



Enabling Grids for
E-science in Europe

<http://cern.ch/egee-jra1>

*EGEE All Activities Meeting
13th September 2004, CERN, Switzerland*

JRA1 Middleware

Frédéric Hemmer

on behalf of

A central globe is surrounded by various computer hardware including a server rack, a laptop, a desktop monitor, and a tower PC. Yellow arrows radiate from the globe to these devices, symbolizing a network or data flow. In the foreground, a person's hands are visible typing on a laptop keyboard.

**Alberto Aimar, Maite Barroso, Predrag
Buncic, Alberto Di Meglio, Steve Fisher,
Leanne Guy, Peter Kunszt, Erwin Laure,
Francesco Prelz**

- Summary of work since last meeting
- State of Deliverables for PM4-PM6
- Risk Analysis
- Issues related to other Activities
- High priority steps between now and Den Haag
- Hiring Status & Manpower level
- Scope & objectives of M5/6 deliverables
- Status of gLite
- DJRA1.2 – Design of Grid Services
(incl. PTF contribution on WDSL)
- Migration Strategy
(with SA1)

Summary of Work since Last Meeting

General

- DJRA1.1
 - Architecture document
 - <https://edms.cern.ch/document/476451/>
 - <https://edms.cern.ch/document/493614> (Feedback)
 - Release Plan
 - <http://edms.cern.ch/document/468699/1.0> (As delivered)
 - <https://edms.cern.ch/file/468699> (Current version - reviewed weekly)
- DJRA1.2
 - Design document
 - <https://edms.cern.ch/document/487871/>
- MJRA1.3
 - Test Plan
 - <https://edms.cern.ch/document/473264>
- Developer's Guide
 - <https://edms.cern.ch/document/468700>
- Continuous contributions of UW-Madison to the prototype
 - GridManager/Condor-C
 - Prototype testbed support, debugging and development
- Definition of requirements for evolution of the Gatekeeper has been discussed with (at) Globus
- 3 design team meetings, including one in Madison/Argonne
- [JRA1 Web site](#) and sub-webs updated significantly

Summary of Work since Last Meeting

Information Services

- SCM compliant RPMs available
- Specification document completed
 - <https://edms.cern.ch/file/490223/2/Specification.pdf>
- APIs
 - Finalised new Java API – and wrapper to old code
 - Finalised new C API – and wrapper to old code
 - Almost completed new C++ API
 - First attempt at new Python API – and wrapper to old code
- Service publisher rewritten in Java
- Site, Service, SE and CE published on prototype

Summary of Work since Last Meeting

Data Management

- Available for integration into the prototype:
 - Metadata Catalog
 - gLite-IO based on AIO and GFAL
- In internal testing:
 - Replica Catalog
 - File Access Service
- In development:
 - File Transfer Service based on Condor Stork
 - Combined Catalog service (service coordination)
- Looking at more components that may help us
 - PHEDEX (aka TMDB)
 - Condor Chirp, Parrot and NeST

Summary of Work since Last Meeting

Work Load Management System



- Transition of existing WMS components to SCM and the EGEE CVS completed.
- Prototype information collection and re-match code developed.
- CE monitoring code developed.
- Prototype LB system released for testing. This is backwards-compatible for access from LCG UIs.
- LB security layer switched from SSL/GSI to GSS API (over Globus GSI implementation, using Globus GSS extensions).
- Prototype Accounting HLR (Home Location register) released for testing.
- Prototype WMS node on the new EGEE reference architecture was made available end of July. All components come from the EGEE, SCM-compliant build system, except the Network Server and the Workload Manager, which are pending interfacing with the new Data Management services.
- Discussions on interfacing WMS with GAS services and other prototype components held.

