the statutes. Any VRCs (or VOs) willing to commit to a lightweight MoU are able to become voting members of EGI's User Community Board (UCB), thus, they are able to participate in the governance of issues relating to the use of the infrastructure. User representation in also provided for in other policy groups within EGI.
4	8	5	*"on a non-governmental level"* – It should be clearer here that e-IRG wants the governmental representation in ERIC to be removed or relaxed to make pan-European non-governmental collaboration easier. At the moment the comment it is not clear.
5	7	Strategic and Financial concerns	Unnecessary distinctions between national and international e-infrastructures are certainly a concern. E-IRG should also consider the unnecessary separation of governance structures that currently exists between networking, grid, HPC and potentially now data infrastructures. A recent survey within the EGI Council indicated some support for a single e-Infrastructure ERIC that would govern all of these functions and help provide an integrated approach to European e-Infrastructure provision.
6	14	Recommendations	There is no consideration in the recommendations on AAI to help tackle the issues of non-proliferation. This is a great concern to many NGIs within EGI and is becoming a blocking issue to greater cooperative use of e-Infrastructure services.
7	17	Recommendations	We would add one more recommendation related to the development of Green IT best-practice for the e-Infrastructure community by facilitating discussion of these topics. There is a lot of general discussion in this area which seems to deal with more general issues rather than the issues of e-Infrastructure providers.
8	17	Recommendation	Software management procedures will generally have no impact on an operational infrastructure. The focus should be on providing more service management procedures.

9	26	4 (section 7.1)	At various points this section seems to ignore the operational existence of e-Infrastructures that already support large-scale pan-European communities - certainly EGI and probably DEISA/PRACE. There are already multi-institutional organisations that know how to deal with federated data resources, to provide access in a coordinated manner, to manage policies/procedures process through consensus that span organisations. For example, EGI has experience in dealing with large quantity of data through supporting current heavy users of the infrastructure, such as high energy physics, computational chemistry and life sciences. We do not deny that more needs to be done to improve the organisational and technical aspects in this area and the authors are correct in establishing this is an area of crucial importance in the collaboration needed for building up the scientific data e-Infrastructure. However, there should be a greater recognition that we are not starting without some prior experience.
10	29	EGI.eu definition	EGI.eu would like to change EGI.eu definition to: EGI.eu is an organisation established on 8 February 2010 to coordinate and manage the infrastructure (EGI) on behalf of its participants: National Grid Initiatives (NGIs) and European Intergovernmental Research Organisations (EIROs). EGI.eu is a foundation recognised by Dutch law and headquartered in Science Park Amsterdam, the Netherlands with 20 employees.
11	30	European Grid Initiative definition	EGI.eu would like to change European Grid Initiative (EGI) definition to: The European Grid Initiative refers to the planning and community consultation supported through the EGI_DS project to establish a federation of resource providers set up to deliver sustainable, integrated and secure computing services to European researchers and their international partners.
12		European Grid Infrastructure (EGI) definition	The European Grid Infrastructure (EGI) is a federation of national and domain specific resource providers coordinated by EGI.eu that operates seeks to sustainably operate a secure integrated production infrastructure for a multi-disciplinary user community across Europe and their international collaborators.
13	32	VOs definition	There is a conceptual difference between VOs and VRCs that EGI.eu would like to emphasize: Virtual Organisations are the structures implemented in the e-Infrastructure for the on-line access to resources that span multiple administrative and geographical boundaries. Often used in grid computing. Virtual Research Communities are collaborations of 'like-minded' individuals that are grouped by discipline or computational model that can span multiple VOs. Effectively, they are the human representation of a community that complements the VO representation within the e-Infrastructure.

4. gSLM project, Matti Heikkurinen

Dear e-IRG White Paper team,

In addition to sending my personal comments sent earlier, I would like to report on the discussions related to the e-IRG White Paper in the recent gSLM project (http://gslm.eu/) meeting. As a conclusion of these discussions, the gSLM project would like to offer the contribution below for your consideration:

Being involved in the formal definition of the e-Infrastructure services and their management and reliability aspects, the gSLM project welcomes the recommendations of the chapter 6. In general, the project would be very interested in supporting any follow-up activity (task forces or production of other more detailed e-IRG documents) that stem from these recommendations!

We would also suggest mentioning the issue of cross-organisational service level management frameworks as an important issue in the governance chapter (e.g. in the second paragraph on page 6). Perhaps making the issue more explicit in the e-Infrastructure services part (chapter 6) could also be considered. This is motivated by our belief that formalising the quality and management aspects of the current service provision practices and complementing them with tools and procedures from the established IT service management discipline would have many benefits. As an example, this would allow users, service providers and funding agencies understand more clearly what each individual service really stands for in practice, and how they can be studied and analysed in order to determine which services are best suited for each individual use case.

