Subject:	**EGEE REVIEW FEEDBACK ON RECOMMENDATIONS (PROJECT CONFIDENTIAL)**
Author(s):	**Marc-Elian Bégin (editor)** **PEB Activity Leaders** **Project Office**
Distribution	**CB, PEB, PMB**

This note contains the answers to the EU Reviewers' recommendations, following the first EU Review of the EGEE project. In this issue of the note, we give an update (identified with the words *follow-on* at the end of the response) on our previous responses, when appropriate.

Table of content

Note

<div style="text-align: right">

Doc. Identifier:

**EGEE-NA1-TEC-588116-
AnswerToEuReviewersRecomRevi
ew1-v8.0**

Date: **03/10/2005**
</div>

1. RECOMMENDATIONS

1.1. PROJECT MANAGEMENT (NA1)

1. Recommendation: People make the difference. Invest in staff retention immediately. Good staff retention management will be key to success.

    Response: Since the current phase of EGEE represents only the first 2 years of a multi-year proposal, hiring and keeping high-quality staff is one of our priorities. Clearly, the short duration of projects such as EGEE have a negative impact on staff retention. The consortium intends to submit a follow-on proposal to the EU following the success of the 1st EGEE review. This proposal will be submitted in early September 2005 (6 months before the end of EGEE). At this point, we will know the structure of the proposal and requested manpower levels which can be made known to partners. Assuming the response from the EU to the proposal is favourable (i.e. that the consortium is invited for negotiation), we will encourage the major partners to take the risk of renewing/extending contracts for personnel before the negotiation process is completed.

    Follow-on: We submitted the EGEE-II proposal in September. We are planning to have a smooth transition, without a gap between the phases of the project, which should ensure that we do not have staff retention problems.

2. Recommendation: Consider further methods to reinforce project management.

    Response: The project is constantly looking for possibilities to strengthen the project management. The assignment of a Deputy Project Director, funded 100% by CERN resources outside EGEE, is a first step in this direction. The Project Office will be reinforced with additional administrative staff to cope with the present work overload situation.

    We have encouraged the other partners to strengthen their share of the project management structure. The Administrative Federation Meeting, not originally foreseen in the program of work, is funded outside of the funded project resources and is a good example of the general good will in the EGEE consortium.

    Follow-on: During the preparation of the EGEE-II proposal, we have reinforced the management of the project by confirming the role of the Deputy Project Director. We have also introduced a new group, the Technical Coordination Group (TCG), lead by the Technical Director, and including representatives from all technical activities and supported application communities. The TCG will be rolled-out this year, to make sure that its operations are optimal by the time EGEE-II starts.

3. Recommendation: Make document deliverables lighter, better restrict the content to the abstraction level of the respective document. Reduce the redundancy among the different documents wherever possible.

    Response: This recommendation will be taken into account for all deliverables from PM12 onwards.

4. Recommendation: Make the PPT tool used for time and effort recording available to other projects and investigate commercialisation.

    Response: CERN will provide a feasibility study of productising PPT in time for the intermediate focused review following PM18.

    Follow-on: To better support and promote support tools for the execution of large EU project such as EGEE and it's related project, a proposal was submitted in September called Grid Relationship

Note

Doc. Identifier:

**EGEE-NA1-TEC-588116-
AnswerToEuReviewersRecomRevi
ew1-v8.0**

Date: **03/10/2005**

Management (GRM). Here is an extract of the project's abstract: "The Grid infrastructure established by the Enabling Grids for E-sciencE (EGEE) project will continue to be developed by its proposed successor project, EGEE-II. This infrastructure is now the basis for a cluster of 13 Related Projects (RPs), and a at least a further three projects are being proposed in this call. In addition, EGEE collaborates closely with DEISA, a high-performance computing Grid project, for specific applications. This large number of interdependent projects makes effective communication between the projects an important and demanding task"

5. Recommendation: Prove in all activities the added value of largely distributed versus centralised developments.

   Response: In general, the successful advances the project has made in the first period are largely due to the high quality of the personnel involved in the project. Such specialists are not centralised in one location and hence it is necessary for the project to be structured in such a manner to ensure they can contribute effectively. Similarly, there are a number of activities that, due to their nature, need to have a large geographical coverage in order to be able to provide localised support for the infrastructure, applications, training and dissemination. One of the strengths of EGEE is its coordinating, collaborative and consensus building role of the many national programmes as well as related Grid projects. A wide scale adoption of the technology produced by the project requires that all partners in the project contribute to the development of this technology. This large distributed approach also allows for a high level of scalability by being able to integrate resources not otherwise readily available. Clearly, there are intrinsic overheads and additional costs in a large distributed development approach but there are no obvious alternatives and the trend worldwide is going in this direction as demonstrated by leading open source projects such as Linux, Apache, etc.

   Specifically:

   NA1: CERN only but with connections with all partners and including distributed groups such as AFM, CB etc.

   NA2, NA3: experience shows they need a partner in each participating country to cater for local support, cultural and language differences, etc. There are currently a few gaps in the coverage but the activity managers are working to establish informal contacts.

   More specifically, for NA2 the benefit of ensuring that EGEE is disseminated to as wider audience as possible, in as many different geographic areas as possible is strategically important. To manage the dissemination and training centrally would limit the number of audiences targeted and potentially reached.

   SA1 is developing a worldwide infrastructure, which by its nature can only be highly distributed.

   NA4: Needs a large distribution to establish contact with potential user communities and provide localised support.

   For NA5, the distributed nature of the project means that the policy-level interaction between the practitioners and policy-makers will reach as large a community as possible. This large-scale community building ensures that the policies will be based on consensus of the largest possible group of stakeholders.

   JRA1: Distributed among a limited number of development clusters with enough critical mass in each cluster in order to maintain good communication. There is a compromise to be made between a distributed development model and concentrated resources to ensure high quality control and a fast development cycle.

JRA2: Small activities with restricted set of partners providing specific expertise and contacts.

JRA3: restricted to Northern federation, which has a rich expertise in security aspects of Grid computing.

