Approach to Sustainable e-Infrastructures: The Case of the Latin American Grid

Roberto BARBERA[1,2,], Francisco BRASILEIRO[3], Diego CARVALHO[4], Inês DUTRA[5], Marcio FAERMAN[6], Philippe GAVILLET[7,8], Herbert HOEGER[9], Maria Jose Lopez POURAILLY[10], Bernard MARECHAL[7,11], Rafael Mayo GARCIA[7], Leandro Neumann CIUFFO[2], Alexandre Nobrega DUARTE[3], Raul Ramos POLLAN[7], Diego SCARDACI[2], Michael STANTON[6]

[1]Dipartimento di Fisica e Astronomia dell'Università di Catania, Catania, Italy
[2]Istituto Nazionale di Fisica Nucleare, Sezione di Catania, Catania, Italy
[3]Universidade Federal de Campina Grande, Campina Grande, Brazil
[4]Centro Federal de Educação Tecnológica, Rio de Janeiro, Brazil
[5] Universidade do Porto, Porto, Portugal
[6]Rede Nacional de Ensino e Pesquisa, Brazil
[7] Centro de Investigaciones Energéticas Medioambientales y Tecnológicas, Spain
[8]CERN Centre Européen pour la Recherche Nucléaire, Genève, Switzerland
[9]Universidad de Los Andes, Mérida, Venezuela
[10]Cooperación Latino-Americana de Redes Avanzadas
[11]Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

Abstract: The EELA-2 (*E*-science grid facility for *E*urope and *L*atin *A*merica) Project, co-funded by the European Commission, is providing a high capacity, production-quality, scalable Grid Facility for a wide spectrum of applications from European and Latin American scientific communities. This paper presents the approach and the strategy being developed to ensure the long-term continuity of the e-Infrastructure, beyond its term, in March 2010. It also discusses how such an approach could inspire African countries in setting up, collaboratively, Grid infrastructures for e-Science.

Keywords: Grid, e-Infrastructure, e-Science, **NGI**: National Grid Initiative, **NREN**: National Research & Educational Network, **EGI**: European Grid Initiative, **LGI**: Latin American Grid Initiative, **GridLGI**: Latin American Grid Infrastructure, **CLARA**: Cooperación Latino Americana de Redes Avanzadas.

1. Introduction

For several years, the relevance of e-Infrastructures to promote e-Science towards several publics (Scientific, Social, Industry, and Government) is becoming a reality in "emerging" countries. Several presentations of Grid projects, amongst them EELA-2 [1], have been contributed at former IST Africa Conferences [2], with the aim of illustrating how Grid technologies are playing a key role in providing substantial distributed computing resources to scientific collaboration members, regardless of geographic location, enabling institutions without enough funding to afford supercomputers or set up computer farms to benefit from available Grid infrastructures to perform research activities in suitable conditions.

The majority of the existing Grid initiatives have been instigated as undertakings supported by regional, national authorities and for the largest ones, as EGEE [3] is the flagship example, as collaborative projects sponsored by International Programmes such as the FP6 [4] and FP7 [5] European Commission Framework Programmes. These initiatives

have demonstrated that e-Infrastructures are becoming indispensable in many scientific domains e.g. Life Sciences, Earth Sciences, and High Energy Physics, among others... They are now mature enough to envisage their future, beyond the term of the corresponding projects i.e. without the current financial support. Therefore the current concern of e-Infrastructures is to establish the conditions of their sustainability. EELA-2, as most of the other e-Infrastructures, is in the process of defining its strategy towards its long-term future, beyond the end of the project, in March 2010.

This paper presents the EELA-2 sustainability plan and the steps already taken to reach, within one year, the autonomous operation and support of the Latin American part of the EELA-2 Grid, the European part of EELA-2 being cared of in the European context.

Although it basically deals with future perspectives of current well-established Grids, the paper illustrates how several key elements, e.g. the model of sustainability, could be most helpful for Grid infrastructures being created in Africa, both in directing them, from the beginning to all relevant issues and in suggesting them to collaborate with existing project(s) during their deployment phase.