Also when building a consensus in the sensitive multi-party negotiations that touch governance models and financial issues, we would see clear benefits to applying relatively high level of discipline or formalism in clearly specifying what the "services" actually encompass and what are the rights and responsibilities of all parties (also in failure situations). The service level management approach could also play a role in avoiding commercial lock-in situations (mentioned on page 7), as a formal analysis of the reliability of the service level management procedures tends to uncover "single point of failure" issues - also in this category.

Concerning Authentication, Authorisation and Accounting section, we would welcome expanding the discussion to cover also accounting. For example a policy recommendation regarding making the necessary site- and national level information for monitoring the quality of the service available and accessibly through open and well-documented interfaces would be most welcome, as this would build up the technical basis for more comprehensive application of service level management in the e-Infrastructures.

The gSLM project leader (Thomas Schaaf, CC'd) also mentioned that he would be very happy to provide additional details or clarifications, as well as to discuss possible follow-up actions - also beyond the scope of the White Paper consultation.

Best regards,
Matti

Matti Heikkurinen - MD, Emergence Tech Ltd
matti@emergence-tech.com - http://www.emergence-tech.com/
+41 79 703 7347 - Geneva, Switzerland

5. GISELA project, Philippe Gavilet

2. PREAMBLE

Having followed regularly the various e-Infrastructures Reflection Group (e-IRG) Workshops over the course of both the EELA-2 (http://www.eu-eela.eu/) and GISELA (http://www.gisela-grid.eu/) EU funded Projects, we would like to bring a few comments on the e-IRG White paper 2011.
We first acknowledge the quality of the e-IRG White papers, this last one being very complete in its intention to cover crucial issues for the long-term success of e-Infrastructures.
As a preamble we want to stress the multiple nature of our comments:
• Most of them expressed our views considering the situation of e-Infrastructures (i.e. of e-Science) in Latin America;
• A few are directly coming from the professional experience of some of us;
• Several are personal opinions of some of us that we thought fair to witness. They are given in the context of the "Foreword" section.

2.1. FOREWORD

Data deluge: Some of us believe that it is not a new characteristic of today scientific data. The volume of raw data has always been directly linked to the granularity of the measurements that is governed by the precision of the instruments, both being ultimately bounded by the available computing capacity (CPU and storage) to process the data. In other words, mega/gigabytes of raw data produced by the previous generation of instruments were as challenging to analyse with existing computing resources as the Tera / Petabytes produced nowadays by modern equipments and processed in the Distributed Computing Infrastructures.
Hence the data deluge is the natural consequence of the tremendous increase of computing capacity (per €), not an independent phenomenon as it is often presented.
Eco-system: Some of us would appreciate a precise definition of what is a "research ecosystem". Indeed ecosystem is more a life science concept. Its extension to other domains needs to be precisely defined especially if self-sustainability is meant.
Digital divide: It is mentioned in several occasions to stress the capacity of e-Infrastructures to reduce the inter-regional differences. Surprisingly nothing is said on opportunistic desktopbased
computing rather popular in developing countries because it is more affordable. On the other hand, several Desktop Grid Projects are funded by the EU (e.g. DEGISCO).

2.2. INTRODUCTION AND SUMMARY

No special comment.

3. GISELA COMMENTS ON THE E-IRG DOCUMENT SECTIONS

1. e-Infrastructure Governance

"The essence of the proposed approach is that e-Infrastructure governance will have to show a shift towards a user-driven approach"
We are very sensitive to this approach.
As a matter of fact, in 2009, EELA-2 has proposed to the EGEE Collaborating Projects to realise an e-Infrastructure survey to gather from DCI end-users, their appreciation of the usefulness of einfrastructures,
their possible dependability on them for their research, how they see the need for
national (NGI) and international (EGI) coordination, etc.. The objective was to get the widest support for DCI computing as relevant feedback on the existing projects, e.g. EELA-2, and moreover to support ongoing NGI, EGI and LGI initiatives. The survey has been realised by eResearch2020 (www.eResearch2020.eu) in collaboration with Belief (http://www.beliefproject.org/) after agreement of the respective Project Officers. The outcomes of the survey have been acknowledged at the

"eResearch2020: The role of e-Infrastructures in the creation of global virtual research communities"
Workshop held in Brussels on February 24th , 2010.

2. Future of Research Networking

"Networks are an essential element of the e-Infrastructure. The goal of this chapter is to make recommendations for national and international policy makers for the further advancement of research networks, given the expected technological developments, the needs of the user communities and the evolution in the markets for information and communication services."

We are very concerned to discuss Network and Computing as a whole.

Indeed the originality of the GISELA approach is to associate, from the very beginning, CLARA and the Latin American NRENs to the e-Infrastructure activities, currently under the Latin American JRU /

NGI responsibility. It takes advantage of the fact that NRENs are rather recent in Latin America, as compared to Europe, and can more easily integrate e-Infrastructures services to their network services offer.