Summary of Work since Last Meeting

Prototype

- Interfaces to CE and SRM components made available
- Prototype of GAS service deployed on the testbed
- Prototype of the Package Manager deployed on the testbed
- AliEn build process being adapted to SCM
- SRM
 - Enabling castorgrid (main LCG SRM)
 - Fixing issues with gSOAP, security, etc
- Metadata Catalog
 - Client available for GAS integration
- Data Management RGMA integration (publication of services)
- Reinstallation and system administration work (new location, new machines)
- Bug fixes and introduction of new features in collaboration with ARDA project

Summary of Work since Last Meeting

Tools



- Coding guidelines: the CodeWizard tool has been run on the code in CVS and is being integrated in the build system
- Naming conventions: are being discussed/agreed among the JRA1 clusters
- Testing: CppUnit, PyUnit and Qmtest are among the tools used for the JRA1 testing
- Quality Assurance: the existing QA reports have been adapted to the metrics specified in the JRA1 SCM document, and will be in production shortly.
- Savannah: several improvements requested have been implemented (sequence of states and notifications)
- J.Benard joined from CNRS as second FTE to SPI

Summary of Work since Last Meeting

Integration



- Implemented automated nightly build system and deployment of gLite packages to gLite web site
- Started work on common service configuration, logging and error handling for gLite services
- Added automated checks for Java coding conventions and unit test coverage in build system and CVS (C++ on the way)
- Launched official gLite web site
- Almost completed first release of gLite Developer's Guide
- Coordinated and published JRA1 Quality Assurance Plan for JRA2

Summary of Work since Last Meeting

Integration (III): gLite Web site

The screenshot shows a web browser window displaying the gLite website. The browser's address bar shows 'EGEE > gLite'. The website has a blue header with the eGEE logo and tagline. Below the header is the gLite logo and the title 'gLite Lightweight Middleware for Grid Computing'. A navigation menu on the left lists various categories: GLITE SUBSYSTEMS (ALIEN, COMPUTING ELEMENT, DATA MANAGEMENT, ACCOUNTING, LOGGING AND BOOKEEPING, MONITORING, SECURITY, WORKLOAD MANAGEMENT), DOWNLOAD, DOCUMENTATION, PACKAGES, ABOUT GLITE (EGEE JRA1, EGEE JRA3, EGEE JRA4), and ABOUT EGEE. The main content area features sections for 'What is gLite?', 'gLite News' (with a sub-section 'New gLite web site unveiled (13/09/2004)'), and 'gLite People' (listing CERN, INFN, and Datamat). A footer area contains a 'Page updated: 12/09/2004' notice and a 'W3C XHTML 1.0' logo.

eGEE
Enabling Grids for
E-science in Europe

gLite
Lightweight Middleware for Grid Computing

What is gLite?

gLite (pronounced "gee-lite") is the next generation middleware for grid computing. Born from the collaborative efforts of more than 80 people in 10 different academic and industrial research centres as part of the [EGEE Project](#), gLite provides a bleeding-edge, best-of-breed framework for building grid applications tapping into the power of distributed computing and storage resources across the Internet.

Want to know more about gLite? Read the following [presentation](#).

gLite News

New gLite web site unveiled (13/09/2004)

The new gLite web site has officially gone online on Monday 13 September. The web site offers a single point of access to public documentation, installation packages and guides and loads of other useful information. The web site has been developed by the gLite [Integration Team](#) with the collaboration of all project members using original web templates from [TERENA](#).

gLite People

The gLite software is produced as part of the EU EGEE Project funded by the European Communities. The following academic and industrial research centres are collaborating to the development of the software organized in three different Activities: [JRA1](#) (data management, workload management, monitoring, accounting, computing element, logging and bookkeeping), [JRA3](#) (security) and [JRA4](#) (network monitoring and provisioning).

 The European Organization for Nuclear Research (CERN)

 Istituto Nazionale di Fisica Nucleare (INFN), Italy

 Datamat Spa, Italy

Page updated: 12/09/2004
[About the website](#)

W3C XHTML 1.0

Summary of Work since Last Meeting

Integration (II): Build System

- Currently running on RH Linux Enterprise 3.0
- 134 gLite modules built continuously every 60 minutes
- Automatic build failure notifications to developers
- Automatic packaging of modules in three different formats (source tarballs, binary tarballs, RPMs)
- Complete nightly builds run every night at 03:30, all packaged deployed to the gLite web site. All builds are tagged for full reproducibility
- Prototype build system on Windows available (Java modules only)
- Coding guidelines, unit test and coverage automatically run for all packages (Java only, C/C++/Perl on the way)