The paper is organised as follows. Section 2 presents the Strategy objectives towards a sustainable Latin American Grid. Section 3 develops the concept of long-term sustainability applied to e-Infrastructures. Section 4 presents the model for Operation and Support of the Grid and deduces operation manpower cost. The interactions with the User Communities and Decision Makers to get a sustainable Grid Infrastructure in Latin America are discussed in Sections 5 and 6, respectively. The need for the Grid-related projects to adopt a common approach is argued in Section 7. The progress already achieved and the major remaining challenges, common to all e-Infrastructure's future, are summarised in Section 8.

2.   Strategy Objectives

The Latin American contribution to the whole EELA-2 Grid computing resources amounts to **1521**/3007 (50%) Computing Nodes, located in the **28**/39 Resources Centres (72%) of the **35**/50 (70%) Member Institutions of the **9**/14 (64%) Partner countries.

The EELA-2 strategy to ensure the long-term sustainability of this sub-Grid called the Latin American Grid Infrastructure (GridLGI), beyond the project end, in March 2010, is being developed along four well defined axes, each being split in precise lines of actions:

- Establish the conditions of a sustainable GridLGI by;
    - o Adopting a suitable architecture model;
    - o Anticipating its implementation and evolution;
    - o Deriving the Manpower Effort to operate and support it;
    - o Preparing a transition strategy from EELA-2 to GridLGI.
- Interact with User Communities to:
    - o Progress in the creation of National Grid Initiatives (NGI);
    - o Get a wide consensus on all beneficial aspects of Grid computing and support to sustainable Grid Infrastructures;
    - o Increase the amount of people promoting Grid computing;
    - o Look for new Applications of interest and potential impact.
- Interact with Decision Makers to:
    - o Present the GridLGI;
    - o Disseminate the usefulness and impact of Grid technology;
    - o Work out optimized implementation scenarii in each local context.
- Collaborate with Grid-related projects for common actions towards sustainability, by:
    - o Maintaining permanent concertation to coordinate actions;
    - o Discussing and launching initiatives of common interest.

3. Concept of Long-Term Sustainability Development

Sustainability is, in a general sense, the capacity to maintain a certain process or state indefinitely [7]. The term has its roots in ecology, as the ability of an ecosystem to maintain ecological processes, functions, biodiversity and productivity into the future.

Sustainable development is a pattern of resource use that aim to meet human needs while preserving the environment so these needs can be met both in the present, and in the indefinite future. World's sustainable development is based on three fundamental pillars:

1. Social development, also known as social change, refers to:
   a. Change in social structure: the nature, the social institutions, the social behaviour or the social relations of a community of people, and so on;
   b. Any event or action that affects a group of individuals that have shared values or characteristics.
2. Environmental protection is the process of making sure that current processes of interaction with the environment are pursued with the idea of keeping the environment as pristine as naturally possible;
3. Economic development: is made of three building blocks, i.e. information, integration and participation.

Grid-based e-Infrastructures can be compared to the real world if one makes the following fundamental analogy: **Grids are complex "ecosystems" of services "sold" and "bought" by virtual communities**. This analogy allows to make the correspondences between the real world's and the e-Infrastructures' pillars of sustainable development shown in Table 1.

Table 1: Correspondences Between "Real World and e-Infrastructure"
Pillars of Sustainable Development

Real world	e-Infrastructures
Social development	Virtual Organisations
Environmental protection	Resource Centres + Service Providers = National Grid Initiatives
Economic development	Middleware services, Application support, and Training

