



Enabling Grids for
E-science in Europe

*AWG face to face meeting
CERN, October 13, 2004*

NA4 biomed status report

Johan Montagnat



- Hiring completed (as of September 9th)
 - 5 funded project engineers
 - 2 in Spain (UPV Valencia, CNB Madrid)
 - 3 in France (IBCP Lyon, LPC Clermont, CREATIS Lyon)
- 23 persons registered in the VO from 15 laboratories (3 countries)
 - mainly application developers
 - end-users (physicians, biologists...) weakly represented yet
 - disseminated and loosely coupled community
- AWG representatives: Johan Montagnat, Ignacio Blanquer
- PTF representatives: Johan Montagnat, Christophe Blanchet
- MWSG representative: Christophe Blanchet
- Task Force leader: Christophe Blanchet
- VO manager: Yannick Legré

- Task (from technical annex)
 - "Once the applications are selected, the proposing institutes will be invited to carry on with their deployment on the EGEE infrastructure. A significant fraction of the resources allocated to this activity will be devoted to supporting the migration of these new applications"
- Strong pressure to get large scale biomedical applications running
 - Borderline regarding the initial goals of biomed
 - Suddenly happened in September
 - Convincing a large user community is a long run process that involves setting up a high quality service first
- Focus of project engineers
 - Strong emphasize on LCG2 deployment until September (contribute and use resources moto)
 - Strong emphasize on applications development from then
- Train a technical team
 - LCG2/glite experts able to port apps/to help in apps gridification

Applications

- Applications description page
 - <http://egee-na4.ct.infn.it/biomed/applications.html>
- Pilots
 - GATE – therapy assistance
 - GPS@ – bioinformatics
 - CDSS – diagnosis assistance
- Internal
 - SiMRI3D – magnetic resonance image production
 - gPTM3D – interactive segmentation of medical images
 - xmipp_MLrefine – alignment of electron microscopy images
- External
 - Mammogrid – diagnosis assistance in breast cancer

- LDAP server at CC-IN2P3
 - vo-biome.in2p3.fr:389/636 (with ssl), Base DN: o=biomedical,dc=lcg,dc=org
- Resource Broker
 - Production: CNAF (egee-rb-01.cnaf.infn.it), IFAE (lcgrb02.ifae.es)
 - Test and backup: UPV (xperseo.dsic.upv.es)
- RLS
 - Production: CC-IN2P3 (under configuration and testing)
 - Temporary: CNAF (datatag2.cnaf.infn.it)
- 5 UI: LPC, CC-IN2P3, UPV, IBCP (GPS@ portal), CNB
- 8 CE: LPC (2), CC-IN2P3, CGG, UPV, CNB, SCAI, LAL
- 334 WN
 - LPC (72), LPC (50), CGG (4), UPV (10), CNB (8), SCAI (14), LAL (10)
- 0.9 TB SE (+10 TB coming soon)
 - LPC (2), CC-IN2P3 (1), CGG (1), UPV (1), CNB (1), SCAI (2), LAL (1)

Still missing services

- MPI enable nodes accessible to the biomed VO
 - CNAF announced availability within the week
- Short queues for short jobs (penalizing for several apps)
 - Latency induced by the resource broker (difficult point)
 - Having dedicated resources to process short jobs (installation problem)
- Service certificate
 - Authentication of a service giving an anonymous user the ability to run a restricted number of jobs on behalf of the service
 - Not a problem anymore? Accepted policy in several US CAs.
- Stable RLS service
 - It seems that HEP is experiencing problems with the RLS service too
- Interactivity handling
 - Firewalls are often preventing connection between WNs and the UI.
- Fine grain security
 - File access control. Encryption. Anonymization.
 - Will soon become critical to enlarge the community of medical users.

Future work planning

- Demonstrations development for review
- User community enlargement
- NA4-SA1 communication
 - Need for SA1 contact to solve application deployment problems
 - Need for SA1 information on maintenance operations
 - NA4-SA1 joint group (first meeting tomorrow)
 - Privileged contact with French and Spanish rocks
- NA4-SA1 joint group mandate
 - **Define the VOs associated with new application community (number and internal structure)**
 - **Assist in negotiations for resources for new application area**
 - **Ensure ROC/CIC services available to new area – eg. Operational and user support**
 - **Ongoing brokerage for resources for application Vos**
 - **Ongoing review of community satisfaction and development**
 - **Encourage provision of test facilities by resource centres for prospective new application areas**

Demonstrations

- The Hague, November 22-26
 - gPTM3D
 - CDSS
 - GPS@
 - What is the demonstrations agenda?
- EGEE review
 - GPS@ as the primary demonstration (genomics, largest user community)
 - CDSS as a fallback (diagnosis assistance)
 - or both!?
- gLite demonstration?
 - UPV has developed a volume rendering application (based on the virtual human dataset) on gLite

Biomed task force

- Objectives
 - prepare for the review
 - application oriented support
- Leadership
 - Christophe Blanchet
- Members
 - 5 project engineers
 - 1 SA1 correspondent per application developer (TBC tomorrow)
- Actions
 - applications plan with steps and milestones
 - meeting and deviations tracking
 - training (on LCG2)

First demonstrations roadmap

- **CDSS**
 - October: packaging (RPMs) and deployment on 3 sites (UPV, CNB, LPC).
 - November: consolidation (first scenario and deployment at CC-IN2P3).
 - Until review: add more services (search engines) and clinical usecases.
- **GPS@**
 - October: use RLS to distribute replicate biomed databases on several SE. Test GFAL for accessing registered data.
 - The Hague: test short queues CEs (performance). Demonstration scenario. Deploy on 5 sites with 5 to 10 CPUs each (CNB, UPV, CC-IN2P3...).
 - December: Improve results visualization. Make some scale testing.
 - Until review: Test compound jobs. Add processing algorithms.
- **gPTM3D**
 - will be finalized next week.
 - The Hague: require 20 CPUs.
 - Later: widespread the applications in hospitals (where the non-gridified version is already available).

Tasks assignment

- Vincent Lefort (IBCP)
 - GPS@ porting and deployment on LCG2
 - The Hague demo / EU review
- Christophe Pera (CREATIS)
 - gLite testing, user reports
 - biomedical data management middleware
- Yannick Legré (LPC)
 - Biomed VO administration
 - NA4-SA1 joint group leader (with F. Harris)
 - VO organisation issues (VOMS...)
 - Resources and requirements gathering
- Angel Merino (CNB)
 - bioinformatics applications development on LCG2
 - LCG2 quantitative testing (report on apps testing and updates)
- Ignacio Blanquer (UPV)
 - CDSS deployment on LCG2, work on attracting more users
 - gLite testing

Other works in progress

- Write Once Read Many
 - Analysis of WORM model completed.
- VOMS
 - study of VOMS current capabilities and gLite integration
- Security of biomedical data
 - encryption and anonimization
 - medical metadata

Next meetings

- Task force next face to face meeting and training
 - November 3-5, at CERN
 - Conflict with LCG2 activity, date to change
- The Hague meeting
 - Task force
 - Biomed
- Biomed meeting in Sophia Antipolis
 - January
- Participation to ARDA
 - October 20-22