In fact GISELA is promoting tight integration of Network and Grid services over the Latin American continent, under the auspices of CLARA. This looks rather natural nowadays, thanks to the improved performance of networks, to associate network and computing as the basic components of all etechnologies

(as exemplified by the emergence of Web 2.0 services, Cloud computing, etc.). This integration could not be done easily in Europe, because at the time of creating DCIs, GEANT-2 and European NRENs had, independently of them, their proper history. The situation is rather different in Latin America where CLARA and NRENs have been created more recently. The opportunity to foresee a coupled provision of network and DCI services, now recommended at the e-IRG level, could be envisaged from the beginning in Latin America.

3. Authentication, Authorisation and Accounting

No special comment.

4. Energy and Green IT

"Changing the cooling medium from air to better heat conductors such as water can increase the cooling efficiency. Such a change may be difficult or even impossible to apply in existing facilities as it may be limited by the infrastructure of the data centers or proximity to lake or river."

We are directly interested by the Green Data Centre approach as the solutions chosen in Europe or in the US are generally adopted in Latin America.

One of us (PhG) has coordinated the development of the infrastructures of the Online computer farms of the LHC experiments: typically 2500 CPU boxes (i.e. 10000 cores) housed

underground in rooms not larger than about 80-100 m2.

The only possible approach has been water-cooling. A complete solution has been worked out. It is currently operational at CERN where it can be seen. It is based on a universal (vertically) water-cooled rack now commercialised (http://www.ciat.fr/). It is by far the most optimised solution especially from the density, cooling / energy recuperation and safety point of views, adaptable with minimum modifications to any Data Centre.

References:

• "*A water-cooling solution for PC-racks of the LHC experiments*", Ph.Vannerem, N.Elias, LHCb-2004-035 DAQ http://cdsweb.cern.ch/record/732069?ln=en

• "*Large CPU-farm implementation in a HEP experiment with tight constraints*", L. Brarda, B. Gaidioz, D. Ruffinoni, P. Gavillet, G. Decreuse, RTC'05 Proceedings of the 14th IEEE-NPSS conference on Real time 2005, p325

• We also mention the US Uptime Institute (http://www.uptimeinstitute.org/) that is, in our opinion, one of the best references on Data Centre infrastructures.

5. Exascale computing and related software

No special comment.

6. e-Infrastructure services

No special comment.

7. Data infrastructures

No special comment.

6. Kees Neggers, Surfnet

Dear Rosette,

--cut-

In this light I just want to add that I fully support Steven Newhouse's comment (nr 13) that we need to carefully distinguish between VO's and VCR's and should check the various chapters of the WP on the consistent use of the terms 'virtual' and 'virtualisation'.

Kind regards,

Kees

7. MAPPER project, Alfons Hoekstra

Dear colleagues,

on behalf of the MAPPER project (www.mapper-project.eu) I would like to share some thoughts and ideas stemming from the project.

Best regards,

Alfons Hoekstra
project director MAPPER

We've taken a look at this and here are our thoughts about this.

Chapter 2 Future of Research Networking:

- The focus here lies on an open, federated and flexible approach for wide area networking. However, currently many wide area networks are poorly integrated with each other. We think it's crucial that the document also emphasizes the need to integrate the different wide area networks, and to provide optimal interoperability performance and a uniform networking interface to the user.
- The document recommends to "Develop a transparent and sustainable structure for user participation in network governance.". Some kind of EU-wide network reservation system would be ideal for this.

Chapter 3 Authentication, Authorisation and Accounting
- We largely agree with the e-IRG recommendations here.

Chapter 4 Energy and Green IT
- No comments.

Chapter 5 Exascale computing and related software

- "* Resources must be dedicated to studying new programming models, algorithms and languages, porting software libraries and software tools to exascale environments."

OK, but this recommendation says nothing about the requirements of such models, algorithms, tools and libraries. These developments should emphasize the seamless integration with existing software, so that as far as possible existing code does not need to be rewritten
for each new generation of machines. (Of course, that may sometimes become impossible to realise.)

- "* The partnership between users of exascale computing, industry and computer scientists must be encouraged, and scientists should liaise with programming experts."

We think here it should be "scientists should be given the opportunity to liaise with programming experts". The e-IRG cannot dictate how scientists should behave, but it can recommend for appropriate facilities.

- "* Specialists must create training materials including robust and straightforward to use "cook books" for users, especially for those who are not computer scientists."

A cook book alone will not suffice to attract new users, or educate total outsiders. Training courses and workshops can fill this gap.