By the same analogy, the pillars of e-Infrastructures' sustainable development can then be depicted as shown in Figure 1.
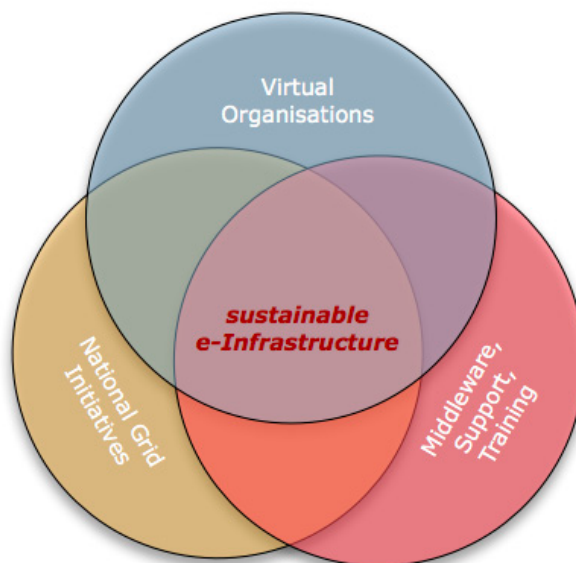

Figure 1: The Pillars of e-Infrastructures Sustainable Development

www.IST-Africa.org/Conference2009

One of the conditions for the real word's sustainable development is the occurrence of a social change (also referred to as social development). By the same analogy stated above, this means that an e-Infrastructure can be sustainable only if a change occurs in the way we consider and support Virtual Organisations of users. Any model of long-term sustainable e-Infrastructures should then put the user at the centre and be scalable and dependable.

The model of a sustainable e-Infrastructure in Latin America and the Caribbean is described in the next section.

4. Conditions of a Sustainable GridLGI e-Infrastructure

The basis of the EELA-2 approach to a sustainable e-Infrastructure has consisted of: i) defining the best model for its future organisation, ii) studying its implementation from its current state, iii) deducing the manpower effort to operate and maintain it on the long term and iv) anticipating a transition phase in the second year of the Project.

4.1   Definition, Adoption of the suitable Architecture Model

The GridLGI is composed of three Grid and Network levels (see Figure 2), each with its well defined functions, its specific operation needs and proper funding source.
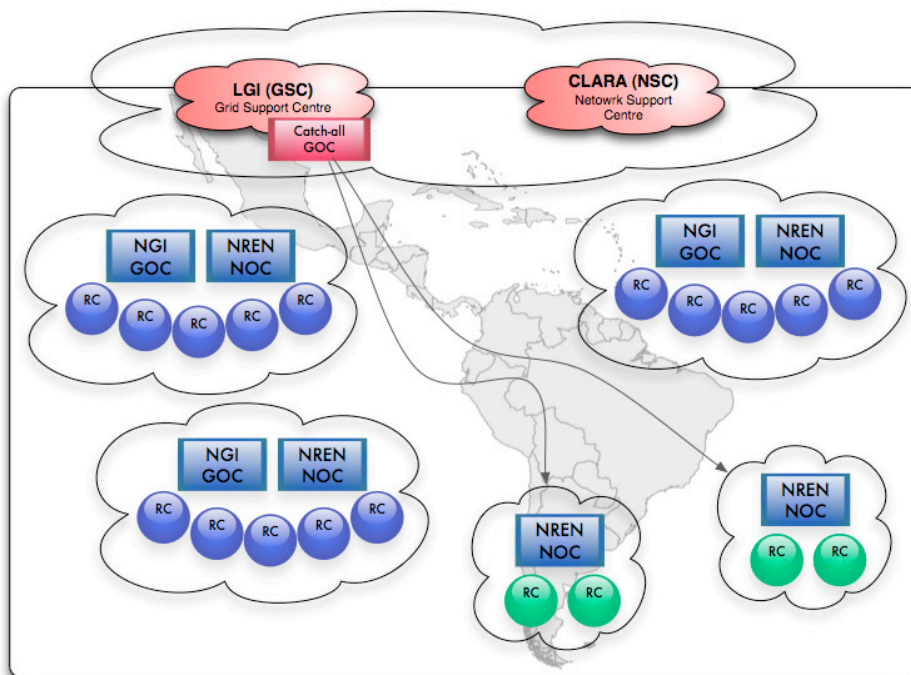

