Enabling Grids for E-science in Europe

www.eu-egee.org

NA3 Induction Course, 17th May 2004

The EGEE Project

Fabrizio Gagliardi
Project Director

EGEE is a project funded by the European Union under contract IST-2003-508833

Contents

Enabling Grids for E-science in Europe

- Why EGEE?
- The EGEE challenges
- CERN's role in the Grid
- LCG: LHC Computing Grid
- EGEE partners
- EGEE applications
- Pilot applications
- Related Projects
- EGEE project structure
- EGEE Management Structure
- The Project Office
- Timesheets: what and why
- Project Office issues

Why EGEE?

Enabling Grids for E-science in Europe

- A lot of investment from previous projects both at national and international level
- For once Europe is not lagging behind (yet) more advanced IT regions (US and Japan)
  - NYT article on 11/11/03 gives EU a 12-18 lead to Europe on Grid deployment
- Important to keep momentum and preserve the human asset and resource investment so far O(100 MEuros) in FP5
- 100 M Euros already invested in first FP6 phase, another 160 M foreseen in second phase
- More investment possible in FP7 (if success in FP6 continues)
- Project Director and senior partners already working on this

The EGEE challenges (I)

Enabling Grids for E-science in Europe

- A large investment in a short time (32 M Euros/ 24 months):
  - The rationale is to mobilize the wider Grid community in Europe and elsewhere and be all inclusive
  - Demonstrate production quality sustained Grid services for a few relevant scientific communities (at least HEP and Bio-Medical)
  - Demonstrate a viable general process to bring other scientific communities on board
  - Propose a second phase in mid 2005 to take over EGEE in early 2006
- Move from R&D Middleware and testbeds to industrial quality software and sustained production Grid infrastructure performance
- Implement a highly distributed software engineering process while maintaining efficiency and a fast release cycle (development clusters)
- Harmonize EGEE activities with national and international activities
- Cope with new FP6 rules and different and often conflicting EU Grid plans and activities

The EGEE challenges (II)

Enabling Grids for E-science in Europe

- On a more technical ground:
  - How to keep the present GT2 based production middleware running on the production infrastructure, while developing a “simple” prototype from different and disparate building blocks?
  - Are the above two processes going to converge in the short time of the project life?
  - Where is the overall architecture developed? Is everybody convinced we need one (a part from his/her own?)
  - Do we have a process in place to integrate new VOs in SA1?
  - How to support effectively new VOs other than HEP in NA4?

CERN's role in the EGEE

Enabling Grids for E-science in Europe

- LHC poses **unprecedented computing challenges**
- **LCG project** and **Grid technologies** are **CERN responses**
- Also for this reason CERN is the **lead partner** for the **EGEE** project which will provide a grid infrastructure for several application domains

LHC Computing Grid Project (LCG)

Enabling Grids for E-science in Europe

- **EGEE builds on the work of LCG to establish a grid operations service**
- **LCG:** a worldwide collaboration of
  - *The LHC experiments*
  - *The Regional Computing Centres*
  - *Physics institutes*

- **Mission:**
  - Prepare and deploy the computing environment that will be used by the experiments to analyse the LHC data
- **Strategy:**
  - Integrate with EGEE in SA1 (Grid services) and JRA1 (Middleware)
  - Coordinated management structure
- **Status**:
  - LCG service up and running with LCG-2 mware – successfully being used for LHC data challenges