JRA4: Distributed development absorbs effort and distributed project management (at all levels) is expensive. However, in the case of JRA4, there is no single, participating organisation that has all the necessary authority, skills, connections and interests to carry out all the tasks involved. Networks are links between domains. It is not possible to assert a common interface: one can only agree interfaces between interested parties (EGEE/GEANT); and each party has its own, exclusive part in the implementation of the schemes involved. It is thus naturally the mandate of JRA4 (and SA2) to achieve active collaboration with the GN2 and GEANT projects.

1.2. DISSEMINATION AND OUTREACH (NA2)

6. Recommendation: Better capitalise on success stories from all activities through a constant solicitation of the activity leaders. Special emphasis is to be given to innovation in scientific areas triggered by the deployment onto EGEE of key applications.

    Response: Until now, this has been achieved predominantly through the weekly PEB meetings. However, for the future, we will reinforce the collaboration between the Project Office and the NA2 team, such that any success stories within each activity can be quickly turned into news items by NA2.

    Follow-on: A number of application press-releases have been issued since the first EU review. You can find the list here.

7. Recommendation: Improve the appeal of flyers and publicity material to better target executive and politician audiences.

    Response: Much of the material developed so far was developed with a wider community in mind and is not purely targeted at scientific audiences. However, more material will be developed over the next few months targeted particularly at executives, politicians and industry members.

    Follow-on: In order to refocus our high-level public relation message, we have developed a new series of printed and multi-media material. For example, the EGEE brochure and the EGEE short version video have been produced to directly target executives and decision makers.

8. Recommendation: Encourage more participation from the Industry Forum.

    The feedback of Industry Forum partners on various EGEE dissemination and outreach products will be collected in a more systematic way than previously thanks to the Industry Forum technical Working group, led by Yann Guerin (IBM) which has recently been put in operation. In particular, documents concerning gLite software: design, features, performances (and limitations if any) will be circulated to all IF members for comments.

    The Industry Forum Newsletter, which is edited in close partnership with EGEE, will also be used as a vehicle for dissemination and feedback.

    Follow-on: In collaboration with the Industry Forum, the EGEE PO and NA2, 2 specific flyers have been produced, targeting the training t-Infrastructure and the EGEE middleware. Further, EGEE-II includes a stronger link with the CERN openlab, a collaboration between CERN and several high-profile leading IT companies. [Extract from EGEE-II Proposal] "One of the main goals of EGEE-II is to better integrate the Grid middleware with industrial solutions in both hardware and software. As a means of extending the platforms on which EGEE-II is available, the project has established a close link with the CERN openlab project". The collaboration with openlab should be an opportunity to explore Grid Economies.

9. Recommendation: Continue to have strong participation in international meetings and increase presence at key HPC international events (e.g., SC in the US or ISC in Europe).

    Response: With the assistance of the whole EGEE community, EGEE has made excellent progress in targeting key events worldwide. However, a more structured approach will be taken by NA2 towards Event Management to ensure EGEE continues to be promoted at major events. TERENA is currently developing a database of key events to target and a system whereby individuals in the project can request dissemination material for events they are attending. The database will be ready by May 2005.

    Follow-on: Awaiting inputs from Jo on the status of the DB. Awaiting inputs from Hannelore on

the coverage of EGEE at SC2005. The following EGEE partners, with a stand at Super Computing 2005, have agreed to display EGEE material (poster, brochure and folders):

- *Budapest University of Technology and Economics*

- *Argonne National (Healthgrid)*

- *Institut fr Graphische und Parallele Datenverarbeitung der Joh. Kepler Universit t Linz*

10. Recommendation: Publish press releases for each new production-quality service which goes live, portraying its value-add to EGEE user communities.

Response: NA2 has established closer links with NA4 (NA4 now have a dedicated contact for NA2) and will liaise closely with NA4 to be informed of any new applications running on EGEE and can ensure news releases are issued. This model has already been implemented with the issue of a news release about the first industrial application running on EGEE in early March.

Follow-on: See follow-on response to recommendation 6 above.

11. Recommendation: Put more effort into making information sheets available in most European languages.

Response: TERENA, as NA2 lead partner, will encourage all the NA2 partners across Europe to translate as much of the new material as possible. To date 35 information sheets have been translated into three different languages and more Federations are going to translate material in the next two months. Many Federations can share language translations (for example Austria and Germany) and will be encouraged to share the work. TERENA will continue to inform the NA2 partners when an information sheet is written or updated so they can take the necessary actions.

Follow-on: Will be updated with results of translated material from DNA2.6.3 (PM18)

1.3. USER TRAINING AND INDUCTION (NA3)

12. Recommendation: Ensure that training experience remains good and is supported by a high-quality T-Infrastructure and high-quality trainers.

Response: the training and induction team are making strenuous efforts to maintain quality by evaluating all courses and using the feedback to improve material, presentations, curricula and planning. The training consortium throughout the 22 partners undertakes this model and supports each other in its execution.

The t-Infrastructure requirement is well recognised. We are investing in more software development and more clusters to improve facilities and increase resources. Much greater investment could be usefully deployed in developing more widely used and general-purpose t-Infrastructure. This has been the subject of recently submitted SSA proposal.

Follow-on: [Malcolm, David] Provide inputs on t-Infrastructure from EGEE-II and ICEAGE.

13. Recommendation: Give appropriate weight to training activities in future plans, in particular during the migration to gLite phase for application developers and system administrators.

Response: This is in the NA3 plans and is the primary topic of planning at present. It will receive considerable attention at the EGEE conference in Athens and a ramp up of courses with this target is already scheduled.

Follow-on: We published press releases to publicise the success of training events, such as summer schools (see more details here).

14. Recommendation: Improve the exploitation of training success in the dissemination activity. Such success stories are easy to communicate and excellent for executives and politicians.

Response: This is a good idea. We will certainly generate regular training success stories for newsletters, press releases and the project web site.

15. Recommendation: Expand the use of e-Learning through recorded material and demos. Investigate the use of the t-Infrastructure for self-trainee exercises.

Response: We agree that this is important. It is the topic of a train-the-trainers meeting in Athens and of a discussion session. The work is now led by Dr Boon Low, an expert in e-Learning, recruited in Edinburgh to help us meet this goal.

Follow-on: A demo on e-Learning will be held during the EGEE Conference in Pisa in October 2005. [Malcolm, David] Need status of e-Learning.

