Input for EGEE-2

Bob Jones
Revised November 12, 2004

Introduction

This document is a summary of my thoughts about the possible content and structure of a follow-on project to EGEE (referred to as EGEE-2 in this document) and is intended as an aid to the planning work of the PMB working group formed to study the subject.

These notes are based on my observations of the EGEE project from its preparation, through the first 7 months of existence and interpretation of EU commission future priorities that could be gathered from the EU commissions presentations at the concertation meeting held on September 16-17th 2004 in Brussels. The document is organized into areas that roughly follow the activity definitions for the current EGEE project.

Organisation

Federations

The federated structure of EGEE has proved successful but lacks an official recognition within the project with respect to the EU. A legal mechanism for ensuring that this structure is recognized is needed for EGEE-2. This is especially important if EGEE-2 is to encompass more regions of the globe.

Empowering the group of federation administrators (i.e. the AFM in EGEE) and giving them specific responsibilities would ensure the distributed structure is entrenched in the project and would reduce the administrative load currently supported by the PO. Each country should organise its partners internally and the project would deal with a national representative who should appoint a dedicated administrator. This approach is consistent with the goal of EGEE to build on national initiatives and clarifies the interface between the project and such national grid projects.

Activity presence

There are a number of activities that need a presence across all the participating countries: grid operations (SA1), computer network support (SA2), dissemination (NA2), training (NA3) and application support (NA4). In EGEE there are gaps in this coverage which means that the project is less effective in supporting applications and disseminating information than it could be.

Partner classifications

Partner membership of the project should be more clearly organised into categories to minimize the amendments required to the contract as new organisations become involved. Organisations which want to contribute computing resources to the services (production, pre-production) or testbeds and those that wish to bring in applications should not need to sign contracts or be funded directly within the project.

External Advisory Committee (EAC)

The EAC is a small group of external advisors that review the progress of the project and give recommendations on its strategy. The independence and existence of this group is important but has not proved as useful as I would have hoped mainly because the members are very busy and do not have sufficient time to spend on the project. Their intervention is currently limited to attending the project conferences and reading the overall quarterly reports. We still miss one or more individuals that could review the deliverables produced by the project, after they have been reviewed internally and before they are sent to the EU. In DataGrid this task was performed by Mark Parsons but we have been unsuccessful in finding someone to fulfill the role in EGEE.

Deliverables and Milestones

The definition of deliverables and milestones for the project is a contractual obligation and a means of measuring the project's progress. In EGEE the deliverables and milestones were defined to be useful within the project (i.e. we did not want to simply produce mountains of paper) but, during the project preparation phase the list grew and we became aware of more EU imposed deliverables. For EGEE-2 it is important to minimize the number of deliverables to only those which are required by the EU and essential to the running of the project since the production and review of these documents is a very labour intensive activity. A few lessons learnt about milestones and deliverables are:

- The definition of deliverables and milestones at the proposal stage must include an abstract describing their purpose and scope (see JRA1 table BB in the EGEE Technical Annex as an example).
- The activities assigned to review deliverables and milestones should be defined at the project proposal stage.
- When possible, milestones should be used instead of deliverables because their reporting requires less effort (i.e. the internal review process for milestones is lighter than for deliverables and they only need to be mentioned in the quarterly reports rather than sent to Brussels individually).
- Synchronization between milestones and deliverables in different activities must be maintained during the project preparation stage to avoid inconsistencies.
- No more than one milestone or deliverable should be produced by any activity during each quarter.
- Deliverables should not be scheduled on the same months as quarterly and periodic reports (i.e. avoid months 3, 6, 9, 12 etc.).
- Deliverables should not be scheduled during the July and August holiday season.

Reporting Period

The reporting period for EGEE is 9 months. This was favoured over the more usual 12 month period because of cost-flow concerns: given the level of advanced payments permitted by the EU, a 12 month reporting period would have left the partners in the red towards the end of the period. While financially this is prudent, from an organizational point of view it is very awkward since EGEE will have 2 reviews during the 2005 calendar year and a final review 6 months later. For EGEE-2, pressure should be put on the EU to ensure a 12 month reporting period can be adequately financed given that the consortium will already have a track record of respecting its commitments.