Figure 2: Architecture of the GridLGI

4.1.1   Resource Centre (RC) Level

At the lowest level of a hierarchical Grid Infrastructure sit the Resources Centres providing to their National Grid (GridNGI) their Computing (CE) and Storage Elements (SE) connected to the backbone of the National Research and Education Network (NREN) via the Local Area Network (LAN) and/or the Metropolitan Area Network (MAN) of the housing Member Institution. The main part of the RC operation cost is already part of the baseline budget of the Institution. Only the support of the new Grid-specific Services has to be added to this budget, this for a maximum benefice for the Institution which substantially expands its computing resources by adhering to the GridNGI, as well as improves the performance of the local activities by leveraging on the collaborations for the exchange of data and expertise.

4.1.2   NGI / GOC – NREN / NOC Level

At the intermediate level is the National Grid Initiative (NGI) federating the Institutions members of the Grid Infrastructure in the country (GridNGI) and providing them with Centralised Services through its associated Grid Operation Centre (GOC). The layer capitalises on the well established existence of the National Research and Education Network (NREN) and its Network Operation Centre (NOC) which manages all the GridNGI network transactions and whose use is free of charge in the majority of the Latin American countries.
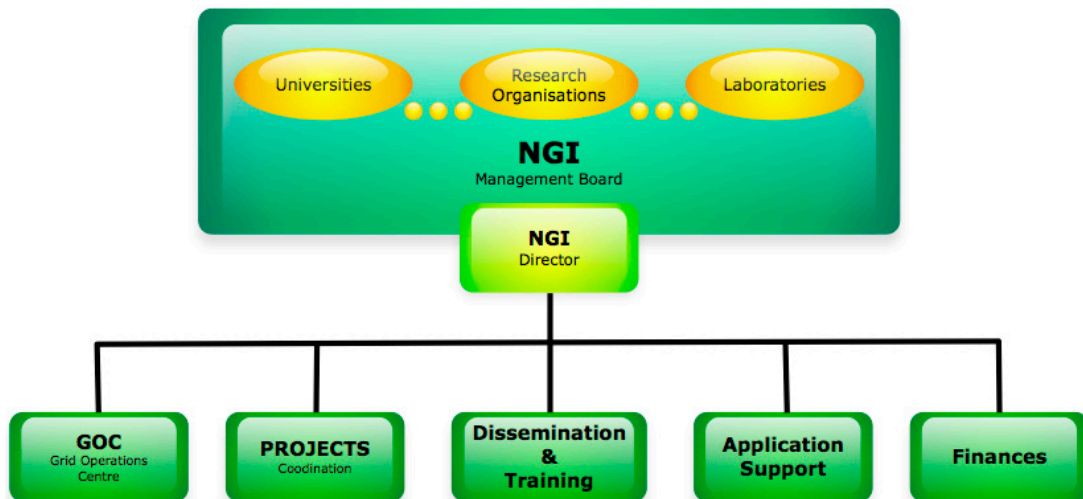