Build System



CruiseControl Status Page

Project	Last build result	Last build time
org_glite	passed	10/09/2004 02:02:47
org_glite.alien	passed	10/09/2004 02:40:46
org_glite.ce	passed	10/09/2004 02:36:46
org_glite.ce.ce-plugin	passed	10/09/2004 02:34:46
org_glite.ce.monitor	failed	10/09/2004 02:32:46
org_glite.ce.monitor-api-java	passed	10/09/2004 02:30:46
org_glite.data	passed	10/09/2004 05:40:55
org_glite.data.api-perl	passed	10/09/2004 05:12:02
org_glite.data.catalog-file-api-perl	passed	10/09/2004 05:15:08
org_glite.data.catalog-interface	passed	10/09/2004 05:10:04
org_glite.data.catalog-meta-api-perl	failed	10/09/2004 05:12:43
org_glite.data.catalog-replica-api-perl	failed	10/09/2004 05:15:39
org_glite.data.catalog-service-fr	passed	10/09/2004 05:10:38
org_glite.data.catalog-service-meta	passed	10/09/2004 05:11:23
org_glite.data.common	passed	10/09/2004 05:09:23
org_glite.data.common-gsoap-api-c	passed	10/09/2004 05:08:14
org_glite.data.config-service	passed	10/09/2004 05:16:55
org_glite.data.io-authz-catalogs	failed	10/09/2004 05:32:55
org_glite.data.io-base	passed	10/09/2004 05:22:57
org_glite.data.io-client	passed	10/09/2004 05:28:29
org_glite.data.io-daemon	passed	10/09/2004 05:24:55
org_glite.data.io-gss-auth	passed	10/09/2004 05:21:32
org_glite.data.io-protocol-rfio	passed	10/09/2004 05:29:53
org_glite.data.io-quanta	passed	10/09/2004 05:20:08
org_glite.data.io-resolve-catalogs	failed	10/09/2004 05:30:55
org_glite.data.transfer-client-java	passed	10/09/2004 05:29:55

Summary of Work since Last Meeting

Testing

- Deployment of distributed testing and validation testbed
 - 24 Machines distributed across the three sites and growing
 - Migrating to RHEL at all sites now
 - Quattor based installation of testbed beginning at all sites
- Installation testing rpms/tarballs produced by build system on all nightly builds.
 - Full automation expected for end September
- Deployment of gLite services on RHEL underway
 - Metadata catalog, R-GMA, gLite IO, Castor SE, dCache SE, WN
- Tools and frameworks
 - Two interesting frameworks identified and still being evaluated in September
 - QMTest and Condor based framework from NMI.
- Active testing of the prototype
 - 91 problems submitted by test team members to date (41% of all bugs submitted)
- Reporting tools written to output test results in XML and translate to reports (HTML,pdf)
- Installation guide for gLite being compiled and tested by test team
- Testing coordination of all EGEE testing activities organized by JRA1 team
- Testing of gLite IO, R-GMA and WMS components underway
- Agreement reach with NA4 to share common tools, framework and reporting format for testing

Summary of Work since Last Meeting Testing (II)

Click on each RPM name to see its log file.