16. Recommendation: Through a better profiling and follow-up of the trainees, encourage top trainees to become trainers and contribute to training events within their region. Think of an incentive for trainees to become "official" EGEE trainers, like a small reward programme.

Response: This has always been a goal of training, and it has to some extent already been met through "train-the-trainers" courses. It is certainly worthwhile and we endorse the view that it should be the subject of explicit encouragement. We are planning follow-up surveys. We will then select likely candidates and approach them to see what is needed to persuade them to train others.

Follow-on: [Malcolm, David] Need an update.

Note

Doc. Identifier:

**EGEE-NA1-TEC-588116-
AnswerToEuReviewersRecomRevi
ew1-v8.0**

Date: **03/10/2005**

1.4. APPLICATION IDENTIFICATION AND SUPPORT (NA4)

17. Recommendation: Have all current applications migrated to gLite with a very good user satisfaction rating on application development support and grid infrastructure operations.

Response: Migration of current applications to gLite is foreseen in three phases:

- *The common testing with other project activities (JRA1,SA1) will allow to set-up the preproduction service*
- *Once the preproduction service is available, a subset of applications will be validated on this preproduction service.*
- *Once these applications are validated, the other applications will migrate in a progressive manner*

Within this common framework, pilot application areas have slightly different strategies:

- *LHC experiments are already using LCG2 for data challenges. They cannot go through a phase where resources disappear and reappear. The migration will take place when they are convinced that the new system is better than the previous one. This evaluation will occur through the deployment of the preproduction service. Two prototypes are already being migrated to the preproduction service. Final validation will be achieved when the four ARDA prototypes run on the preproduction service. Once the ARDA prototypes are validated, the migration of experiment production software will be accomplished using non-NA4 resources.*
- *7 biomedical applications are deployed on the EGEE infrastructure. Of these the UPV CDSS application has been used for early gLite testing in 2004. Three other applications have expressed interest for early migration to gLite: GPS@, gPTM3D, Telemedicine and these will be moved first to the preproduction service once it is available.*

The migration of the GILDA infrastructure to gLite will start with the deployment of the new middleware in INFN, Catania. Once the site is configured and tested, the GENIUS generic services will be adapted to gLite so that the migration is transparent to the generic applications using GENIUS. GENIUS applications used for tutorials will be adapted first, so that they can be used immediately in gLite induction courses. After the first test, the new middleware will be deployed on all GILDA sites. For all of this to happen successfully, it is necessary that LCG-2 and gLite worker nodes can co-exist on the very same physical machines and the access to existing data is granted. Otherwise, due to the small scale of GILDA, there would be a disrupting interruption of service with strong impact also on NA3 activities

Based upon past successful experience, the migration of the GENIUS services of the "official" EGEE generic applications from LCG-2 to gLite will be done during 'retreats' between the GILDA team and the experts of the corresponding communities. The timing and duration of these retreats will be agreed upon with the new communities depending on the timescale of their readiness and willingness to adopt the services of the new middleware.

However, tests of the new middleware by the "official" EGEE generic applications communities and the transition from LCG-2 to gLite should not impact the current process of deployment on the LCG-2/EGEE-0 infrastructure.

Follow-on: Several gLite services have already been deployed in Production. R-GMA, the information system is used by LCG for monitoring. The new File Transfer Service (FTS) is being used as a cornerstone of the LCG Service Challenges, responsible for the heavy data transfer

Note

Doc. Identifier:

**EGEE-NA1-TEC-588116-
AnswerToEuReviewersRecomRevi
ew1-v8.0**

Date: **03/10/2005**

between CERN and the tier-1 centres. Meanwhile, the full gLite middleware stack is being deployed on the Pre-Production Service, and tested, in preparation for deployment on the Production Service.

The PPS is now being used by CMS HEP experiment and the DILIGENT project. Further, DILIGENT has deployed gLite on a separate testbed, in order to build their own infrastructure. gLite is also deployed on the GILDA testbed and has already been used at a number of training events over the summer.

Missing something on application migration.

18. Recommendation: Building on the experiences of previous FP5 Grid projects, capture full requirements of future user groups, assess needs for new Grid services and plan accordingly for later implementation.

Response: Requirements capture is addressed within EGEE by the Project Technical Forum (PTF). The PTF has responsibilities in the following technical areas: service APIs and requirements. By design the group contains representatives from all activities and is intended to be a forum where a consensus on technical issues can be forged. Requirements are managed in a central database and prioritised by the PTF.

Since the beginning of the project, external user groups have been invited to contribute to EGEE. As a result, three European projects (GRACE, DILIGENT, SEE-GRID) have contributed to the user requirement gathering process.

Some workflow system from external projects are currently being investigated. In particular, this is the case of Taverna from the MyGrid UK e-Science project and Triana from the GridLab EU project.

Applications from different user domains (e.g. gPTM3D from Biomedical and DIRAC from LHCb) share similar requirements and have shared similar Grid application solutions. Following from this type of exchange, we are investigating products from the CrossGrid EU project (in particular the glogin connectivity service), to address the needs for connectivity issues raised by the above mentioned applications.

Follow-on: Following the dissolution of the PTF, the project has decided to create a Technical Coordination Group, including dedicated manpower from all technical activities and led by the new Technical Director. As many new mechanisms devised for EGEE-II, we are planning to try out this new group before the end of EGEE, in order to make sure it is fully operational by the time EGEE-II starts. One of the goals of this new group is to plan the direction the EGEE middleware needs to take, according to inputs from all stakeholders.

Further, several middleware external contributions are being ported to interface with gLite, such as GridBench, glogin, the ProActive Framework. MyGrid???

19. Recommendation: Clarify the true motivation of users from new application areas right from the beginning. In the extreme, are users from the new areas really interested in applying Grid technologies to explore new ways to produce more and better scientific results or simply interested in accessing large and cheap resources?

Response: The initial philosophy for the application identification process was to attract a broad spectrum of applications to verify if the Grid model proposed by EGEE is of general interest. To date more than 6 supported application domains are active, and over 50 virtual organisations exist, on the production service. In this two step process, the EGEE Generic Application Advisory Panel (EGAAP) was foreseen to be a first filter acting as a scientific committee. The second technical filter was provided by the Project Executive Board (PEB).

