Subject: **Summary of the EGEE All Activity Meeting – 13TH OF SEPTEMBER 2004**

Author(s): **Marc-Elian Bégin**

Distribution **EGEE Project Members**

The following is a summary of the different presentations given during the last EGEE All Activity meeting held at CERN on the 13th of September 2004, for details please see:

http://agenda.cern.ch/fullAgenda.php?ida=a043604.

About 65 people attended the event.

Common Themes

During the AA meeting from June, we highlighted a few 'Common Themes', which we followed since. This was reflected by the dedicated presentation from Leanne Guy from JRA1 on Testing Coordination. The following lists the common themes. These themes involve several activities and are important to the success of the project. The text in italic represents the update on the original themes.

Management of Requirements

We are now converging on a common tool for the management and tracking of requirements across all activities (http://egee-na4.ct.infn.it/requirements/). The Generic and Bio-Medical requirements have already been ported to the tool. Since requirements will evolve during the lifetime of EGEE, it is important that we bring all the requirements for all application types (i.e. HEP, BioMeb and Generic) under a same roof. This issue concerns NA4, SA1 and JRA1-4 and is being discussed within the PTF (see below). *This is an ongoing theme and needs to be addressed by the PTF.*

Testing Strategy

The coordination of many testing activities is being coordinated in order to address the following issues:

- Avoid too much overlap, gaps and ensure we focus on priorities, since we cannot test all aspects
- Adopt common tools and frameworks, such that testing is harmonised across the activities and tests can be reused across the spectrum of activities
- Ensure mechanism where simulation/reprocessing (i.e. batch jobs) using gLite can be tested with the experiments as well as analysis (ARDA)

An extended Test Coordination Meeting was help at CERN on September 7th (http://agenda.cern.ch/fullAgenda.php?ida=a043635). The result of this meeting and following actions, were reported in the dedicated testing presentation. The different testing teams have agreed to:

- *JRA1/Testing and NA4/Testing will collaborate on the use-cases they are testing (e.g. end-to-end testing of services)*

- *Further to the tests mentioned above, JRA1/Testing will also perform black-box testing of individual service/components*
- *NA4 will start porting and rerunning the LCG-2 test-cases on the gLite services deployed on the Pre-Production Service (when ready)*
- *SA1/Certification will continue to use the LCG-2 test suite and on top of it run the JRA1/Testing tests – this should enforce that during the co-existence phase, both LCG-2 and gLite (partial at first) will provide expected services*
- *ARDA doesn't consider itself as a testing body as such, therefore they will continue to work independently, while participating in coordination efforts.*

JRA1 and NA4 Testing teams will work on the same CVS server (i.e. gLite repository). Further, all testing teams have agreed to support the JUnit XML (defined by JRA1/Testing) report format for tests performed on gLite.

Middleware Migration

The migration between LCG-2 and gLite must avoid "big bang deployment" by evolving existing service to the new ones in a smooth and gradual deployment strategy. An important aspect of this migration is to ensure that LCG-2 managed data (already deployed on LCG-2) is preserved and kept transparent for current users and future users of gLite. A close coordination must take place to ensure that the middleware release plan (JRA1), the grid deployment strategy (SA1) and the pilot applications (NA4) are well synchronised. *An initial dialogue between SA1 and JRA1 started, and need to be followed up. This point is now closely related to the establishment of the Pre-Production Service (PPS).*

The PPS is being setup with the current LCG-2 software stack. It is likely to be upgraded to LCG-2 running on SCL3. Then, the issue is to deploy on it the first gLite services/components, which should be released by JRA1 and then certified by SA1 for deployment in October. The first service/components will be a CE and the Workload Management Service. Now that the PPS is in place, the migration plan can take a more concrete form. We also need to integrate an EGEE schedule, which will include the JRA1 gLite release, plan, but also the certification step (including future LCG-2 releases on the PPS) and integration time required to implement the co-existence when required. The release plan of the gLite/PPS also needs to take into account application requirements, such that we can provide meaningful components and grid services to the application people.