Figure 3: Example of a possible Latin American National Grid Initiative

4.1.3   LGI / GSC – CLARA / NSC Level

At the upper level, a supervisory stage, the Latin American Grid Initiative (LGI), is proposed, similar to the European Grid Initiative (EGI) [8] being implemented in Europe, to care of the overall coordination of the European NGIs. The LGI federates NGIs in a similar way CLARA brings together NRENS, providing respectively the Grid and Network supervision, at the international level, via the associated Grid Support Centre (GSC) and the Network Support Centre (NSC).
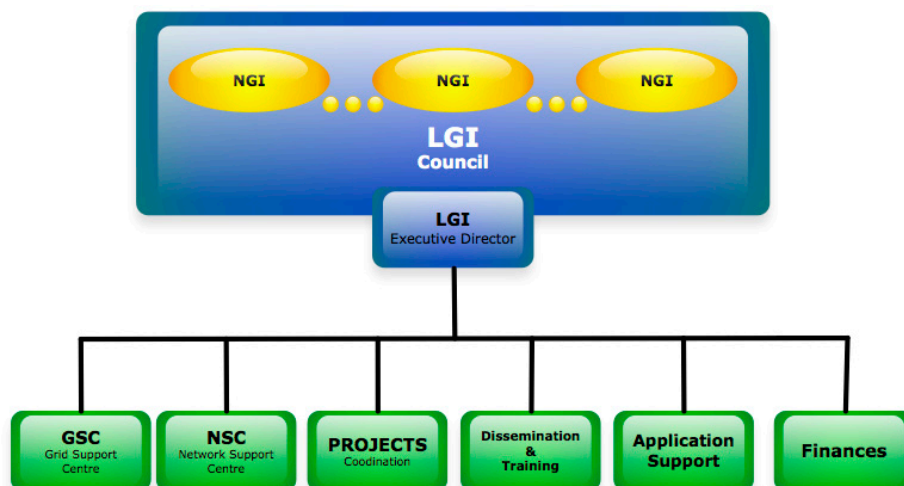

Figure 4: Possible Latin American Grid Initiative

4.1.4 Dissemination & Training

The support to Dissemination and Training activities is foreseen at the three RC, NGI (GOC) and LGI (GSC) levels.

4.1.5 User Support

The future support to users of the GridLGI is seen to be provided through a 3-level scheme, with Local support at the RC site, a GOC User Support service, at the NGI level, in direct contact with the Application developers and a supervision & coordination at the GSC level.

4.2 Implementation Aspects

The implementation of the above architecture will essentially depend upon the state of creation of both the NGIs and the proposed LGI as discussed below:

- **NGI creation**: There are only a few NGIs being created in Latin America and the emergence of new NGIs is difficult to predict. The actions foreseen to overcome this situation, include:
  - Have special sessions of the Collaboration Board Members, e.g. at the occasion of the EELA-2 Conferences, to track the status of NGIs, discuss the conditions of their creation, possibly within the existing NREN structure and agree on a process to monitor their progress;
  - Provide countries without NGI / GOC with a Catch-all GOC (located at the LGI Grid Support Centre (GSC) level) during the phase of setting their own NGI / GOC.
- **LGI creation**: The originality of the EELA-2 approach is to suggest embedding the LGI into the existing CLARA entity. Provided the CLARA agreement is obtained, this would greatly minimize the GridLGI operation manpower effort, while accounting cost effectively for the evolution of Scientific & Educational computing which associates every day more Networks and Computing to answer the community needs.
- **LGI – EGI (and other Large Infrastructures) relationship**: Designated persons of each entity could be appointed as non-voting members of the Council of the other(s)

4.3 Operation Manpower

On the basis of the model described in Sec. 4.1 and of its implementation, the manpower effort necessary for the sustainable long-term operation of the GridLGI e-Infrastructure and its anticipated sources of funding have been estimated (see Table 2). The quoted numbers are for each RC, each GOC and for the GSC & NSC respectively.

Table 2: Manpower Effort

Layer	Type of Operation	FTE	Funding source
RC	Operation of a RC & LAN (~10 to 100 Computing nodes)	1.5[1]	RC Institution
NGI / GOC	GOC Operation	4	NGI
NGI / GOC	Dissemination & Training	2.5	NGI
NGI / GOC	GOC User Support	2	NGI
LGI / GSC	GSC Operation	8	LGI
LGI / GSC	Dissemination & Training	4	LGI
LGI / GSC	GSC User Support	3	LGI
LGI / NSC	NSC Operation	2	LGI

4.4 Transition Strategy

By M16 (July' 09), the EELA-2 Network & Service Activities will have completed their objectives as defined in the Description of Work [6], i.e. the e-Infrastructure should be in steady state running with all anticipated new Application & Infrastructure-related Services available. The last half-year will mainly be dedicated to handover of operational tasks. A transition phase will be put in place from M18 onwards. Its details will depend upon the actual status of implementation of the e-Infrastructure model, but it foresees to set up:

- A management structure with people available to take responsibility beyond EELA-2;
- The deployment, as a function of time, of the agreed Operation scenario;
- The follow up of the transition phase by:
  o Monitoring the level of autonomy of each RC, NGI, LGI levels;
  o Identifying the emerging bottlenecks;
  o Providing the necessary support through the SA1 support scheme.