Note

Doc. Identifier:

**EGEE-NA1-TEC-588116-
AnswerToEuReviewersRecomRevi
ew1-v8.0**

Date: **03/10/2005**

To clarify the true motivation of users, we propose to generalize the usage of Memoranda of Understanding (MoUs) between EGEE and the new users. These MoUs should provide the following information:

- *Identification of the main actors in the different organizations: on the EGEE side, the contact points in NA4, NA3 and SA1, and on new users side, the names of application supervisor, VO manager and main application developer.*
- *Expression of application needs (CPU, storage, training, special requirements for networking, software etc) and evaluation of their impact on EGEE*
- *Resources committed by the application partners and by EGEE*
- *EGEE and application commitments with deadlines on specific tasks*
- *Clarification of the 'Grid added value' for their application*
- *First definition of basic metrics for evaluating user community satisfaction*

A template document has been produced and is being circulated to a subset of the generic applications already deployed on the infrastructure. The goal is to have these MoUs documented for all generic applications by PM18.

The next EGAAP session organized during the Athens EGEE conference will be devoted to discussions with the applications already running on EGEE. Foreseen topics include user satisfaction as well as Grid added value obtained by using the EGEE infrastructure. For the future EGAAP call, all new applications will be required to fill their part of the MoU before submission to EGAAP.

Follow-on: Mention results of user survey and EGAAP findings

1.5. POLICY AND INTERNATIONAL COOPERATION (NA5)

20. Recommendation: Explore collaborations with major national Grid projects in the USA, Japan, as well as other regions of the world.

    Response: The US collaboration was presented by Miron Livny (from University of Wisconsin) during the review, while discussions with Japanese counterparts are ongoing.

    EGEE (mainly through JRA1 and SA1) and LCG are actively collaborating with the US on different topics. Topics covered by this collaboration include security, Grid service interoperability, share of design information, etc. JRA1 have also a direct contact with Globus, and Condor for specific gLite related development

    Japan: The Project Director was invited after the first project review to present EGEE at the annual Naregi conference and profited from this occasion to launch a number of exploratory activities for collaboration between the two projects. These activities will be further discussed at the third EGEE conference in Athens.

    Following the success of the extension model used by SEE-GRID, extensions of EGEE to new regions are being proposed for the Baltic area, South America, China, Mediterranean countries, etc.

    Follow-on: Mention status with OSG and Naregi and related infrastructure projects submitted.

21. Recommendation: Establish a significant presence at meetings such as ISC and SC, either using as an exemplar or even possibly leveraging existing European projects such as the UK e-Science body.

Note

Response: Key EGEE members are participating in the work of the UK e-Science. Several presentations were given at IST last year, including demos, and we are planning to continue our support. Through our partners, we are also actively participating to events such as SC. This information is included in NA2 event list planning and reports.

Follow-on: See follow-on response to recommendation 9 above.

22. Recommendation: Establish top-level collaboration with GGF and other standards bodies so that operational expertise would contribute more effectively and directly to the standardization process in a systematic fashion.

Response: The Project Director and a number of EGEE activity leaders and key project members are involved in international standards bodies and boards. For example, members in EGEE are actively contributing to GGF and occupy important roles as follows:

- *GGF External Advisory Committee member: Fabrizio Gagliardi*
- *(co)Directors in GGF Steering Groups:*
  - *Olle Mulmo: Security Area (he's also co-author of the OCSP recommendations document)*
  - *Peter Clark: Data Area*
  - *Malcolm Atkinson: member*
- *Member of:*
  - *CA Operations Working Group (Olle Mulmo)*
  - *Chairs of both GSM working groups at GGF12 (Peter Kunszt/JRA1)*

- *Participation/contribution to:*

  - *GSM - standardization on Storage Management Interfaces*

  - *DAIS - Data Access and Integration; Database access for Grids*

  - *GFS - Grid File Systems*

  - *INFOD - service discovery, eventing, transactions*

  - *OREP - OGSA replication group, data catalogs*

  - *OGSA - architecture*

  - *NM-WG - Network Measurements Working Group*

  - *OGSA-DAI - Open Grid Services Architecture Data Access and Integration*

  - *OGSA-DATA design team -*

  - *OREP-WG - OGSA REPlica Working Group*

  - *GSM-WG - Grid Storage Management Working Group*

  - *INFOD-WG - Information Dissemination Working Group*

This shows a very active participation by EGEE in GGF and other international bodies (e.g. EUGridPMA).

Further, JRA3 submitted code to the open source project OpenSSL (http://www.openssl.org/), such that the software stack recognises proxy credentials, in order to authenticate secure SSL communications in a grid context. This is an important step forward and demonstrates how EGEE can contribute to not only standards definitions but also implementations of these standards. In the present case, the maintenance of the contribution to OpenSSL (available from version 0.9.7g

onwards) will be provided by OpenSSL.

Follow-on: Work with Joanne Lawson on this

23. Recommendation: Encourage EGEE project members to attend, chair groups, help write standards, at bodies such as GGF.

    Response: See previous answer.

Note

Doc. Identifier:

**EGEE-NA1-TEC-588116-
AnswerToEuReviewersRecomRevi
ew1-v8.0**

Date: **03/10/2005**

1.6. EUROPEAN GRID OPERATIONS, SUPPORT AND MANAGEMENT (SA1)

24. Recommendation: Present a migration path to further reliance of the infrastructure on established, open standards, not just on individual concrete implementations, in order to guarantee the longevity of the built infrastructure.

    Response: It is clear that for the long term, a production Grid infrastructure has to move towards established and open standards. However, at the moment, those moves towards such standards are coming from the bottom-up, as in the case for example of the SRM interface to storage access, as well as from the top-down, as is the case with the move towards web services. It is important that the production infrastructure participate in both activities, contributing what it learns about what really works in practice to the standards bodies, but also ensuring that it is able to adopt standards as they arise. In order to do this, members of the operations activity are active in GGF in many areas and will continue to do so. In addition we work closely with other international Grid projects with the goal of agreeing understood common interfaces that can later be proposed to bodies such as GGF. It is also important to ensure that a production infrastructure should try and maintain a level of performance and reliability, which may mean that new standards-based services cannot immediately be put into production without consideration and understanding of migration strategies from the previous services, support, performance and reliability issues. However, the strong process of integration, certification, and pre-production testing is the mechanism by which these issues can be addressed.