EGEE Partners

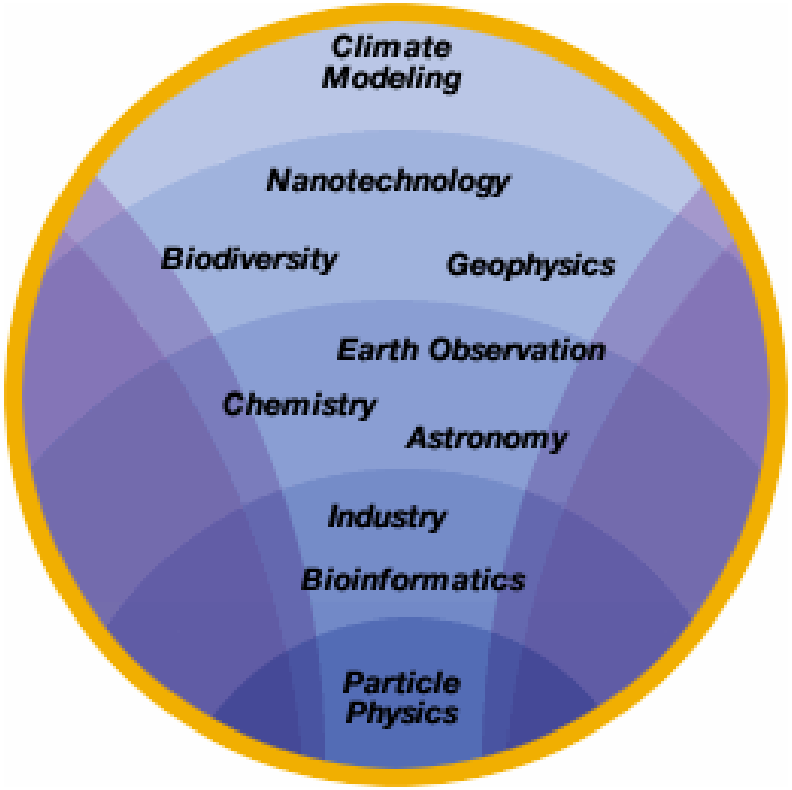
Enabling Grids for E-science in Europe

- 70 leading institutions in 28 countries, federated in regional Grids
- Leverage national resources in a more effective way for broader European benefit

From the EGEE proposal: Applications

Enabling Grids for E-science in Europe

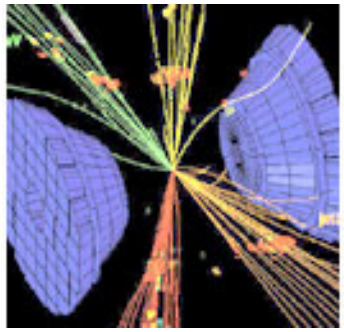
- EGEE Scope : ALL-Inclusive for academic applications (open to industrial and socio-economic world as well)
- The major success criterion of EGEE: how many satisfied users from how many different domains ?
- 5000 users (3000 after year 2) from at least 5 disciplines
- Two pilot applications selected to guide the implementation and certify the performance and functionality of the evolving infrastructure: **Physics & Bioinformatics**



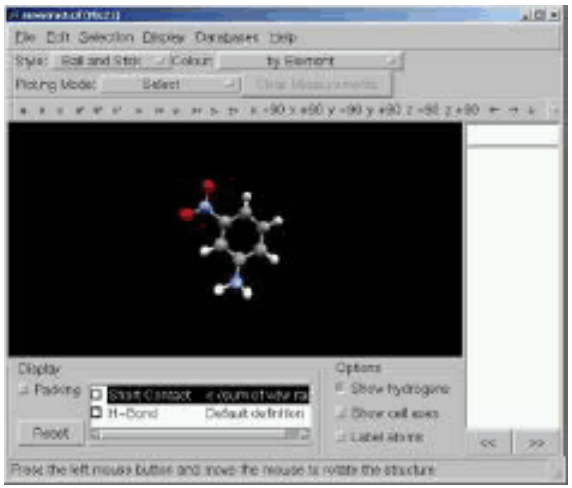
The pilot applications

Enabling Grids for E-science in Europe

- **High Energy Physics** with LHC Computing Grid (www.cern.ch/lcg) relies on a Grid infrastructure to store and analyse Petabytes (10^{15} bytes) of real and simulated data. LCG is a major source of resources, requirements and hard deadlines with no conventional solution available



- In **Biomedics** several communities are facing equally daunting challenges to cope with the flood of bioinformatics and healthcare data. Need to access large and distributed non-homogeneous data and important on-demand computing requirements