Proposal Preparation

Programme of work

For EGEE, a Technical Advisory Board (TAB) was set-up that produced a report which served as the basis for the project proposal. The TAB was given the following mandate:

"The main role of the Technical Advisory Board is to advise to the EGEE Executive Committee on a project structure which is technically sound and in accordance with all requirements coming from the European Commission (call for proposal, work programme, etc...) This project structure shall consist of a series of well identified, self contained workpackages and shall include if necessary, transverse structures.
The workpackage contents and boundaries should be such that small task forces can be launched to write the corresponding section of the proposal, in the required timeframe. It is expected that the following topics will be covered by the Technical Advisory Board: -Grid Middleware re-engineering -Grid production support and operation -Interface with applications.
Non-technical Work packages such as project management, industry relationship, dissemination, relationship with other projects and networks of excellence will be handled directly by the EC.
The Technical Advisory Board can solicit the help of outside experts to accomplish this task. The role of the TAB during a second phase will be to monitor the progress of the various task forces and to guarantee the technical coherence of the whole project proposal. Task Forces will be appointed by the Executive Committee. The TAB will be formally dissolved on submission of the project proposal."

The TAB held an intensive workshop that resulted in the document. I think the TAB was successful because it was well-focussed and short-lived. I am less convinced about the effectiveness of working groups that then took the TAB report and expanded on it since they were less focused and more loosely managed.

If the EGEE-2 task force is to produce a document that is the equivalent of that produced by the TAB then I suggest it is sent to the existing EGEE PEB and their feedback taken into account before being distributed more widely.

Budget

The budget negotiation was performed based on a premise that a person working full-time for one year would cost 100K euros. This was a very useful approach to get the discussions going and provide ball-park figures but proved inadequate in preparing the final budget for the project. Similarly, the effort required for specific tasks got adjusted to maintain desired budget levels for individual partners which is the reverse

of what is required. The role of the federations in this process was not completely clear.
Having discussed this point with the EGEE financial officer, Severine Bergerot, some changes on the EGEE-2 budget preparation will help the consortium converge more quickly. The EGEE-2 consortium membership is likely to be very similar to that of EGEE and the majority of the current staff will be retained so we can use the EGEE financial information to provide more accurate estimates. If we reinforce the role of the federations (as proposed in this document), it makes more sense to take the following approach:

1. Calculate the initial effort estimates for tasks defined in the work programme in terms of total FTEs and assign them to federations (not individual partners). This step could be completed by the EGEE-2 task force.
2. Federations calculate costs for their assignments taking into account the average costs within their federations (this average can be estimated taking into account the costs reported at the end of the first period of EGEE – i.e. December 2004). In general, such averages are quite close to true costs in all federations with the exception of CE and Networks where there is a large disparity between partners. This step will provide an overall ball-park budget for the project.
3. Once the overall ball-park budget and federation allocations are agreed, the federations can make allocations to individual partners respecting their total envelope. For this step the average FTE cost should be replaced with the costs reported by the individual partners in EGEE.

Relationship to LCG

The strategic choice of relying on LCG to provide the operational has proved very successful because it has enabled the project to deliver on the promise of "hitting the ground running" by providing a production service from day 1.
The feedback about deployment issues coming from LCG is helping to shape the new middleware, called gLite.
Having LCG as the interface with the physics experiments has also been successful but there are still some issues that affect the perception of EGEE related to the interactions with LCG:

- Having LCG as an interface to the LHC experiments has meant that the experiments are unsure about the true role and benefit of EGEE which means it is harder to get their support for EGEE-2.
- The parallel but separate structures of EGEE and LCG means there are overlaps (e.g. similar but unrelated boards, committees, groups; multiple reporting lines, deliverables and milestones) which leads to extra work and clashes in scheduling of people's time.
- More effort must be dedicated to ensuring that EGEE decisions are not perceived as being purely driven by the needs of the LHC experiments.
- The majority of individuals working in areas that are of benefit to both projects still see themselves as a member of one project or the other and this encourages a "them v. us" mentality. Certainly, the situation is less severe than existed with LCG and DataGrid, but it is still a brake on interaction and progress. This is perhaps more of an issue for CERN than other partners.

LCG is currently organising its second phase and hence is in a similar situation to EGEE. LCG is counting on EGEE-2 to complement its funding. My understanding is that, with the exception of the US and Japanese partners, all the LCG partners are also partners of EGEE. If EGEE-2 could secure funding for US partners involved in middleware development (i.e. Condor and possibly Globus) and US & Japan grid service operators (i.e. principle grid sites working for the LHC experiments) this could diffuse the sometimes conflictual relationship between the EU and US grid projects that hampers LHC experiments' attempts to promote the adoption of common solutions.

LCG will be writing its Technical Design Report in the same period that it will be necessary to write the Technical Annex (description of work) for EGEE-2. These documents should be complementary and synchronised to ensure that there are no discrepancies or divergences between the projects. The management structures should take into account the existence of the two projects to minimise duplication and simplify/combine reporting lines.

