

## Moving towards EGI – EGEE-III Year 2

Steven Newhouse 3/3/09

### Executive Summary

The EGI\_DS Blueprint, endorsed by the EGI policy board in January 2009, proposed significant changes in the provisioning of e-Infrastructure within Europe. The EGEE-III project provided feedback (<https://edms.cern.ch/document/971629/3>) to an earlier version of the EGI\_DS Blueprint (<http://www.eu-egi.eu/blueprint.pdf>) and has a commitment within its own work programme to review how it can transition to the proposed EGI model. This will require changes to the EGEE-III description of work for the second year of the project described at a high-level within this document. These changes will primarily alter the existing technical, communication and management structures to produce a better alignment to the functions those being proposed for EGI (<http://www.eu-egi.eu/functions.pdf>). This document will be used as a basis for input into the EGI\_DS deliverable D5.5 which proposes a transition plan for the broader e-Infrastructure community in Europe to the model proposed by the EGI\_DS Blueprint, and changes to EGEE's own description of work which will be approved by April 2009.

Input into this document has come from discussion with the activity managers and their contributions into earlier EGI discussions and the EGEE All Activity meeting in Brussels in January 2009. Further review has taken place within the TMB and AMB.

The most significant risk to these changes is the resistance from the gLite development community to working as a single coherent organisation – it is therefore essential that the project (through the PMB) provides its **full** support for these changes within its own organisations.

### Overview of Changes for EGI Transition

**Software Providers:** Within EGI the bulk of the effort currently undertaken within JRA1 and SA3 will evolve into work done by a software provider (mainly the gLite Consortium within EGEE) and the coordination work done by the Middleware Unit (MU) within EGI.org. Certification and integration becomes an activity done by the software provider with software packages delivered to the MU only for verification. At the start of EGEE-III a move was made to start devolving testing effort to being nearer the development teams ('clusters of competence'). For the remainder of EGEE-III we will continue this devolution of SA3 effort to include certification activities – building 'product teams' across the main functional areas and component sets within gLite. Each product team becomes **completely** responsible for delivering production quality software capable of being deployed and run on the production infrastructure with no further intervention. The remaining SA3 effort will be used to perform the relevant work of the MU.

This will involve some significant process changes, including:

- A review of the patch process (moving to scheduled releases for non-critical patches according to an agreed roadmap with the product teams) would make best use of resources.
- A build system will be used across all development teams to manage this release process.
- The physical resources needed to support the large-scale testing and certification activities used by these product teams will be contributed by exiting resources (from within the SA3 and the PPS) supplemented by resources contributed by NGIs as described in the EGI Blueprint.

- The product teams will be explicitly linked up with their early adopting customers (sites and VOs) earlier in the process ensuring developers are aware of the eventual deployment constraints and operational requirements, and that the customers can have access to early prototypes demonstrating new functionality so that they can provide feedback. This brings back elements of the 'Experimental Service' used in earlier phases of EGEE but still allows for community building around certified functionality using 'Pilot Services' (see below).

**Pre-Production Service (PPS):** Currently, software releases are given a deployment test on the PPS by a distributed team under SA1. This activity is mainly used to verify the deployment of the software across environments typical of those found in production but undertaken on non-production resources. With the improvements made in the certification process this pre-release activity is no longer seen as providing significant benefits in terms of user support. In addition, once released to production many regions undertake their own deployment tests before wide scale release by running the software on production sites. It is proposed that these two stages be merged into one by having a formal 'deployment testbed' composed of representative sites (e.g. different batch systems) from the regions that undertake to deploy new certified software release in a timely manner. This would be coordinated from the centre (the EGI\_DS Blueprint has provision for this) but would effectively provide a grid wide 'rollout' testbed – which could be integrated into existing NGI rollout infrastructures as desired. The 'Pilot Service' (putting major new certificated functionalities into production use with early adopter communities) will continue as required when needed using resources contributed by the interested communities. It is envisaged that communities and sites currently interested in supporting 'Pilot' activities will also be interested in getting engaged in earlier phases of a product's development through 'Experimental' services.

**Training Infrastructure:** Currently the project's t-Infrastructure is partly decoupled from the production infrastructure. Within EGI there is no model for a singular dedicated European t-Infrastructure – although an NGI may choose to fulfil their training capability through either a dedicated or a shared resource. It is important therefore to provide coordination for the provision of resources that support the particular requirements of training and this will be done by coordinating and managing a training VO on the Production infrastructure, and supporting the development of a training SSC to encourage and coordinate national provision. Effectively a site's decision to use a particular resource for training becomes a decision to enable (or not) a training VO onto that resource into which trainees can be assigned. The training 'VO' model will be adopted during year II with all current EGEE training resources being integrated into the same operational model as the production service. The existing t-Infrastructure resources will be obligated to minimally enable the training VO and support the training CA while sitting within the production infrastructure. Some sites which support both training and production infrastructures may take this opportunity to provide a single integrated infrastructure while still being able to provide dedicated resources when needed for training events.