5. Interaction with User Communities

The interaction with Member Institutions and, through them, with User Communities is essential to establish the long-term sustainability of the e-Infrastructure on well defined and agreed grounds. EELA-2 is striving to have excellent, proactive communication with all its members via its Web site, the EELA-2 Bulletin, Announcements of all events, Notification of the issue of publications (Deliverables,..) and other Dissemination (e.g. Grid schools).

For the second EELA-2 Quarterly report, the Management invited, for the first time, its Members to provide a status report of their EELA-2 activities underlying the work done, the human effort spent and possible open issues after half year of participation in EELA-2. Only 3/50 Institutions did not provide their report.

In this context and keeping in mind the long-lasting process to create Latin American NGIs, the EELA-2 Management strongly believes that NGI and LGI creation will require:

- **Full support of User Communities**. Indeed, although there are 50 Member Institutions in EELA-2, apart from Brazil, there are no more than 2 - 4 EELA-2 Institutions in each country. Hence the support of the whole community will be necessary when discussing with the Decision Makers of a given country. This could take the form of a strong Letter of Intent / Support from each Institution, witnessing its highest interest and need of a sustainable Grid Infrastructure. (In fact this could be a common initiative of Grid-related projects as discussed in Section 7)
- **Supporting Documentation on the advantages of Grid technology, including arguments such as:**
  o Grid technology is crucial, if not mandatory in several scientific fields. The best illustration is certainly to have show cases of success stories that needed the grid to produce highly impacting results. The more interesting cases will be those for which grids are a fundamental technology, i.e. if there was no grid available there would be no way to produce those results;
  o Grid computing has the capacity to attract new communities, to enlarge existing and stimulate the creation of new scientific collaborations;
  o E-Infrastructures provide each Institution the unique opportunity to benefit of huge distributed computing platform for its research activities. Moreover some initiatives such as Ourgrid http://www.ourgrid.org/ which propose a so-called Opportunistic Grid Computing are well adapted to small and medium Research Laboratories such as most of those of Latin America and Africa;
  o Grid computing eases the development of an e-Society (e.g. e-Government);
  o Grid computing goes beyond e-Science and can be used to leverage business opportunities for the industry especially for the SMEs.

All this supporting information should come up from the users themselves because they are the first to judge the impact of Grid technology on their professional life. On the other hand, as this supporting information is common to all Grid communities, it could be prepared in common by Grid-related projects, as suggested in Section 7.

6. Interaction with Decision Makers

In EELA-2, the interaction with Decision Makers takes the double form of:
- The organisation of Decision Makers Days, Workshops and Information Days;
- Direct contact of EELA-2 representatives with local Decision Makers.

In both cases, the motivation is to present the GridLGI in its local, national and international aspects, explaining the model of organisation and justifying the human resources needed at the NGI & LGI levels. Over the remaining one year and a half of the project, EELA-2 intends to iterate with Decision Makers on subjects like:
- The benefits of Grid technology for e-Science and beyond in each country;
- The best scheme to develop e-Science and industry awareness in a given country in the LGI framework;
- The evaluation of the GridLGI support manpower to appreciate with them what would be "feasible" in each national context and thus come up to present well shaped, "reasonable" manpower requests, still corresponding to the needs but with some real chance to be accepted.