25. Recommendation: Consider the current gLite as a stepping stone towards a more robust standards-based infrastructure, rather than a final deployment solution. Select additional components for integration and deployment through collaborations with other international middleware R&D initiatives.

    Response: As mentioned above, there are ongoing and embryonic collaborations with other international Grid projects. These include Globus, VDT and the Open Science Grid in the U.S. and with NorduGrid in Europe. The strategy for the deployment of new services is that this should be driven in future by the needs of the applications and considerations from the operations experience. With that basis, components and services should be selected and tested in the integration and certification process. It is clear that in order to be in this situation, the interfaces and functionality of services must be well understood and defined, and the standards process must help here.

    Follow-on: Mention how EGEE-II will take s/w from other sources (SA3)

26. Recommendation: Clarify which components of gLite as well as which established operational procedures are being contributed to the standards activities of international organizations such as GGF, IETF, OASIS, etc.

    Response:

    The SRM (Storage Resource Manager) interface was brought to GGF as a potential standard after it had been demonstrated as useful as a common interface at several Grid sites.

    From the operations activity point of view, the security groups and teams are extremely active within GGF. In addition the deliverable DSA1.5 and DSA1.7 (EGEE Planning Guide - "cookbook"), which is intended as planning guides to assist new participants join or build components of the infrastructure, including resource centres and their administrators, operational organisations (ROCs, CICs, and VOs). It will provide templates and checklists to help to design such facilities, understand what resources are required, how they should be configured etc., and what considerations should be understood in this process. It should be detailed enough to allow

administrators to understand the limitations of the system and how they can be addressed; for example what services can be run on a single machine, how to configure them, and so on. The document will draw on the experiences and expertise in the OMC, CICs, ROCs, and staff in the resource centres.

It is anticipated that this document will represent a summary of the experience gained in building and operating the EGEE Grid infrastructure, and include case studies of why certain decisions were made, contrasting with alternatives in other projects as far as possible.

The document will be contributed also to GGF (Production Grids Management RG) as a case study.

Follow-on: JRA1

27. Recommendation: Continue to conduct application-driven investigation that may result in complex usage scenarios and consider how the advanced middleware and infrastructure would support them in a viable manner. As such, keep a keen eye on new generations of production-level Grid middleware from various international groups that go beyond gLite features.

Response: The High Energy Physics community through the LCG project is planning extremely complex and demanding service and data challenges during the next two years. In the past year, the LCG data challenges were essential in finding the real functional and operational problems in the EGEE infrastructure. This activity will continue and will become increasingly demanding and sophisticated. It is important that as the applications progress, the infrastructure is able to react to the changing requirements through deploying new tools and services as appropriate. These will be drawn from all available sources, as long as the mechanisms for support for those tools is understood and agreed. It would be very advantageous to the project and the long term usability of the Grid infrastructure if other EGEE application communities were also to plan and conduct similar application challenges. This is to be strongly encouraged.

28. Recommendation: The migration path to gLite needs to be better planned, as it is inherently difficult to support two different grid software stacks indefinitely. More specifically, establishing a fixed time-line for migration as well as deprecation deadlines for LCG-2 services, plus possibly identifying who would be the earliest adopters from the application side and the time-line for their possible early commital, would be essential; otherwise, existing users may not be motivated to migrate.

Response: The migration strategy is being planned now in detail. The strategy is one of migration, rather than "big-bang" style changes. The workload management system in gLite has been designed with this strategy in mind. The other new gLite services should be able to coexist with the LCG-2 services. This verification of coexistence is work that is being done on the SA1 certification testbed, and will be further demonstrated in the pre-production service. Once SA1 receives the complete first gLite release and has performed the basic levels of verification, a full migration strategy and plan can be produced. Once that is understood, the idea would be to gradually deploy the gLite services in parallel with the LCG-2 services, first in pre-production, and then in production. This will allow applications to use the existing or new services, and migrate from one to the other in a controlled way. Once the point is reached where no applications use LCG-2 services, those services will be withdrawn. Early adopters of the new services certainly include the LCG community. Other specific applications ready to test new gLite services are being identified and assisted by NA4.

Follow-on: JRA1/SA1, give status of gLite deployment production service and which components will be deployed by summer 2006.

Note

Doc. Identifier:

**EGEE-NA1-TEC-588116-
AnswerToEuReviewersRecomRevi
ew1-v8.0**

Date: **03/10/2005**

1.7. NETWORK RESOURCE PROVISION AND NETWORK SERVICES DEVELOPMENT (SA2 AND JRA4)

29. Recommendation: Identify performance limitations of the networking part and the extent to which performance bottlenecks will be experienced in the network rather than the end systems. Make a preliminary plan to address these performance bottlenecks.

    Response: Two ways to identify performance limitations of the networking part:

    - *To be proactive: Collect user requirements, compare them with the capabilities from the network providers, find the network architecture in a partnership with the network providers. This means that the network requirements must be well defined, mainly in terms of bandwidth, but it is not generally the case today.*
    - *To be reactive: Grid operations monitor performance using monitoring tools, if and when bottlenecks become apparent then they will be signalled through a procedure.*

    When a bottleneck is located in a NREN domain, the Technical Network Liaison Committee (EGEE and GEANT/NRENs) will deal with the problem; otherwise when the bottleneck is located in the last mile network, the site will deal with the provider (Regional, Metropolitan, Campus).

30. Recommendation: Clarify the extent to which GEANT/NREN's can be modified to support the projected functionality.

    Response: The network requirements are submitted through the Technical Network Liaison Committee (TNLC) to GEANT/NRENs. From then the network providers can propose solutions and planned deployment, they also can dismiss the requirement if it is not possible to fulfil.

31. Recommendation: Specify expected capabilities of the end system to complement limited network adaptability to provide projected user requirements.

    Response: The network resource provisioning activity does not deal directly with the end systems, this activity can only provide recommendations in case of performance problems due to the end-systems, optimizations or lack of functionalities from the network to the applications/middleware.