Grid Operations

I believe the underlying principle of EGEE, that no money is foreseen in the budget for buying hardware, has helped ensure the project is open to new-comers. This principle must be maintained in EGEE-2.

Even if EGEE achieves its goal of producing gLite middleware that is more reliable and simpler to install and manage than LCG-2, I still expect grid operations to continue to require the largest proportion of EGEE-2's budget.

The transition from the LCG tier based organisation to the EGEE OMC/CIC/ROC/RC structure is on-going but the exact roles, responsibilities and relationships still needs to be clarified. The effectiveness of this hierarchy should be reviewed after the first EU review, but its establishment in EGEE and the attribution of the funds for this structure was a contentious issue during the project negotiation stage and continues to impact the work performed.

If this structure is maintained for EGEE-2 then the operations funding should be organised so that each federation has the equivalent of a CIC and ROC.

The issue of Quality of Service (QoS) will be a more important aspect of the grid operation in EGEE-2. The specification of Service Level Agreements (SLAs) to address QoS are a deliverable of EGEE but it is unlikely that such specifications will be implemented during the project's lifetime. SLAs will gain in importance as the eInfrastructure grows and supports more applications (notably business applications). The implementation of SLAs is thus likely to be an important task of the grid operations in EGEE-2.

To support this work, better monitoring tools will be required and it is expected that these can be produced based on the experience gathered with the current grid service. Billing and accounting infrastructure will also need to be further developed.

Application Support

Application support has a direct impact on the usefulness of the project to end-users and its appreciation by the EU commission.

The identification of the 2 pilot application domains, HEP and BioMedical (BIO), has proved successful and their role should be continued.
EGEE's NA4 activity has made a good start in this area but I would advocate a significantly increased budget for this activity in EGEE-2. From experience, we see that dedicated task forces are an effective mechanism for ensuring applications and user communities can successfully exploit the grid infrastructure. This is effectively what has happened with NA4 manpower involved in the ARDA project and BIO applications.
The proposed increase in manpower for application support should go primarily to the BIO and generic applications since the involvement of the HEP community is so strong that they are already successfully making use of grids and such an action would help to reverse the perception that EGEE is only addressing HEP needs.
The differences between the BIO and HEP communities are now well known (BIO is less well structured and has less experience at working together across organisational boundaries) and it is proving difficult to ensure EGEE has a significant impact on the BIO community.
The BIO involvement in the project is still lacking a flag-ship application (i.e. the equivalent of the LHC for HEP that rallies together physicists across the world for a common goal). We should try to identify such a BIO flag-ship project or organisation and use this as the principle interaction point with the community for EGEE-2. There are a number of initiatives that now exist (i.e. HealthGrid http://www.healthgrid.org/ EMBRACE, INFOBIOMED http://www.infobiomed.org/, eHealth http://www.ehealthinitiative.org/ etc.) and we should evaluate their role and performance in the BIO community and enter into partnership with the leading group.
The relationship to the European Molecular Biology Lab (http://www.embl.org/) and European Bioinformatics Institute (http://www.**ebi**.ac.uk), which are rich and powerful BIO organisations, and their potential role in the project needs to be understood. The principle contact here is Peter Rice (he is leading a group that investigates potential use of eScience & grids and has already visited CERN). EMBL-EBI provides funding for myGrid (http://www.mygrid.org.uk/) which is a well-known BIO project in the UK who we are trying to attract to EGEE infrastructure.

GILDA
The relationship of GILDA to the production service needs to be clarified (i.e. running exactly the same version of the middleware found on the production service and making sure it acts as a stepping stone to the production service and not an alternative) but its establishment has proved very useful and is an excellent mechanism for introducing new user communities. Its existence and support should be formalized in EGEE-2.

Middleware Development and Support
The situation concerning middleware is evolving and the EU Commission documents separate it into two areas: Grid Foundations Middleware (otherwise known as "Underware") and Grid Services Middleware (otherwise known as "Tupperware".)
EGEE is essentially involved in the support and re-engineering of Grid Foundations Middleware : low-level middleware which provides the basis for an application independent eInfrastructure.