EGEE Related projects

Enabling Grids for E-science in Europe

- From the EGEE mandate, be open and play an infrastructure role:
  - SEE-GRID, South Eastern European Grid-enabled eInfrastructure development: extends EGEE to South East Europe http://www.see-grid.org/
  - DEISA , Distributed European Infrastructure for Supercomputing Applications: Supercomputing grid http://www.deisa.org/
  - Diligent: A Testbed Digital Library Infrastructure on Grid Enabled Technology: (in advanced negotiation) starts in September or October 2004
  - GRID-CC (in advanced negotiation): Real-time Grid applications
  - US projects (Trillium, GRID3, OSG etc.)
  - BioMedical and other EU projects from the current round of EU negotiation (will be known by June)
  - Other countries have expressed strong interest in the project: Korea, Taiwan, Egypt, Pakistan, India, Cuba, Chile, Iran…

EGEE Project Structure

Enabling Grids for E-science in Europe

32 Million Euros EU funding over 2 years starting 1st April 2004

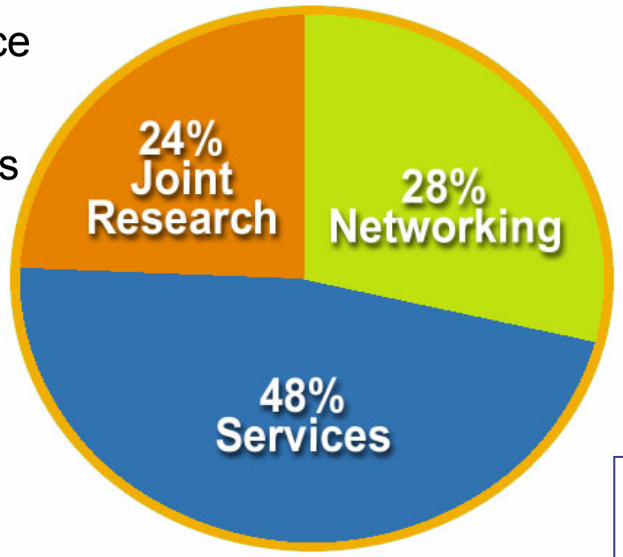
24% Joint Research

JRA1: Middleware Engineering and Integration

JRA2: Quality Assurance

JRA3: Security

JRA4: Network Services Development



28% Networking

NA1: Management

NA2: Dissemination and Outreach

NA3: User Training and Education

NA4: Application Identification and Support

NA5: Policy and International Cooperation