32. Recommendation: Address the gaps that exist currently between projected functionality in providing performance guarantees and current limited capabilities in end node co-scheduling and coordination to achieve it.

    Response: We are working on the definition of an architecture for the network service provisioning, it will be one component of the gLite reservation and allocation architecture. When all the components of this global architecture will be defined and available the co-scheduling (co-allocation) could be envisaged.

33. Recommendation: Clarify how NPM required interoperability and interfacing among many different standards and monitoring tools will be ensured and, more importantly, maintained as the infrastructure evolves.

    Response: The JRA4 NPM activity has introduced a standard interface for independent NPM frameworks to expose data through their own implementations. This interface, based on GGF's Network Monitoring Working Group (NM-WG) schema recommendation[1] is described in section 6 of DJRA4.2[2]. The rational behind this decision is that EGEE should not and cannot aim to enforce

---

[1] "Network Measurements Working Group Schema-related work":

http://www-didc.lbl.gov/NMWG/#schema

[2] "Specification of Interfaces for Network Performance Monitoring"

the uptake of a specific NPM framework across the diverse EGEE fabric, but rather provide the mechanism that allows the administrative domains to publish their data to suitably authorised individuals and agents. Figure 1 shows how anyone with the right credentials can access data from various NM-WG-enabled frameworks. Assuming that NPM frameworks of any origin (including home-grown end-site frameworks) and any type (end-site or backbone) are enabled with JRA4's standard interface, end-users and administrators with the correct authorisation can access NPM statistics from a client of their choice or invention that can call the interface.

This interface has been implemented already on the EDG-produced monitoring tools and (by GN2) on GN2's Perfmonit backbone monitoring framework as part of DJRA4.2. Its use was also demonstrated at GGF13 on 16 March 2005 as part of the NM-WG sessions. The access method for these demonstrators (and indeed all of the forthcoming NPM deliverables and milestones) assumes a central access point, also publishing data through an NM-WG interface. This is depicted in Figure 2.
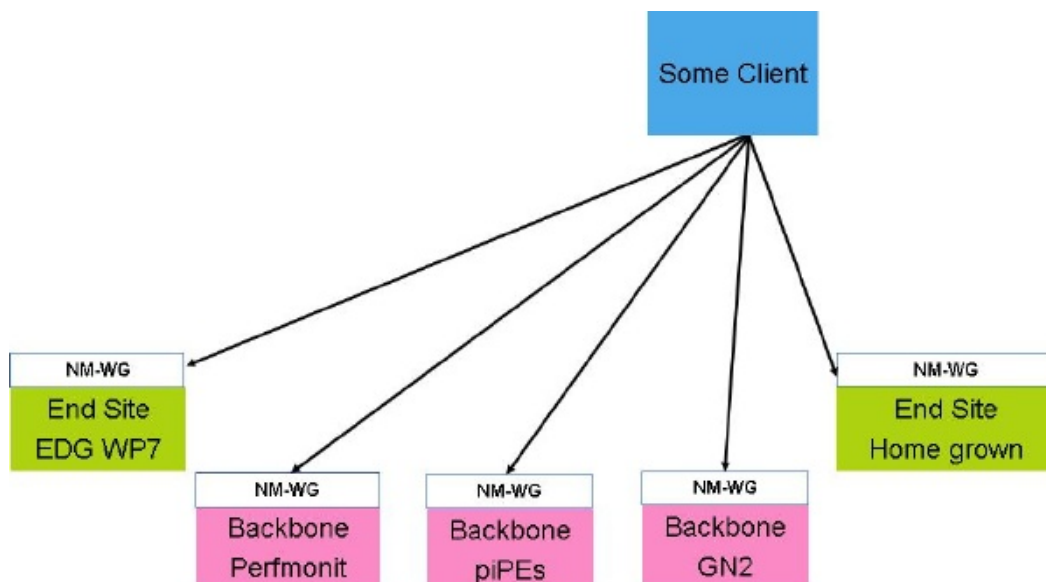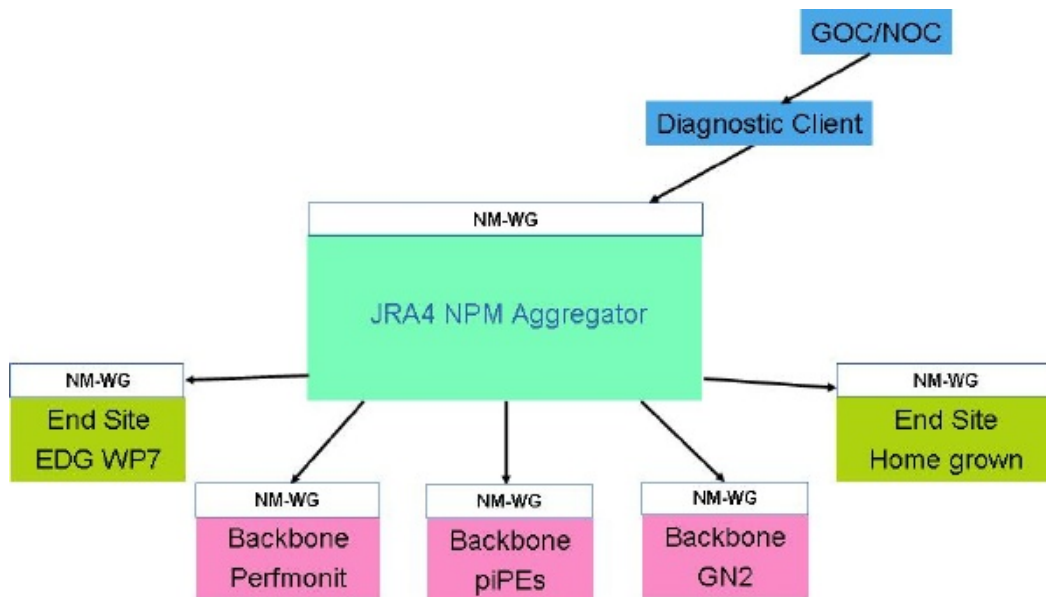

Figure 1: Accessing data from various NM-WG-enabled frameworks

https://edms.cern.ch/document/533215/

Figure 2: Accessing data from various framework through JRA4's single NM-WG interface

34. Recommendation: Even though the sophisticated network provisioning functions are beyond the scope of the current contract, ensure that decisions and actions taken in the project at present will not create conflicts and potentially undermine potential future requirements.