**EGI Networking Support Centre (ENSC):** Establish and operate the ENSC by migrating from the current ENOC by documenting the operational tasks that can take place through the centre and the contributions that need to come from NGIs to operate this centre. More generally, establish the contribution that needs to come from NGIs in order to operate the network activity. Make available any software code related to the networking function in a simple package with sufficient documentation and installation support. Rotate the NGI ENSC amongst interested NGIs (e.g. current SA2 participants) to verify operational procedures and build experience.

**Specialised Support Centres (SSC):** Migrate the interfaces and communication structures currently within NA4 into the EGI model. Each application SSC (current NA4 application domain, support functions and NA3 training support) establishes a co-ordinating contact point, support contact point, middleware contact point & operations contact point. These may be different people or a single person depending on the workload contained within these areas by the SSC. These representatives of the

'new' SSCs will form the interfaces into other emerging EGI bodies. The SSC co-ordinators would form a User Forum Steering Committee (UFSC) chaired by the current NA4 activity leader – effectively identical to the current NA4 Steering Committee but with the representation from NA3 training. Other SSC contact points would interface into the governing structures of the external middleware providers and EGI.org's internal management bodies for middleware, user community and operations units.

**Establish at EGEE oriented MCB:** The TMB renames itself (again) and becomes an EGEE only MCB (Middleware Coordination Board) with representatives from the current operations (SA1), middleware (gLite Open Consortium – JRA1 and certification SA3) and user community (NA4). The new TMB will remain to deal with management issues that will exist within EGEE but not within EGI. One of the new management responsibilities of this body will be to define procedures that monitors the contractual obligations of the software providers.

### **Overview of Critical Changes still needed for EGI Transition**

**Metrics Automation:** The support for automatically extracting information from the operations tools to provide metrics relating to individual NGIs is not progressing at a rate to be in place by the start of EGI. This has found to vital in undertaking effective country reviews. It is critical that a NGI 'dashboard' be in place soon and will require additional focus from the operations teams.

**Operations Tools:** The tools used by the operations team have generally been developed by *ad hoc* teams of developers working within the ROCs. Before the end of the project we need to ensure that these tools are exportable from their current environments and are shown to be deployable and work by other operations teams.

## Detailed Changes - Introducing the Software Provider

Within EGEE-III the software maintenance activity is structured around six areas which include a number of different institutions. These groupings currently form the basis for 'product groups', i.e. the logical groupings where software needs to be delivered as an integrated, certified, working and deployable unit. These groupings will be reviewed (after UF) to see if they deliver appropriate consistency as product groups. Also the maintenance effort currently being dedicated to a component will be reviewed against current requirements to verify that our limited maintenance effort is being effectively used. This may result in a decision that some components enter a strict critical 'bug' fixing only mode allowing resources to be used on higher priority components. This will not affect development effort for functionality which is still seen as having a clear deployment need.

The initial product groupings are:

- Security Infrastructure (???)
  - Authz Service (SWITCH, HIP, CESNET, NIKHEF)
  - VOMS (INFN)
  - VOMSAdmin (INFN)
  - Proxy and attribute certificate renewal (CESNET)
  - Shibboleth interoperability (SWITCH)
  - LCAS/LCMAPS (NIKHEF)
  - glExec (NIKHEF)
  - Delegation Framework (CERN, HIP, STFC)
  - Gsoap-plugin (CESNET)
  - Trustmanager (HIP)
  - Util-Java (HIP)
  - Gridsite (STFC)
- Information Systems (???)
  - BDII (CERN)
  - GLUE Schema (CERN)
- Compute Element (???)
  - CREAM (INFN)
  - CEMon (INFN)
  - LCG-CE (Legacy support)
- Job Management (???)
  - L&B (CESSNET)
  - WMS (INFN, ED)
- Data Management (CERN - ???)
  - CGSI\_gSOAP (CERN)
  - DPM (CERN)
  - GFAL/lcg\_util (CERN)
  - LFC (CERN)
  - FTS (CERN)
- To be assigned
  - Hydra (HIP)
  - dCache (External)

It is proposed to extend these teams (the clusters of competence) with additional resources from SA3 to undertake certification (in addition to development and testing) to form 'product groups'. A single product manager (names TBD) would be responsible for balancing staff effort within these areas between development, testing and certification. That manager must ensure:

- The software released from the area meets all of the specified testing and certification criteria in all the required environments needed for its production deployment.
- Maintenance effort is to be split between explicit bug fixing and support for agreed 'campaigns' such as platform portability, IPv6 support, error code/messages, dependency removal, etc.
- Establish contact with relevant & representative customers (deployers of the software and end-users of the software) through the relevant support function so that feedback can be given to early prototypes, the hosting of 'experimental services' and the eventual deployment of 'pilot services'.

All staff will use the agreed tools and environments and the specified processes. This will be kept to the minimum necessary but will be required (e.g. specific use of ETICS, testing processes). Compliance is not optional!

