



Enabling Grids for E-science

Grid Observatory Cluster

NA4 F2F meeting
03/28/2008

www.eu-egee.org



- **Goals: Integrate the collection of data on the behaviour of the EGEE grid and users with the development of models and of an ontology for the domain knowledge**
- **LRI (CNRS), Universita Piemonte Orientale, London IC**
- **Services and clients**
 - Data Portal
 - Public access to *a posteriori* traces of EGEE activity
 - Computer science and engineering
 - Possibly internal quality insurance, metrics
 - Metrics: access counter, registered users,
 - Analysis
 - Propose realistic methods to answer real operational issues: from fault detection and diagnosis to BDII scalability. No operational development.
 - Metrics: a ranking by SA1 would be nice; number of scientific publications.
- **Newest and smallest cluster**
 - Risk: From none to unrealistic requests
 - Catalyze projects
 - Collaborate with related projects eg answer to RENCI request
 - No metric, but should contribute to evaluation

Data Collection & Publication

- **Comprehensive collection: oriented to long-term, systematic collection**
 - Identification and evaluation of sources
 - Lightweight sustainable acquisition protocol
 - Legal issues and validation process
 - Database schema – depends on users requirements
- **Partial collections: be able to collect example but comprehensive data soon after the creation of the GO, and to publish them.**
 - Bootstrap internal and external collaborations.
 - Drive the comprehensive collection subtask
 - From downloadable files “as it is” to partial implementations of the comprehensive collection scheme.
 - *Raw traces: L&B; internal logs of WMS,SE and MAUI; diff BDII – secured thanks to the DEMAIN programme and LAL contribution*
 - *Summary data: the Real Time Monitor (GridPP), IC contibution*
- **Portal partially outsourced – funding secured for one year**

Analysis

- **Fault detection and diagnosis**
- **Time-series models**
- **BDII scalability if external support**

- **Procedures**
 - Data access
 - Excerpts freely downloadable, complete logs need registration
 - Analysis
 - None so far, open question
- **Communication**
 - Cluster
 - Wiki, mailing list, indico: currently not coherent
 - Other EGEE bodies
 - SA1, data collection: evaluation of the acquisition protocol, deployment. TCG?
 - SA1, TCG and Quality insurance, analysis task: see analysis
 - PEB: validation
 - External
 - The portal is the major communication medium
 - Advertise its creation in relevant newsletters and sites, RSS
 - Participation in EGEE communication actions if invited
 - Conferences: EGEE UF, ICAC, CCGrid