Response: The network service provisioning architecture, as part of the general gLite reservation and allocation architecture, is in the process of defining a common SA2-JRA4-JRA1 work plan.
One of the objectives is to build an open architecture for the future network provisioning functions.

Note

1.8. MIDDLEWARE ENGINEERING AND INTEGRATION (JRA1)

35. Recommendation: Make the migration from LCG-2 to gLite1.0 more application focused. To achieve this, identify exactly which application should be migrated when, who will be the first to use gLite1.0 and what kind of special treatment they are going to receive. The EGEE team should be prepared to provide support resources to encourage and aid in the porting of applications to gLite. If at all possible, announce a timetable for the migration of user applications from LCG-2 to gLite.

The gLite team is working closely with the HEP ARDA group whose task it is to explore interfacing the LHC experiment frameworks with gLite. Similar efforts are ongoing with the Biomed community. A timetable for porting applications will be produced once the gLite software has been released and feedback from its initial usage received and this will be discussed at the project level. Dedicated task forces, which combine experts from applications, operations, and middleware, will be formed to help migrate the applications.

To provide a migration environment and a platform to the applications, SA1 is deploying the gLite middleware on the pre-production service together with the existing LCG-2 middleware. SA1 will support user communities in their use of that service, and together with JRA1 will address issues of functionality and reliability. It is expected that the application user communities will provide effort to assist in this process - in particular to adapt the applications to the new services as needed.

The File Transfer Service is a component used by the LCG Service Challenges, where the Middleware engineering has refocused all efforts to satisfy the immediate needs of this important application. See http://egee-jra1-dm.web.cern.ch/egee-jra1-dm/transfer/index.htm for more details.

36. Recommendation: The current gLite should not be considered to be the final deployment solution, but rather as a stepping stone towards a more robust standards-based Grid infrastructure. Rather, collaborations with other international middleware R&D initiatives are strongly desirable, to select additional components for integration and deployment in mind.

Response: JRA1 will continue involvement in GGF, collaboration with NGS, NorduGrid, OMII, OSG and other initiatives, but this is clearly labour intensive and stretching resources. Efforts on harmonizing interfaces to compute resources have been started including Condor, EGEE, Globus, INFN, LCG, NAREGI, NextGrid, NorduGrid and UniGrids. Common Data Management strategies are being explored with NorduGrid.

The Service Oriented Architecture of gLite is aimed at facilitating the interoperability of different service providers.

Follow-on: Same as 25

37. Recommendation: Presuming that the decision is made that a co-scheduling facility is essential for many aspects of successful Grid computing, a timetable for its implementation should be announced so that both users can work on applications and developers on utilities that require it. The actual implementation should not be considered until the end of the current contract.

Response: Co-scheduling requires the ability to claim (reserve) multiple classes of resource. For the time being reservable resources are not available on the EGEE infrastructure. Efforts are ongoing in JRA4 to provide reservation functionalities for networks. This functionality will be exploited by the planned reservation service based on WS-Agreement.

This will however only be a first step towards co-scheduling but seems the only possibility to advance within the lifetime of EGEE. Co-scheduling is also still an unresolved research problem hence it is too early for publishing a timetable.

Note

1.9. QUALITY ASSURANCE (JRA2)

38. Recommendation: Work within organizations such as GGF to try to establish common QA procedures throughout the Grid community. The EGEE should not underestimate its *puissance* within the community and should use it for the good of all.

Response: The QA procedures put in place for EGEE should be promoted from the beginning for EGEE phase 2 through GGF or other appropriate bodies with the longer term objective of defining a process/service that will deliver "Grid quality label".

An SSA proposal has been recently submitted to the EU on an integration and testing project (ETICS) that has the aim to foster the usage of standard QA mechanisms in software development projects.

Follow-on: Expand on ETICS status

39. Recommendation: Establish a stated QA procedure for external software packages, which guarantees that the same level of QA is provided as for EGEE implemented components.

Response: External packages are divided into two groups: those we can influence (Condor, Globus, dCache) against those which are more widely used (mySQL).

For the first set, dedicated test suites are used during the certification process (currently for dCache). One proposal that should be made to external software packages partners is to agree on mutual QA process. This allowing "in trust" relationship.

For the second set we review the packages, based on pedigree, reputability, large user base, etc.

As a guiding principal, we avoid using the very latest versions to remain more mainstream and safe from the bleeding edge. We are also careful in only upgrading versions when clear benefits are identified (e.g. important upgrades, fixes), while remaining users of well supported versions (i.e. avoiding unsupported packages).

Follow-on: Say SA3 in EGEE-II will expand on this work

40. Recommendation: As much as possible, have members of the EGEE team "embedded" within the development infrastructure of the external providers (the review panel is pleased to see that some of this has already been done).

Response: We are already using this collaboration model successfully with Condor and Globus. Based on this success, we are looking in applying it for other collaborations if and when beneficial.

41. Recommendation: Establish a stated version update procedure that balances the needs of early adopters with production users. The existence of such a procedure will make it easier to respond to urgent requests in a measured and authoritative manner.

Response: The process and mechanisms already in place to balance the needs of early adopters with production users are the following:

- *LCG/ARDA team is actively participating as early adopters*
- *Incremental releases with core and non-critical services/components are released at different rates*
- *The releases are available on pre-production environment before being launched on the production system. Very early access to new developments is provided to application representatives via the JRA1 development testbed.*
- *Further updates will take into account feedback from the applications and operations*

Note

1.10. SECURITY (JRA3)

42. Recommendation: Establish future training courses on non-technical aspects of security. Possible course of action is to establish a dialog with the centre for e-social science in the UK.

    Response: Today's focus on the security aspects in the training are highly technical, and to further improve on the non-technical (that is, organizational and behavioural) security aspects of training, JRA3 will work closely with NA3 User Training and Induction.

    Follow-on: *Need status update*

43. Recommendation: Further investigate the industry security requirements through a synergetic study with the EGEE Industry Forum.