Chapter 6 e-Infrastructure services

- The conclusions recommendations in this chapter are heavily geared towards SaaS and Cloud Computing, a service model which is largely incompatible with peta- /exascale computing resources. The recommendations aim rigidly at providing IaaS (Infrastructure as a service) and Cloud Computing abilities to the e-infrastructure as a whole. However, recommendations shouldn't aim primarily at rigidly introducing additional services, but instead aim at the requirements of the users. The recommendations describe a top-down imposed modification of services without any integration of the user communities and their needs, and will therefore likely lead to an enormous waste of resources. This is very unwise and we must avoid falling into such traps for the umpteenth time.

- The unconditional addition of yet more service-based software will make e-Infrastructures more complex, more difficult to maintain, and likely both slower and less reliable as well (considering reliability generally scales inversely with the size of the software stack). Besides, existing individual software components already cover individual needs of users, as it is the lack of an integrated environment that causes most of the critical problems for e-infrastructure users.

What e-infrastructure truly needs is an integration of the existing software frameworks into a connected, interoperable environment that fullfils the needs of the users. This requires a pragmatic and user-centric approach which aims to meet the user requirements by incorporating and strictly enforcing standards in current software tools, and by integrating the individually operating software tools into a seamlessly interoperable and well-connected environment. Such an environment does not exist today. There is no reason why it couldn't be put into production however.

In addition, active efforts should be undertaken to fulfill the user requirements using a *modular* environment of light-weight software tools, so that stability and performance problems can be easily diagnosed and resolved.

Chapter 7 Data infrastructures

- We largely agree with the e-IRG recommendations here.
--

Alfons G. Hoekstra
Section Computational Science
University of Amsterdam
+3120 5257543

http://www.science.uva.nl/research/scs
http://www.science.uva.nl/~alfons

8. Thomas Brunner, SWITCH

Dear Gudmunt and Rosette
Thank you for letting us participate in the White Paper consultation
process. Here are our comments:

Re 2. Future of Research Networking
Géant and Glif are both important aspect of the network, but what we
have to consider is that the two items should be treated at different
levels.

Géant is the basic building block of the European Network, underpinning
the European e-Infrastructure from a technical, but also from an
organizational and political point of view. Remember that Géant is a
flagship project of the EU! Therefore, given its scope and importance
not only at the technical level, it deserves a particular consideration
in the chapter.

On the other side GLIF is an innovative technology. GLIF is an example
of innovation in the frame of the Network, but it is a technological issue.

Our suggestion is therefore to mention both Géant and GLIF, but to
allocate them in the chapter at the right level of abstraction.

Re: 3. Authentication, Authorisation and Accounting
The first paragraph of the AAI-Chapter could be improved by a more
simple but striking explanation of the basic concept of AAI, namely the
separation of the responsibilities between Identity-Provider and
Resource-Provider. Because the addressees of the White Paper are
positioned at the political and strategic level, the message should be
accordingly formulated.

Kind regards

Thomas Brunner
 SWITCH
Serving Swiss Universities
Thomas H. Brunner, Managing Director
 Werdstrasse 2, CH-8021 Zurich, Switzerland
 phone +41 44 268 1515, fax +41 44 268 1568
 thomas.brunner@switch.ch, http://www.switch.ch/

9. Neil Geddes, STFC

STFC response to the IRG White Paper

Overall we find the white difficult to read with a very large number of draft recommendations. These are made at a number of different levels ranging from the technical to the very abstract long term. Many of the recommendations are not specific or simply unclear. For example there are a number that read much more like statements or conclusions than recommendations, with no indication of priority.

The first section of the paper talks about the strategy and funding being user led, and urges that projects and users should budget for costs. It then recognizes that long term guarantees are needed for the infrastructure to appear attractive to users. It is not clear how the gap between these statements was being, or should be, addressed. At a practical level, does this mean that there would be an increased need for (UK research) projects to budget for costs that are currently funded elsewhere (and maybe not even by the UK). How would this increased project burden be addressed? So whilst the principles are reasonable they could have major hidden or unintended implications.

All reference to the public/outreach seems to be in transmit mode, i.e publicizing the benefits of the e-infrastructure. There should be recognition that we should be listening to the views of the public, for example on green computing, or the cloud, and engaging in a two way dialogue.

Building a European data infrastructure risks being top heavy and bureaucratic, and maybe a difficult or impossible thing to do. Intermediate or less grand goals may be more appropriate in this developing area.

Regarding the recommendation to

Encourage the development of European hardware technology in order to compete and cooperate with the currently HPC leading countries.

We are concerned about using this as a way to stimulate or support hardware vendors. Our view is that we should be free to buy the best hardware and then use it to do the best science – the hardware is a means to an end, not an end in itself.

In this same vein, the final recommendation in the HPC section also needs some careful thought. We would not support a strong publicity push for exascale computing in the absence of a strong science/use case, hence this recommendation should be couched in terms of developing and publicizing the science case for exascale computing.