RPM Name	Files in RPM	Internal Dep Specified	Libs in \$LOCATION/lib	Bins in \$LOCATION/bin	Correct info in header	Doc Present	Licence Present	RPM Installs	RPM Erases	RPM Relocatable	Outcome
glite-data-catalog-interface-0.1.0-1.noarch.rpm	Files	Yes	N/A	N/A	Yes	Yes	Yes	Yes	Yes	Yes	Pass
glite-data-catalog-service-meta-0.1.0-1.noarch.rpm	Files	Yes	N/A	N/A	Yes	Yes	Yes	Yes	Yes	Yes	Pass
glite-data-common-0.0.1-4.noarch.rpm	Files	Yes	N/A	N/A	Yes	Yes	Yes	Yes	Yes	Yes	Pass
glite-data-config-service-0.1.0-1.i386.rpm	Files	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Fail
glite-data-io-protocol-rcio-0.1.1-1.i386.rpm	Files	Yes	Yes	N/A	Yes	No	Yes	Yes	Yes	Yes	Fail
glite-data-io-quanta-0.1.1-1.i386.rpm	Files	Yes	Yes	N/A	Yes	No	Yes	Yes	Yes	Yes	Fail
glite-data-io-qs-auth-0.1.1-1.i386.rpm	Files	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Fail
glite-data-io-lase-0.2.0-1.i386.rpm	Files	Yes	Yes	N/A	Yes	No	Yes	Yes	Yes	Yes	Fail
glite-data-io-client-0.2.0-1.i386.rpm	Files	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Fail
glite-data-io-daemon-0.2.0-1.i386.rpm	Files	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Fail
glite-data-io-authz-catalogs-	Files	---	NOT	BUILT	---	---	---	---	---	---	Fail
glite-data-io-resolve-catalogs-	Files	---	NOT	BUILT	---	---	---	---	---	---	Fail
glite-data-catalog-service-fe-0.1.0-1.noarch.rpm	Files	Yes	N/A	N/A	Yes	Yes	Yes	Yes	Yes	Yes	Pass
glite-data-transfer-interface-0.1.0-1.noarch.rpm	Files	Yes	N/A	N/A	Yes	Yes	Yes	Yes	Yes	Yes	Pass
glite-data-transfer-client-java-0.0.0-0.i386.rpm	Files	Yes	N/A	N/A	Warning	No	Yes	Yes	Yes	Yes	Fail
glite-data-transfer-service-0.1.0-1.noarch.rpm	Files	Yes	N/A	N/A	Yes	Yes	Yes	Yes	Yes	Yes	Pass
glite-das-pa-clients-0.0.0-0.i386.rpm	Files	Yes	Yes	N/A	Warning	No	No	No	Yes	No	Fail
glite-ll-client-interface-0.0.0-0.i386.rpm	Files	Yes	N/A	N/A	Warning	No	No	Yes	Yes	Yes	Fail
glite-ll-common-	Files	---	NOT	BUILT	---	---	---	---	---	---	Fail

State of Deliverables for PM4-PM6

Milestone	Month	Date	Description	Status
MJRA1.3	M5	08-2004	Integration and testing infrastructure in place including test plans (Release 1)	Done.
DJRA1.2	M5	08-2004	Design of grid services (Release 1)	Document circulated, TBR. WSDL location to be defined with PTF

Risk Analysis (I)

- Data Management
 - Biomed Storage Element: no effort available to build DICOM-enabled SRM
 - Delays due to
 - dependencies on/issues with external components: Stork, AIO
 - rearrangement of priorities and new requirements coming from NA4 and SA1
 - late involvement of integration and testing in the prototype
 - deployment work on the prototype which was not foreseen
- GAS
 - Remains to be proved in the practice.
 - Some doubts were expressed on scalability
 - Applications can still bypass the GAS and use services directly, but might need an additional “reliability” layer that is otherwise provided by the GAS interface components.
- Package Manager
 - Prototype must evolve to take into account the requirement for various experiments to preserve the current practices. This can turn out to be difficult given the fact that each experiment has different approach to the problem and try to impose their solutions
- WMS
 - Having multiple ways to handle jobs (AliEn, EDG RB+) may cause difficulties to assign, track and resolve bugs

Risk Analysis (II)

- Integration
 - Changing requirements in terms of supported platforms and external dependencies may require additional work and cause delays
 - Adapting existing software to common build procedure and quality constraints may cause delays and long discussions
 - Aggressive timelines in producing new or modified functionality may weaken the focus on quality
 - Lack of resources may cause shift of manpower from Integration to other activities
- Testing
 - Lack of system administrator (CERN only)
 - A sys admin to manage the testbed at CERN is needed to help with the administration of testbed machines and the move to automatic installation using quattor. The current ad-hoc system is time consuming, inefficient and poses security risks.
 - Late release of components
 - There is pressure to release to SA1 ASAP. If component delivery is late then the time available to test will be reduced, resulting in a poor quality release.
 - Lack of installation instructions
 - Components usually come with incomplete installation logs only
 - Currently spending a lot of time understanding how to install and configure services
 - Hard to get help sometimes as developers are busy developing
 - The test team is writing the gLite installation guide(s).
 - In principle, we should be given an installation guide to test