    Response: Together with the EGEE Industry Forum representative, JRA3 will further investigate the industry security requirements.

    Follow-on: *Need status update*

44. Recommendation: Perform stress tests of the security infrastructure by planned, deliberate external attacks conducted by 3rd party contractors.

    Response: The testing of the security infrastructure is divided into two parts, service attacks and the operational side's response to the attacks.

    At this stage, and for the rest of the project, security analysis is our favoured method of testing security, with walk through and exercises for operational side.

    For the middleware, inputs are given from the operational side – e.g. logging done properly. For actual holes in middleware security, code reviews of the critical parts and basic smoke testing are being performed. For example, we now have a set of test certificates, running the set against all the services gives us a valuable input whether the authentication works as it should.

    Follow-on: Mention Vulnerability Group

45. Recommendation: Because of divergence in the national policies within the EU in providing digital signatures for their respective citizens, the EUGridPMA should widely advertise its role and experiences in order to help provide technical grounds for a more standardized treatment of online security across the EU, which would also be compatible with global policies.

    Response: The EUGridPMA in its daily activities addresses the risk of divergence in the national policies within EU. Worldwide, EUGridPMA is a driving example for similar groups in Asian-Pacific and the Americas, and based on EUGridPMA requirements a global federation of Grid identity providers has been formed, the International Grid Federation (IGF).

1.11. FUTURE WORK

46. Recommendation: Testing and software packaging will be critical to success. Reinforce these also intellectually very demanding activities even further.

Response: This is an area where SA1 has invested considerable effort and staffing over the past year building on the LCG efforts. It is essential that this effort be continued and strengthened together with staff from JRA1. Even if the project were not developing new middleware this activity is still crucial to success. Staff from all relevant activities (SA1, NA4, JRA1 and JRA3) and their partners are being actively canvassed to contribute in this area. Since the review, several new staff have joined the testing effort.

Follow-on: Give testing status, including contributions from partners: Russia, Imperial, CE, etc.

47. Recommendation: Work hard on event-based monitoring techniques, triggering preventive maintenance actions, to improve the stability of the Grid infrastructure.

Response: This is part of the operations work that is planned for the remainder of the project. The many different monitoring tools that are currently in use are being integrated into a single top-level view of the monitoring system by the use of R-GMA. This will then allow alarms and actions to be triggered based on this high-level view, potentially with the correlations of different views of a problem from different tools.

Follow-on: Need status update on site certification test process and suite

48. Recommendation: Improve the middleware deployment process (technical, organisational) even further to increase the stability of the infrastructure and consequently improve the job success rate and reduce the load on the support team.

Response: This work is in hand, the deployment process has been reorganised to streamline it and make it more efficient and to prepare it for the gLite deployment. The various levels of services, core services, other services, client tools, etc have been separated to allow different deployment schedules and update policies. This will permit critical fixes, security patches, improvements to client services to be deployed in a more flexible and rapid way, avoiding major upgrades. This in turn will improve stability and thus reliability. A significant effort has been put into improving the software configuration mechanisms for the deployed middleware. Site misconfiguration has been a source of problems and this work to simplify the configuration of a site will go a long way to improving that.

Follow-on: Need update from SA1

49. Recommendation: Implement a strong mechanism to quickly isolate unstable sites in the production Grid.

Response: The information system allows a site to be easily removed from those available to an application. This can be done currently at the level of individual VOs or users. We are working on a mechanism to remove bad sites from overall visibility. This has to be done in conjunction with a well defined and objective process based on defined tests before sites are marked as unstable or bad.

Follow-on: Need update from SA1

50. Recommendation: Choose new application areas with greatest care: capture user and application requirements in detail and determine the impact on Grid infrastructure and services. The decision to deploy any new application (and if yes, when) should be based on a very careful analysis of all associated technical and non-technical risks.

Response: This recommendation is addressed by requiring new applications to provide information needed for a Memorandum of Understanding with EGEE before they are evaluated by EGEE Generic Application Advisory Panel. These MoUs will contain key information identifying the key requirements and level of commitment between the applications and EGEE.

Follow-on: No new EGAAP call (text from EAC note)

51. Recommendation: The ambition to quasi-fulfil industrial or commercial requirements should be relaxed until the end of the current contract.

    Response: Through the Industry Forum, we are keeping contact with the future needs of the industry. However, the main focus for the project is to deliver the best possible Grid service for the benefit of e-Science communities.

52. Recommendation: If examples can be found of scientific work that could not have been done without the EGEE development (e.g., necessarily distributed process, truly enormous calculations), they will be a strong motivation for future development.

    Response: All applications demonstrated the added value of the Grid. Some of the science behind the applications deployed today on EGEE could have been achieved without a Grid. The level of benefit often depends on the type of usage the application makes of the Grid services.

    The advantage of the Grid for applications is not limited to access to large computing power. Access to largely distribution and federated data is another key feature of the Grid and can be leveraged by Grid applications. This is especially true for the biomed community, which often is composed of heterogeneous data distributed across a wide range of locations. In this case, uniform access to this type of data is a unique feature the Grid can offer. In addition, the EGEE computing infrastructure provides access to large computing resources in areas where otherwise would not have been available. For example, contrarily to HEP or industrial partners, biomedical applications often cannot rely on established computing centres and large system engineering teams for maintenance of computing facilities.

    Deployment of biomedical and generic applications is foreseen at a larger scale in the next months and significant scientific results are expected to come out of this exploitation. NA4 is also looking for opportunities to deploy "data challenges" similar to High Energy Physics to show the full capabilities of the EGEE Grid Production Service beyond HEP.

    Follow-on: Mention regular application press releases

53. Recommendation: The concertation effort among the various grid activities that was undertaken by EGEE needs to be formalised in order to be effective. A deliverable reporting on the progress and results of the concertation effort will facilitate that task. It is suggested that such a deliverable be prepared for the next review, close to the end of year 2005.

    We propose documenting the concertation activities as part of the deliverable DNA5.4 "European Grid project synergy report (in collaboration with DEISA, SEE-GRID and other relevant initiatives and EU projects)".

    Since the DNA5.4 is an update of the document DNA5.2, which covered a short report on concertation activity, this would be more natural than the creation of a new deliverable.

    Follow-on: Need update from NA5