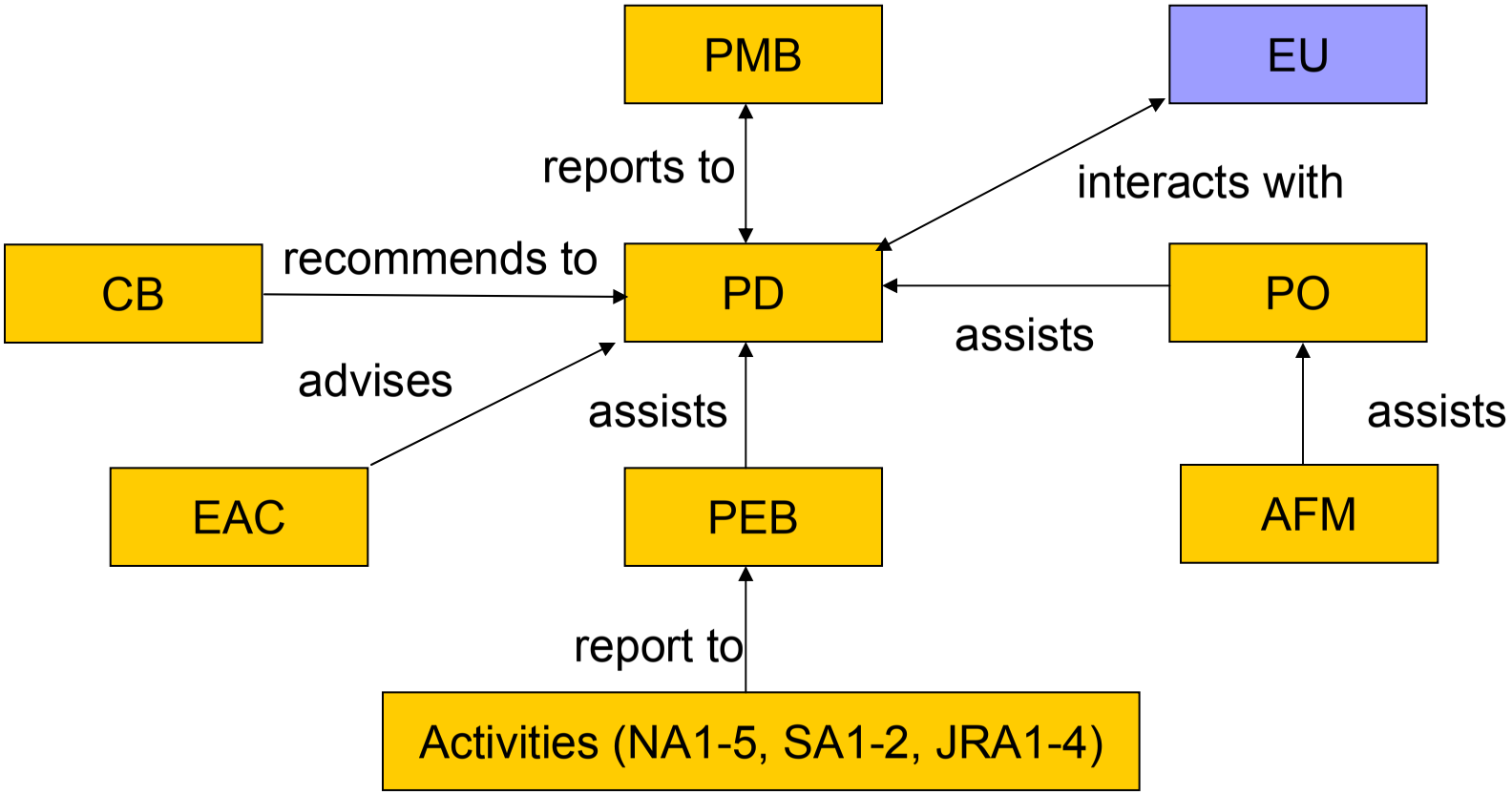
48% Services

SA1: Grid Operations, Support and Management

SA2: Network Resource Provision

Emphasis in EGEE is on operating a production grid and supporting the end-users

Management structure



CB	Collaboration Board
EAC	External Advisory Committee
EU	European Union
PD	Project Director

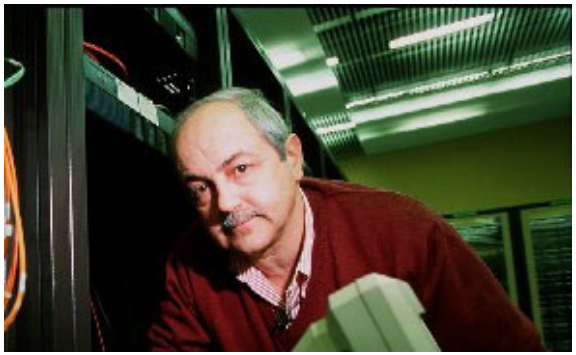
PEB	Project Executive Board
PMB	Project management Board
PO	Project Office
AFM	Administrative Federation Meeting

The EGEE Project Office

Enabling Grids for E-science in Europe

- The PO provides support to the Project Director and coordinates information with all 70 partners

Project Director: Fabrizio Gagliardi



Project Technical Director: Bob Jones

Project Requirements

Enabling Grids for E-science in Europe

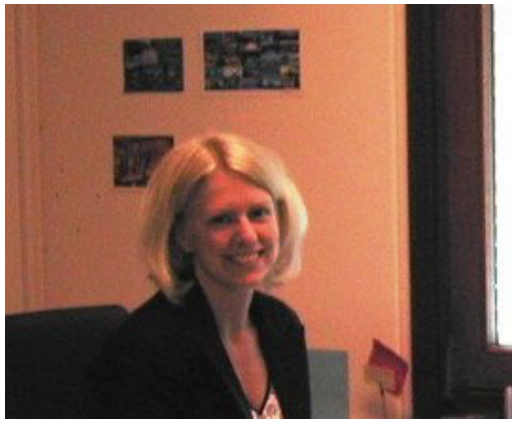
- The Project Director (Fab) is the pivotal figure interacting with the partners on the one hand, and with the EU on the other
- The Technical Director (Bob), coordinates all technical activities and is deputy to the Project Director
- The project follows the FP6 rules set by the EU and therefore a set of specific deliverables and requirements are imposed
- The PO is there to communicate and request the various administrative matters from partners to ensure that the EU's requirements are met
- Let's meet the team…

