



*EGEE 3rd parties SA1 operational policy training,
Athens 20-21/01/05*

Enabling Grids for
E-science in Europe

EGEE Generic Applications Advisory Panel (EGAAP) and related procedures

Vasso Kotroni
NOA



EGEE is a project funded by the European Union

Presentation Content

- Why was EGAAP created
- What is EGAAP
- EGAAP Terms of reference
- Procedures
- Results up to now

Why EGAAP (1)

- EGEE main goal:
 - provide a reliable and efficient computing infrastructure for academic or industry-based applications.
 -
- EGEE application types:
 - "pilot" applications: to test EGEE middleware and to evaluate performances (under the responsibility of NA4 funded partners in charge of HEP and biomedicine)
 - "internal" applications: they involve EGEE partners in collaboration with institutes external to EGEE
 - "external" applications: come from collaborations external to EGEE and need support for deployment on EGEE
- EGEE has defined a process to host new applications interested in:
 - benefiting from this infrastructure and
 - progressively integrate their own computing resources in this framework, in order to boost their scientific reach.

Why EGAAP (2)

- The driving principle is to help each selected community to deploy their applications by dedicated
 - manpower,
 - tutorial
 - training facilities
 - access to test beds, etc.
- Since:
 - the manpower to do this is limited and
 - it could be counterproductive to deploy too many new applications simultaneously,

EGEE has set up the EGEE Generic Applications Advisory Panel (EGAAP)

EGAAP Objectives

- EGAAP acts as an interface between EGEE and its potential users in the scientific community
- EGAAP is an advisory body to EGEE management with a role to:
 - Review new applications from as many scientific fields as possible
 - Make recommendations on how to best integrate these applications in the EGEE infrastructure
 - Review the applications deployment and results

EGAAP Membership

- **5 External Members**

- Francesco Beltrame (Genova)
- Wolfgang Hiller (Bremerhaven)
- Ladislav Hluchy (Bratislava)
- Vasso Koltroni (Athenes)
- Brigitte Plateau (Grenoble)

- **3 EGEE Members**

- Malcolm Atkinson (Edinburgh)
- Vicente Hernandez (Valencia)
- Guy Wormser (Orsay) EGAAP CHAIR

- **5 Ex-officio**

- Vincent Breton (Clermond-Ferrand) - NA4 manager
- Roberto Barbera (Catania) – NA4 Generic applications Manager
- Christian Saguez (Paris) – NA4 Industry Forum Coordinator
- Bob Jones (CERN) – EGEE Technical Director
- Fabrizio Gagliardi (CERN) – EGEE Project Manager

EGAAP Meetings

- EGAAP will meet at least twice a year
 - 1st meeting: CERN, 14 June 14 2004
 - 2nd meeting: Den Haag, 25 November 14th 2004
 - Next meeting: Athens, April 2005

EGAAP terms of reference (1)

- **The EGAAP**
 - is in charge of facilitating the deployment of potential new applications on the EGEE infrastructure
 - advise the Application team in EGEE on the allocation resources.
 - collects the necessary information about the application candidates, identifies the needs to achieve this goal
- **The EGAAP makes recommendations** to that effect to:
 - the NA4 management for allocation of the NA4 resources to the applications
 - the EGEE Project Execution Board where the technical consequences of this new deployment will be examined.

EGAAP terms of reference (2)

- **Criteria used for recommendations**

- scientific interest of the proposed work, with particular emphasis on the grid added-value,
- added value for EGEE to have such an application running on its infrastructure
- coordination of the corresponding community,
- grid-awareness of this community
- minimum requirement that a small team followed the EGEE training, dedication of the community to this application,
- agreement to the various EGEE policies and especially the security and resources allocation policies.

EGAAP terms of reference (2)

- EGAAP will hear regularly reports from the deployed applications on the EGEE infrastructure
- In the case of industrial applications, the EGEE Generic Applications Advisory Panel can require input from the EGEE Industry Forum

Procedures

- The Virtual Organisations should submit a request to have their applications deployed on the EGEE infrastructure.
- The proposal must be sent to EGAAP chair, Guy Wormser
- The proposal should describe
 - the program of work of the Virtual Organisation,
 - its internal structure,
 - its technical contact points, their expertise and availability,
 - the training needs, etc.

The questionnaire is available on the web page:

<http://egee-na4.ct.infn.it/questionnaire/na4-genapp-questionnaire.doc>

Supported VOs



- Proposals evaluated during the 1st EGAAP Meeting:
 - **Astrophysics (MAGICS) – H. Kornmayer**
 - **Earth Sciences (ozone maps, climate) – M.Petitdidier**
 - **Computational Chemistry – A. Lagana**
 - **Industrial applications (SIMDAT Project)**

The three proposals have been recommended to NA4

For the last proposal EGAAP recommended to pursue collaboration with
EGEE

Evaluated VOs

- **Proposals evaluated during the 2nd EGAAP Meeting:**
 - **Earth Sciences, by M. Petitdidier**
 - **Cosmology (PLANCK), presented by F. Pasian**
 - **Drug Discovery , by M. Hoffmann**
 - **GRACE search engine**
 - **EGRID.by I. Stefano**
- **Medical imaging GRIMI, by A. Riposan**
 - For the Earth Science: the committee viewed very favourably the further deployment of two new applications, Solid Earth Physics and Hydrology.
 - For the PLANCK project the committee recommended the acceptance.
 - For the Drug Discovery project the committee recommended the acceptance.
 - For GRACE project the committee recommended the acceptance.
 - For EGRID recognized the high potential scientific value of the application, but did not recommend its approval at this stage.
 - For GRIMI project the EGAAP recognized the high potential scientific value of the application, but did not recommend its approval at this stage.

Requirements

Contents

- Earth Sciences Requirements (short form)
- Earth Sciences Requirements (long form with priorities)
- Astro-particle Physics Requirements



EO Applications Requirements

- Improvement on metadata handling (RMC, Spitfire)
- Improvement on security (restricted access, groups and roles within a VO)
- Improvement on scalability (e.g. number of files, file sizes).
- Improvement of the performance of most of the functionalities.
- Support for parallel programs (e.g. MPI, PVM)

Climate Applications Requirements

- Handling of large data sets stored in different places and formats: This may include data sets combining hundred-thousands of files or other data objects for one analysis
- A secure but comfortable access to the mass storage archives (single-sign-on, data security, resource sharing)
- Resource broker services for transparent access to data and processing tools
- Concepts for load balancing, quality of service and assurance of data quality
- Definition of the role and permission for each partner of a VO to access metadata and data
- Integration of different scripting languages and COTS.

Solid Earth Physics Applications Requirements

- MPI with Mirinet, SCI or other fast network and Fortran 90: The simulation proposed requires the possibility to submit a script and not only the executable. Modeling and imaging are also intrinsically parallel applications using large memory, medium to large exchanges between computing nodes with MPI, local scratch disk space and are heavy cpu intensive. Another requirement, not encountered in DataGrid, is the consistency between the compiler and libraries available on the UI and the ones needed by the cluster at the working node where the job is executed.
- Definition of the role and permission for each partner of a VO to access metadata and data
- Operational databases: update, mirroring, integrity, secure access and metadata handling.
- Secure access to external database such as the GPS and seismological databases