7. Collaboration with Grid-related Projects

EELA-2 recognises the high value to collaborate with Grid-related projects, not only because they face common problems that call for common solutions, but because of the international nature of Grid organisations illustrated by the many e-Science Applications spanning several Grid Infrastructures. Dealing with the support to sustainable Grid Infrastructures, EELA-2 would see Grid-related projects:
- Pursuing the ongoing concertation, in the spirit of the EGEE'08 Concertation Meeting, i.e. for coordination of all future activities on long-term sustainability, the next milestone being ICT'08 Conference (25th - 27th November - Lyon, France);
- Envisaging common actions such as:
  o Exchanging views, defining common procedures to assess regularly the progress of NGI creation.
  o Identifying activities which would benefit from joint participation (e.g. development of more Standards tools and procedures for Grid Operation and Monitoring, Dissemination & Training, middleware developments) and thus minimise the corresponding support manpower;
  o Iterate on the evaluation of the NGI and EGI / LGI support to refine the amount of support manpower for each post at all levels.
- Discussing the opportunity to:
  o Have a common survey of User Communities to collect Statements of interest to Grid Technologies and Letter of support to sustainable Grid Infrastructures;
  o Prepare a comprehensive set of "universal" supporting documentation illustrating all the benefits of Grid technology.

8. Conclusions – Outlook

EELA-2 has much progressed over the last months in evaluating the conditions of a sustainable Latin American e-Infrastructure. A model of operation & support has been defined, its gradual implementation anticipated and the necessary Manpower effort has

been estimated. Close interactions with User communities and Decision makers are taking place to guarantee the expected support.

On the other hand, EELA-2 sees two important risks not to get the GridLGI on the long term in Latin America, namely a too few number of NGIs and the difficulty to get the LGI accepted. These risks and possible mitigation actions are briefly discussed below:

- Insufficient number of NGIs by the end of the project.
  This is a real risk as the advance of Grid computing is not yet at the level of what it is in Europe. As outlined in Section 4, EELA-2 has foreseen to:
  o Focus on NGI creation through regular discussion with its Members;
  o Support a Catch-all GOC (located at the LGI Grid Support Centre (GSC) level) during the phase a country is setting its own NGI / GOC.
- Difficulty to get the LGI concept accepted and supported.
  This is another risk as the concept of Continental Grid Initiative is not so widely spread in Latin America. EELA-2 believes to have taken good steps towards a LGI creation by:
  o Suggesting the most cost-effective implementation of the LGI as a Department of CLARA (Partner of EELA-2);
  o Opening discussion with CLARA on the feasibility of such a common Grid - Network Coordination Organisation.

In many respects, the e-Infrastructure situation in Africa compares with Latin America, even if progress is slightly ahead in the Latin American context. This means that African Grids are likely to face the same problems that EELA-2 is experiencing currently, especially in getting necessary User and Decision Makers support to create NGIs and later a continent-wide African Grid Initiative (AGI). The hope is that the EELA-2 endeavour towards sustainability as described in this paper could help the progress of African Grids.

References

[1] www.eu-eela.eu/
[2] B. Marechal, D. Carvalho, L. Ciuffo, I. Dutra, P. Gavillet, R. Mayo. Creating and Operating a Grid Infrastructure: a Use Case from Latin America, IST-Africa 2008 Conference Proceedings, Paul Cunningham and Miriam Cunningham (Eds), IIMC International Information Management Corporation, 2008, ISBN: 978-1-905824-07-6, available at www.ist-africa.eu/home/default.asp?page=paper-repository
[3] www.eu-egee.org/
[4] http://cordis.europa.eu/fp6/dc/index.cfm?fuseaction=UserSite.FP6HomePage
[5] http://cordis.europa.eu/fp7/home_en.html
[6] http://documents.eu-eela.org/record/1086/files/ (DoW: Description of Work)
[7] www.un.org/esa/sustdev/natlinfo/indicators/isd.htm and links therein.
[8] www.eu-egi.eu

---

[i] Only 1/3 of the estimated manpower, i.e. 0.5 FTE, is related to the support for the new Grid Services; the remaining 2/3 are already part of the budget of the Institution.