Training

Training requires that the links between NA2, NA4 and SA1 are maintained to ensure constant supply of participants. *Since the start of the project, NA3 has held 17 training events. An NA3 Open Meeting in Karlsruhe 23rd September is being organized, to which all interested parties are welcome. Following the release of the first gLite service/components, we will need to start looking into a training programme for gLite powered applications and site administration.*

VO Management

A new theme added during this AA meeting is the management of VOs. Both from the point-of-view of policy definition, as well as the deployment/definition by site administrator, we need to clarify the 'how to' around Virtual Organisations. For example, we need to further clarify the role/responsibility/scope of VO managers, and the process by which agree on new VOs, create them

and get them setup in the selected sites. The selection/nomination process of VOs also needs to be taken into account. This last process also needs to be put in the context of the already existing structure of the EGAAP, NA4/SA1 Working Group, SA1 Weekly ROC Managers Meeting (projected), etc.

These themes need to be addressed quickly and agreed at the PEB level.

NA1

The project is, as a whole going well. The Quarterly Reports were sent to the EU with an agreed delay since it was the first quarter but we will not have the luxury of such an extension for the next quarter. The PM4 deliverables are on track, with the exception of the DJRA2.1, MJRA4.2 and DSA2.1 which are only just going into review.

We need to ensure sufficient non-HEP application feedback is provided on key deliverables, such as the architecture and design documents from JRA1 and JRA3. The timing of the feedback is also important – it must be provided during the review process for it to have sufficient weight.

Reviewers should take their role seriously: take the time to review, in depth (content not only form) and in time **this is fundamental to the success of EGEE**.

At the project level, hiring is basically complete. However, there are still outstanding hiring issues within the PO. This matter currently leaves the PO stretched in several fronts (e.g. PD deputy, dissemination).

On the reporting front, the Quarterly Reports will be easier and lighter for the 2nd quarter. This will be possible thanks to a simplified report template, taking into account feedback from all activities on the first quarter and guidance from the EU commission.

PPT

A dedicated presentation on PPT, including a live demo, was given to the AA audience. This was the opportunity to launch the long awaited online/automated timesheet capture and processing tool, which should go live in time for October reporting. This will automate the processing of the timesheets and simplify this monthly task for the project members. We are asking each project member to re-enter their timesheet for the month of July and August, which will save many man-weeks of effort to the PO.

NA2

Several key events are being organised, including:

- IST2004 in November: http://europa.eu.int/information_society/istevent/2004
- EGEE-02 Conference in November in Den Haag: http://public.eu-egee.org/conferences/2nd/

Following the agreement of the EGEE "Style Guide", key dissemination material is soon to be released, including the Power Point presentation template, different posters as well as folders and brochures.

The new look-and-feel of the external EGEE website was released: http://public.eu-egee.org/

NA3

NA3 mentioned during their presentation that they need to gather data about retreats already organised by other activities. Please ensure that NA3 are informed of any such events.

Another important point was that NA3 not only trains but also tries to help activities train themselves. This is a service that all activities should use when ever required.

NA4

EGAAP will meet again during the next EGEE Conference in Den Haag. This group is part of the 'Virtuous Cycle'.

The process to allocate generic application resources to the new applications approved for deployment is under definition. A proposal (from Roberto Barbera) is to allocate resources from PEB/PMB approval to the following EGAAP meeting. This needs to be further discussed.

With respect to the Industry Forum:

- The industry forum newsletter will be available next week on the industry forum web site
- A session at next IST2004 is under preparation
- Two work groups are being set up and are expected to report on their activity at next EGEE conference
- A new charter is under discussion

BioMed Application deployment

An applications description page has been setup at the following address:

http://egee-na4.ct.infn.it/biomed/applications.html

Currently three types of application are being pursued:

- **Pilots**: LCG2 compliant applications at day 0
- **Internal**: from project partners, to be deployed on EGEE
- **External**: from other projects, to go through a selection procedure

From these families of applications, the deployment of the following is being prepared:

- **Pilots**: GATE, GPS@, CDSS
- **Internal**: SiMRI3D, PTM3D, xmipp_MLrefine
- **External**: Mammogrid

A number of limitations are reported by NA4. The following is a summary of these limitations:

- Limitations induced by infrastructure
  - Limited VO acceptance (CNAF RB)
  - No RLS service for biomedical VO
    - CC-IN2P3 proposal to host the service, under study in September
    - CNAF proposal to temporarily set up a RLS service meanwhile
  - No MPI-enabled resources available for parallel applications
    - INFN reported that an MPI-enabled cluster exists on INFN-Grid
  - Lack of service redundancies
    - High sensitivity to CNAF RB health
- Future plans
  - Focus more on applications and less on infrastructure deployment
  - Set up a communication channels
    - Notify users of SA1 maintainance operations and problems
    - Provide feedback to SA1 on problems encountered
  - Set up a new community integration procedure with SA1 including:
    - VO creation
    - RB registration
    - RLS provision

Nevertheless, a number of sites expressed their willingness to help support the deployment of non-HEP applications by providing computing resources.

Generic applications

A number of *Generic* applications are being supported by NA4. The following applications are at different stage of readiness (they are being deployed on the GILDA testbed).

- Earth Observation
- Geophysics
- Chemistry
- Hydrology
- ESA
- Astrophysics
- Astroparticle-physics
- GRACE

ARDA

ARDA is an LCG project whose main activity is to enable LHC analysis on the grid. Using CERN NA4-HEP resources, ARDA contributes to EGEE NA4 by evaluating the very latest of the EGEE middleware: gLite. This should allow ARDA to be a key player in the HEP application evolution from LCG-2 to the EGEE infrastructure.

Further, ARDA is providing early and continuous feedback to the JRA1/JRA3 developers and guarantee that the software is what the experiments expect and need.

ARDA also helps in adapting/interfacing the HEP applications with the re-engineered middleware.

NA5

NA5 recently held a kick-off meeting. The activity is led at CERN by:

- Matti Heikkurinen (CERN) - Overall NA5 coordination and member of the PEB.
- Hannelore Hammerle (CERN) - eIRG "Virtual Office", PO liaison.

Following the feedback from the kick-off of the activity, NA5 is currently establishing new working methods. The Execution Plan has been produced. The eIRG meeting is also in preparation, as well as the "concertation" meeting, which will take place during the next EGEE conference. Finally, several white papers are being prepared.

An email list has been created: "project-eu-egee-na5" and so far 20 people are listed on the new NA5 website: http://egee-na5.web.cern.ch/egee-na5/.

NA5 needs the involvement of all other activities in order to ensure its success.

JRA1

Several key gLite deliverables are progressing well. These deliverables include:

- Architecture (gLite and security)
- Design
- Test Plan

However, delivery dates of services and components are slipping with respect to the original release plan. This situation is aggravated by other delays coming from JRA3 (security components and services). Therefore, we need to revisit and define a clear and realistic release plan up to functionality freeze. The functionality freeze is currently scheduled for November. This date also needs to be re-evaluated, including a clear list of priorities. This schedule also need to be put in the context of the Pre-Production Service and user needs.

Information Service

A large amount of the gLite codebase has been integrated in the SCM. Several services and components can now be packaged as RPMs. Several APIs have also been finalised:

- new Java API – and wrapper to old code
- new C API – and wrapper to old code

Meanwhile, the new C++ API are almost completed and a first attempt at new Python API – and wrapper to old code – is under development. Service publisher are being rewritten in Java. The SE and CE have been published on prototype testbed.

Data Management

The Metadata Catalog and gLite-IO are available for integration on the prototype testbed. The Replica Catalog and the File Access Service are currently being internally tested. Finally, the File Transfer Service based on Condor Stork and Combined Catalog service (service coordination) is under development.

The DM team is also looking at more components that may be of interest, namely PHEDEX (aka TMDB), Condor Chirp and Parrot and NeST.

Status of the Prototype Testbed

The prototype testbed is now installed with interfaces to CE and SRM components. A prototype of the GAS service has also being deployed. Further, a prototype of the Package Manager was also deployed.

On the SRM front, the castorgrid (main LCG SRM) is being enabled, while issues with e.g. gSOAP, security, etc are being fixed. A client to the Metadata Catalog is now ready for integration with the GAS. Services can now publish information to the Data Management RGMA. The AliEn build process is being adapted to the gLite SCM.