The Project Office - 1

Enabling Grids for E-science in Europe

Project Administrators:
Kristina Gunne

Karin Burghauser



Pierrette Colaci

Karin (60%), Kristina and Pierrette (50%, she leaves on June 30, 2004) assist the Project Director in his numerous managerial and administrative tasks, and provide support for all administrative issues. They are multi-tasking experts!

The Project Office - 2

Enabling Grids for E-science in Europe

Project Secretary: Anna Cook

Anna deals with general project administration (contracts, reporting, meetings, etc.) and she manages the Project Office staff

Financial Officer: Severine Bergerot

Severine is the main contact for all financial matters, cost claims, timesheets, etc.

The Project Office - 3

Enabling Grids for E-science in Europe

- Project Office PR and link to dissemination and outreach activities (NA2): Leticia Martignon

Rosy Mondardini (60%)

Marc-Elian Begin

• Support the Project Director in relations to dissemination activity, press and media, serve as a link between the project office and the outside world
• Liaise with Terena (NA2) for production and distribution of information material for various EGEE, Grid and specific public relations events
• Co-ordinate demonstrations of EGEE at various events
• Identify news about the project and disseminate it to the wider community
• Maintain the internal project web site (with temporary support by Marie Laure)

The Project Office - 4

Enabling Grids for E-science in Europe

- EGEE related project section (RP). Part of NA5 activity. Relations to Grace, Diligent, SEE-GRID and exploratory role in other potential partner projects. Roberta Faggian, SL, Juergen Hofer, technical student (Grace), leaving soon, a new TS, Jan Fiete Grosse-Oetringshaus, is being hired; 2 FTEs for Diligent, and 1 FTE for SEE-GRID and eIRG support.

The Project Office - 5

Enabling Grids for E-science in Europe

- Web Assistant: Marie-Laure Bourgeois-Schutz

Marie Laure assists Leticia in the internal web site maintenance and, with Karin, is responsible for registering external EGEE partners at CERN

Her contract ends on September 30, 2004

And finally, timesheets What and why

Enabling Grids for E-science in Europe

- Among the various requirements for the project to unfold seamlessly, timesheets have become the “hot topic” of the moment
- Each person participating in the project must provide a timesheet on a monthly basis. The timesheet must report the number of hours worked in which task of which activity
- This timesheet must be approved by the local management, and passed on to the designated “collector” in their institute, who will then submit them all to the project office for validation by the PEB
- The timesheets must satisfy the partner internal accounting and auditing rules and procedures
- A template has been provided and made available on the project website: http://egee-jra2.web.cern.ch/EGEE-JRA2/TimeSheet/TimeSheet.htm
- An electronic tool is being developed at CERN to help the reporting process. Further information will be provided once it is in place

Project Office issues

Enabling Grids for E-science in Europe

- EGEE Project office does not scale from the previous EDG project office experience (except maybe on the dissemination and PR part)
- Technical co-ordination under resourced (2 FTEs in EDG, only 1 in EGEE). Very critical issue.
- Still in a start-up phase, with the exception of Bob, Karin and Rosy, everybody in the PO is new on board and without prior experience of EU, CERN and similar projects.
- Open to suggestions, criticisms and ready to improve in any possible ways.
- Counting on your patience and support while getting up to speed.

The Project Office - Contacts

Enabling Grids for E-science in Europe

The project office can be reached on the following email address:

Project-eu-egee-po@cern.ch

And for all timesheet matters ☺:

Project-eu-egee-reporting@cern.ch

For further information on the project: www.eu-egee.org