Grid Foundations Middleware re-engineering should remain limited to a small number of partners since I can see definite improvements in this activity of EGEE compared to the performance of the many separate work-packages in DataGrid. Inefficiencies are still present in some clusters because their participants are distributed across several sites which complicates the management and coordination task.
The current approach, starting with a new code (i.e. different from LCG-2) is a high-risk choice not made by JRA1 but imposed on it by outside pressures coming from some of the LHC experiments and related bodies. There is more new development involved in this choice than originally foreseen under "middleware re-engineering". This decision was made at a time when experience with LCG-1 was limited. Since then more experience with LCG-1 and subsequently LCG-2 has been gained and its reliability/stability has significantly improved and it is now widely accepted by the LHC experiments. Most of the problems with the current production service are not related to the middleware (the majority of problems observed come from site related issues), so EGEE-2 should be less ambitious and less revolutionary in middleware developments.
EGEE is now in the position of having improvements being made to LCG-2 for the current production service by SA1 while developing new gLite software in parallel by JRA1. While on-paper this duality appears efficient, in reality it makes migration from LCG-2 to gLite very difficult since they are significantly different software stacks. Having different teams creating/modifying middleware in parallel has lead to a competitive environment with some degree of mistrust which the activity managers are finding difficult to diffuse. Also, improvements for the end-users are appearing more slowly because the effort in the two activities is not fully coordinated since they are addressing different priorities (SA1 fix short-term limitations & bugs in LCG-2 while JRA1 have to tackle the more long-term architecture related points).
Insisting that in EGEE-2 any further middleware developments or re-engineering must start from the code base that will be running on the production service at the end of EGEE and define the highest priority for the developers to be that of addressing problems found with the deployed production software will help ensure EGEE-2 does not suffer the same conflicts.

Inter-operability
The LHC experiments are currently using multiple grid infrastructures (e.g. LCG/EGEE, Grid3 and NorduGrid) and the situation is unlikely to change in the near future. These grid infrastructures do not inter-operate and this obliges the users to develop software at the application level to be able to exploit the resources available via each infrastructure. Manpower and milestones should be foreseen in EGEE-2 to specifically address inter-operability issues with these sister grid projects. Agreements should be formalized (e.g. something like a MoU) with these alternative grid projects in which the other projects also commit manpower to ensure inter-operability is achieved.

Grid Foundations and Services

In terms of relative importance of Grid Foundations Middleware re-engineering in EGEE-2, I currently believe it should be continued at a similar level but this needs to be reviewed based on the results of the large-scale deployment of gLite during the second year of EGEE.

Strategically, EGEE-2 should be concerned with Grid Foundations Middleware and not try to develop Grid Services Middleware since we would be in direct competition with a myriad of other projects. Rather, the project should provide an environment which can accommodate instances of Grid Services Middleware coming from other sources. This will ensure the project is seen as being open to many application domains. As an example, consider how the different application domains approach work-flow. HEP considers tools such as Condor's DAGMAN as sufficient for its needs while BIO and business applications require more sophisticated tooling (e.g. Taverna visual tool along with the Scufl language and an associated execution engine). Work-flow and access to disparate data sources in varying formats are two areas that are important to many application domains but there is as yet little agreement on a common solution (though OGSA-DAI is gathering importance as it matures).

Standards

While we must continue to converge on emerging standards to facilitate inter-operability and eventual technology transfer to industry, my belief is that the WS-RF (or whatever it evolves into as a standard for grid middleware) will not be a fully-defined, stable international standard before the end of EGEE. This means that effort will be required to ensure gLite tracks this standardisation effort in EGEE-2. This convergence on standards is an iterative process that has started with EGEE but will surely need to continue.

Middleware Repository

The concept of an international middleware repository, similar in function to the UK's OMII project, is also gaining ground and is seen as a possible task for EGEE-2.

I see a potential conflict of interest in this area for EGEE in a similar manner to the criticisms that have been raised against OMII (i.e. that the project favours the software developed by its own partners.) To avoid this situation one possibility would be to have the repository follow the definition of middleware used above: EGEE would provide a common set of application independent Grid Foundations Middleware (i.e. gLite) on which a host of Grid Services Middleware instantiations could be accommodated. EGEE would hopefully thus been seen as impartial with respect to the Grid Services Middleware and would need to define the software engineering process to evaluate the quality of the services contributed. A golden rule necessary to preserve this impartiality would be that EGEE does not produce any Grid Services Middleware itself and must publish the criteria (and associated testsuite) that will be used to assess software contributions.

Competition with OMII itself would be an issue but collaboration on a joint repository could be possible in the future (assuming EGEE and OMII could agree on a common set of Grid Foundations Middleware) given the similar timeframe's of the projects.

The EU commission has stated that it is interested in funding projects to produce a network centric grid operating system (GOS) in the FP7 programme. Given the joint CERN/Fermilab work to move from RedHat to the Scientific Linux distribution, if

EGEE-2 proposed to bundle elements of the Grid Foundations Middleware (i.e. gLite) in this distribution it could be a way of funding this on-going work, ensuring we are aligned with EU wishes and getting a common basis with our US colleagues.