The current SA3 teams will be split into:

- Those dedicated to the 6 (or more?) product groups (each containing several clusters of competence) in doing relevant certification tasks for these areas. (i.e. resources that will be part of the gLite consortium's maintenance efforts within EGEE)
- Those dedicated to tasks that will end or be much less significant (e.g. process automation)
- Those that are part of the centre (EGI.org) whose role will be to review the certified software either accepting it or rejecting it. It will be the responsibility of the product team (including its development, testing and certification resources) to debug it.

The new product groups would use local resources for developer testing. For larger scale testing and certification resources we expect resources to be allocated from NGIs as described within the EGI model. A common minimal build and test methodology will be established. It is **not** acceptable that the implementation of these changes and the use of a consistent development methodology and release process can be vetoed or ignored by **any** individual developer.

## Detailed Changes – Additional Activity Milestones for Year II

There are a number of tasks within EGEE-III that seem to have no clear milestones (in terms of delivering something) and within the context of transition to EGI there is a need for further milestones to capture the new tasks. Below is a provisional list of additional milestones – dates & milestone subject to be confirmed – that need to be aligned with existing milestones and deliverables.

- PM 13: SA3: To ensure that all the certification tasks undertaken for each component are clearly documented so that they can be undertaken by the clusters of competence.
- PM 13: SA3: Define the base development environment (the client/SDK environment) that will be used by **all** developers to develop/build/test against and the procedures (with JRA1 if needed) that they should be following to do so correctly.
- PM 13: NA1 (TMB): With SA3 negotiate a product roadmap with the middleware providers (JRA1) and their product managers relating to the scheduled delivery of certified maintenance patches.
- PM 13: SA1 & NA4: Define an API to support operations, application and potentially middleware (e.g. FTS, L&B) messaging requirements that can be supported by the MQ system, can be wrapped to provide a compatibility library for existing RGMA users (if needed) and satisfies current and short term needs of the broader EGEE development community.
- PM 13: SA1: Develop a firm plan to remove RGMA as a dependency within the operations infrastructure. NB: all operations tools will use a single messaging infrastructure to simplify support burdens.
- PM 13: SA1: Document the interfaces currently being used by operations to integrate with the NGIs.
- PM 13: JRA1: To propose **clear** suggestions on error codes/messages (including for authorization) including a strategy to manage backward compatibility for review by the TMB. Once agreed by the TMB this must be adopted by all interfaces.
- PM 13: JRA1: Establish within EGEE a technical and management structure similar to the proposed gLite Open Consortium.
- PM 13: JRA1/SA3: To review the current software process and its supporting tools (e.g. ETICS, CVS, Savannah) in order to define which tools will be used for the remainder of the project and initially by the gLite Open Consortium.
- PM 13: NA3: Define a plan with SA1 for the coordination and integration over 3 months of the t-infrastructure with the production infrastructure and the establishment of a training VO.
- PM 13: SA2: Establish the operational procedures for the EGI and NGI operators of the ENSC and its integration into existing EGEE and future EGI infrastructures.
- PM 13: SA3: Provide a working gLite.org. Currently, gLite is defined as the software that is being run on the production infrastructure. JRA1 provides the bulk of that software, but components from other sources are integrated into it hence why this is an SA3 responsibility.
- PM 13: JRA1: Development website. Enhance the JRA1 website (Twiki) so that it becomes a clearly useful resource for any developer wishing to build upon gLite. This should include CLEAR links to the source code repository(s), issue trackers, documentation, etc.

- PM 14: JRA1: Port gLite to the platforms defined by the TMB.
- PM 14: JRA1: Verify that all software components can produce separate client and server binary packages for deployment.
- PM 18: SA3/JRA1: Prepare plans for the end of the project, in conjunction with the other software providers within EGI, for the current gLite.org website to become 'UMD.org' (for example) and to provide pointers to other software providers, and for gLite.org to become the gLite Open Consortium (JRA1) website.
- PM 15: JRA1: Review, rationalise and reduce the dependency entanglements in the code base so those individuals 'product groups' have minimal interdependencies – ideally none other than the core development environment.
- PM 15: JRA1: To document all public APIs, CLIs, error codes, error messages and implementation of the different services. Fixing error codes and error messages so that they reflect the best practice.
- PM 15: JRA1/SA3: Review the current test coverage and develop a prioritised plan to expand the test coverage.
- PM 17: SA2: Document the availability of all source code for networking operations and support through pointers to the relevant source code repositories or package. , Document the software in order that it can be built by groups other than their development groups. Review the operational experience from operating the ENSC. [Review this milestone with MSA2.3.2 with a view for merging.]
- PM 15: NA4: Integrate the EGI.org UCS transition team into the planning for the EGEE User Forum taking place in 2010.
- PM 16: NA1 (TMB): Update roadmap based on experience to date.
- PM 19: JRA1/SA1/SA3: Review all documentation to ensure that it reflects current practices.
- PM 19: NA3: Review the integration of the t-infrastructure with the production infrastructure and the establishment of a training VO.
- PM 19: NA1 (TMB): Review contractual relationship with middleware provider (gLite Open Consortium)