Issues related to other Activities

- SA1
 - For some components, adapting the build structure to SCM rules is proving to be more difficult than expected. To streamline the process, another round of discussion with SA1 is needed to review initial requirements on s/w packaging and deployment.
 - Changing or unclear requirements from Operations in terms of packaging and external dependencies handling may cause additional work and delays
 - Data management services development in SA1 are overlapping with those of JRA1
- JRA3
 - Data management has strong dependency on JRA3 deliverables for security model
 - No general-purpose TLS code is coming from JRA3. WMS is converting the TLS applications to GSSAPI. WMS will use WS security delegation port type when available.
 - Security testing is being discussed
- NA4
 - Continued collaboration with NA4 on sharing testing tasks is essential in view of the amount of testing to do and NA4's need to use our testbed for their testing activities

High priority steps between now and Den Haag (I)

- Information System
 - Complete C++ and Python APIs
 - Convert CLI tool (ex edg-rgma) to use new Python API
 - Refactor code to Web Services
 - Produce high quality user guide
- Data management
 - Initial versions of all essential data management services integrated into the prototype
 - Find effort to build DICOM-SRM
 - Resolve issues with SA1
- Workload Management
 - Complete WMS integration into the 'glite' prototype, start gaining operational experience via prototype user support.
 - Continue development of LB, Accounting, CE and WMS components.
 - Work on unit tests and documentation for final delivery of 'Release 1' components according to the Release Plan.

High priority steps between now and Den Haag (II)

- **Prototype**
 - Integrate all components into build and testing system
 - Add more resources (Worker Nodes) running SCL3
 - Focus on LSF CE backend
 - Install services delivered as foreseen in the Release Plan
- **Integration**
 - Integrate sets of individual CVS modules in logical deployment modules and provide packages for them
 - Finalize and agree common configuration, logging and error handling guidelines and reference implementations
 - Finalize coding conventions for C/C++/Perl and implement automated checks in the build system
 - Fully understand any issues related to porting gLite to Windows and plan for it
- **Testing**
 - Scale up the distributed testbed from the current 24 machines to ~50
 - Finalize evaluation of the testing framework and deploy
 - Put in place Quattor for automatic installation and configuration of the distributed testbed
 - Set up two more VOs, hosted each at NIKHEF and RAL
 - Understand with the PTF high priority requirements to focus testing on
 - Test all gLite components identified as high priority in the release plan

Hiring Status & Manpower level

Partner	FTE ¹	MM	Assigned ²	To Hire ²	FTE from TA ²	Deviation
CERN	19.3	232	19.3	0.0	32	0%
RAL (CCLRC)	9.0	108	9	0.0	8	0%
INFN	25.7	308	24.2	1.5	16	6%
DATAMAT	5.3	64	4.05	1.3	6	24%
CESNET	4.4	53	4.4	0.0	4	0%
CNRS	2.0	24	2	0.0	2	0%
Uchicago, USC, UW-Madison	N/A					
Total effort	65.7	788	63.0	2.8	68	4%

Resource Plan indicator

Hired or assigned up now	63.0
Total FTE from TA	68.0
Recruitment indicator	96%
Still to hire	2.8

1: People planned & assigned

2: Head count

Scope & objectives of M5/6 deliverables

- DJRA1.2 – Design of Grid Services
 - Describe the implementation of the proposed services identified in DJRA1.1
 - Provide Interfaces (API and WDSL) and semantics for those services

Status of gLite

- Service, ServiceStatus, Site, Se and CE published via R-GMA
- Metadata Catalog integrated
- I/O library integrated
- Replica catalog being integrated
 - But will be replaced
- WMS being integrated
- LSF based CE being integrated
- Prototype of GAS deployed
 - File catalog integrated
 - Metadata catalog integrated

DJRA1.2 – Design of Grid Services (incl. PTF contribution on WDSL)

- DJRA1.2 delivered on September 1, 2004
 - <https://edms.cern.ch/document/487871/>
 - Contains Public Interfaces definitions
 - No feedback yet
- PTF scheduled for tomorrow
 - <http://agenda.cern.ch/fullAgenda.php?ida=a043775>
 - Design Document
 - Security Architecture

Migration Strategy (with SA1)

- Same LCG-2 site infrastructure services:
 - SRM SE
 - LCG-2 Gatekeeper on CE
- New services will be gradually added
 - Statically (like gLite I/O)
 - Dynamically for the CE
 - Demonstrating coexistence with LCG-2
 - Allowing for certification by LCG of these components
- Data and catalog migration will be worked out

<http://cern.ch/egee-jra1>