Fabric Management

It was decided to exclude fabric management from the EGEE project. This decision was based on the experience with EDG and early LCG work which showed that it was impossible to find a single tool that was suitable for all participating grid sites. This decision had the advantage of ensuring that the middleware could be installed manually and so be independent of any particular fabric management tool. However, we now see that the Quattor tool is proving popular amongst a number of grid sites and it has been reported that the most successful (i.e. those up and running as quickly as possible with the least effort) sites for LCG were those using LCFGng (the predecessor to Quattor). The LCG GDB is now proposing to create a group that would work with Quattor to fill the current support vacuum but it is unclear who will commit the necessary manpower and gLite configuration definitions for this tool.

The EU Commission has stated that it is looking to fund system management and coordination aspects of grid nodes in FP7 and would promote the development of network-centric grid operating systems (GOS). The characteristics of such a GOS include many of the aspects we refer to under the subject of fabric management.

It is beyond the scope of our programme of work to embark on the development of operating systems, but including support for a fabric management activity within EGEE-2 while ensuring the tool independence of the middleware would simplify the inclusion of new computing resources in the grid infrastructure and provide a funding line for work that is requested from the community.

Bundling aspects of Quattor as part of the Scientific Linux distribution (our simplified vision of what the EU commission calls a GOS) as an optional aid to cluster management (see **Middleware Development and Support** above) would help provide funding for our license-free Linux distribution work and fabric management efforts while being inline with the EU Commission’s priorities.

Training and Dissemination

Training has started well in EGEE with numerous events (including the GGF grid school) organised across many countries. There is good technical knowledge in the training team and the feedback from event attendees has been positive.

Dissemination output will need to be increased and its targets better defined. More planning will be necessary before the EGEE-2 project starts to better define the dissemination programme and to make strategic decisions. We really need a list of events defined before the project starts to ensure news about the project can be disseminated to user-communities and computing resource providers at regular intervals. Ideally we should also be in a position to make use of dissemination as an aid to ensuring that the project proposal is successful (I'm thinking here of how candidate cities promote their cause when competing to host the Olympics.)

More material must be available earlier in the project and the dissemination groups must make more of an impact at user-community events (i.e. not just grid events to which we get invited). To do this the dissemination activity must be funded to organise stands etc. and be capable of running its own demos and be familiar with the grid technology that the project is offering and what are the benefits for the end user.

It is my belief that the division of dissemination and training into separate activities in EGEE has reduced the overall effectiveness of the two teams. I would recommend that the activities be combined in EGEE-2 to ensure the team as a whole has the necessary skill-set, budget and contacts to achieve maximum impact.

Similarly, dissemination and training is intimately linked with application support since similar skills are required and the user-communities are heavily involved.

The EGEE conferences have proved more expensive than originally foreseen. The decision to link them to EU presidency looked sensible at the outset but has added to the cost and difficulties of organization. The suggestion is that in the future the conferences are separated from the presidency and only the eIRG meetings linked to the presidency.

Security
Security will become more important in EGEE-2 due to the expanding nature of the infrastructure and inclusion of more applications with more severe security requirements.
Standardisation work is on-going and I expect that, while agreement may be reached during the lifetime of EGEE, their successful implementation will be a task of EGEE-2. In particular, the support for WS-Security, WS-Privacy and WS-Policy will need to be addressed.

Computer Networking Support and Research
In EGEE the most important aspect of this work is the link to the GEANT project. The funding level has been set to ensure this link is maintained and to be able to track the developments related to IPv6. This aspect of the programme of work is politically important but is not on the critical path for the project itself. A similar level of funding should bee foreseen in EGEE-2.
Similarly to computing resources, EGEE-2 should not request funding to provide connectivity for partners or contributing sites.

International Collaboration
International collaboration was originally intended to be lead by the activity NA5 in EGEE. In reality, the technical programme of work has been organised so that international (i.e. outside the EU) collaboration has taken place within a number of activities (mainly middleware re-engineering and grid operations).

This has left NA5 with the role of supporting the eInfrastructures Reflection Group (eIRG) and developing a synergy roadmap with the DEISA (i.e. supercomputing) and SEE-GRID (i.e. South East Europe extension of EGEE) EU funded projects.
These are really collaboration tasks with related EU projects. Clearly in EGEE-2, such "concertation" with other EU projects will be required but should be better defined. Assuming the model for a middleware repository (outlined in **Middleware Development and Support** above) is adopted, this collaboration work could be the investigation of the suitability of inclusion in the repository of the higher-level services resulting from the work of EU projects. This task could be simplified by ensuring the projects start by using the EGEE infrastructure from the beginning as is the case now with projects such as GRACE and DILIGENT.