Finally, bug are being fixed and new features introduced in collaboration with the ARDA project. The prototype testbed is also being actively maintained, including reinstallation and system administration work (e.g. new location, new machines).

Tools

CodeWizard tool has been run on the code in CVS and is being integrated in the build system. This tool enforces agreed coding guidelines. Naming conventions are being discussed and agreed among the different JRA1 clusters.

On testing, CppUnit and PyUnit are among the tools used for the JRA1 testing. QmTest is being evaluated.

The existing QA reports have been adapted to the metrics specified in the JRA1 SCM document, and will be in production shortly. Several requested improvements on the Savannah configuration have been implemented (e.g. sequence of states and notifications).

Finally, J. Benard has joined the activity from CNRS as second FTE to SPI.

Integration

The Integration Team (iteam) has implemented an automated nightly build system and deployment of gLite packages to the gLite web site. The team also started work on common service configuration, logging and error handling for the different gLite services.

Meanwhile, the build system has been upgraded to include automated checks for Java coding conventions and unit test coverage in build system and CVS (C++ is on the way). The iteam also launched the official gLite web site: www.glite.org.

Good progress has been performed on the gLite Developer's Guide, with an almost completed first release. The team also coordinated and published the JRA1 Quality Assurance Plan for JRA2.

Testing

A dedicated presentation on testing was given at the end of the event. Please see Testing Common Theme for details.

JRA2

The project Measurement Plan is being finalised (this document is part of DJRA2.1). Further to the software and service based metrics that JRA2 is tracking, we also need to keep track of 'soft' metrics like 'press coverage' following press releases and also collect human values (e.g. survey).

From the current LCG-2 Production Service, initial quality indicators are being collected. We would, however, also like to collect the cause of job failures, as well as:

- Success rates of short/medium/long jobs
- Ratio of working/non-working sites
- Etc

During the discussion of the JRA1 presentation, we identified the need to nominate means/person responsible for collecting each piece of data.

In order to be able to compare between LCG-2 and gLite, we must ensure that gLite will also be capable of collecting the data currently provided by LCG2.

Finally, since not all the sites are taking part in the monitoring process, we have to facilitate the steps sites need to follow in order to contribute to the collection of the quality metrics.

JRA3

Now that all the information is available (e.g. architecture and design documents), it is important to check that JRA3 OGSA recommendations are consistent with JRA1 plans.

JRA3 is working both in terms of high-level security policy definition, but also as software providers to gLite. In the later case – i.e. development work – JRA3 is integrated in JRA1 as cluster, like all other JRA1 clusters.

The delivery schedule of the JRA3 cluster is also slipping which will cause a knock-on effect on the gLite release plan.

JRA4

JRA4 have made public a document, which provides guidelines and recommendations for software developers for writing IPv4/ IPv6 independent software.

The delivery of the deliverable "DJRA4.1 - Specification of interfaces for bandwidth reservation services" is being very tight.

SA1

On the next release of LCG-2, accounting uses R-GMA.

SA1 informed us that scripts exist for assisting in VO creation. It was also pointed-out that we need less intrusive accounting techniques and tools, since they can take a substantial amount of resources and access rights.

The main subject of discussion and interests from the SA1 participation was on the Pre-Production Service (PPS). Here are a few highlights of the new service:

- Needs to be up and running with LCG-2 by end of September
- The originally devised 5 nodes/site is likely not to be enough with the current requirements of LCG-2. 10 nodes/site would be required.

During the presentation of SA1 on the schedule of the release of the first gLite services/components on PPS, serious doubts on the feasibility of the schedule were raised by the audience. This schedule needs to be integrated in the 'Migration Common Theme' in order to have a complete picture.

During the same discussion on the PPS, different views on application use-cases came across. Therefore, we also need to take into account the user perspective, and create/manage the VOs for the applications in order for them to be able to run useful tests and deploy their application on the service.

SA2

From the presentation and the following discussions, we now need to coordinate with SA1 in order check that the advertised or agreed QoS is delivered in the current running service.

Conclusion

With the Project Conference in Den Haag in November, no other AA meeting will take place this year. Next meeting will be the dressed rehearsal beginning of February.

Thanks to all EGEE participants and carry on with the good work