

Grid Observatory Cluster

NA4 F2F meeting 9-10 June 2008



www.eu-egee.org

INFSO-RI-031688



Newest and smallest cluster with a most ambitious goal

Its aim is to develop a scientific view of the dynamics of grid behaviour and usage. The Grid Observatory will make available traces of users and middleware activity on the grid for study by computer scientists, based on the existing extensive monitoring facilities. These data will excite researchers in both the grid and machine learning areas and improve the connection between the EGEE project and computer science, with a specific impact sought to the emerging field of autonomic computing. Modelling the dynamics of both the grid and the user community will contribute to a better understanding of the areas of dimensioning and capacity planning, to performance evaluation, and to the the design of self-regulation and maintenance functionalities such as real-time fault diagnosis Eventually, the improved understanding of grid activity may improve the reliability, stability, and performance of the grid itself



- LRI (CNRS), Universita Piemonte Orientale, London IC
- Data Portal
 - Public access to a posteriori traces of EGEE activity
 - Target: computer science and engineering, possibly internal quality insurance
 - 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?*
 - Scientific research
 - Metrics: a ranking by SA1 would be nice; number of scientific publications.
- Focusing interactions with CS
 - Catalyze projects
 - Collaborate with related projects eg answer to RENCI request
 - No metric, but should contribute to evaluation

egee

- Enabling Grids for E-sciencE
- 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 (CoreGrid, Autonomic Computing, RENCI)
 - 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 GRIF/LAL contribution
 - Summary data: the Real Time Monitor (GridPP), IC contibution
 - To come: user-oriented workflow logs
- Portal web design and hosting partially outsourced to HG funding secured for two years (CNRS grant)
- Released planned September 2008
- Urgent: short docs of the formats and presentation see below (IC?)

egee

- 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
- Ontology was in the initial plans, not a practical goal with the actual level of funding
- Interesting starting point: Common Base Event Model <u>http://www.ibm.com/developerworks/library/ac-cbe1/</u> possible collaboration

✓ Assessment



- Practical issues
 - Risk1: From none to unrealistic requests
 - Risk2: useless developments
 - Need for a structured interaction with SA1 and JRA1
 - Request for a TCG meeting, should be prepared by preliminary consensus
 - <u>Tracker form</u> as a small starting point
 - Tracker implementation and maintenance (Cern savannah?)
- Scientific research
 - Fault detection and diagnosis
 - Time-series models
 - HMM
 - Reinforcement learning
 - 2 year postdoc on time-series and RL, with expert collaboration
 - ✓ Report on required data schema
- Shared: how to build the trust in autonomic tools



- 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
 - ✓ Killer Demo?
 - Wiki and other collaborative tools (knowledge base etc.):
 - Which ones?
 - Who?



- Wiki, mailing list, documents
 - <u>https://listes.cru.fr/sympa/info/gridobs</u>
 - https://listes.cru.fr/wiki//gridobs/
 - <u>https://listes.cru.fr/sympa/d_read/gridobs/</u>
 - Subscribe!
- Meetings
 - Periodicity to define



Adminstratrivia

- Register
 - Send a copy of your passport to
 - eu-egee-registration@cern.ch
 - Fill the form
 - http://egee-technical.web.cern.ch/egeetechnical/procedures/Registration/EGEE3/Egee3-ExternalForm.htm

Need to know who is your administrative supervisor

• Each month: fill the PPT timesheets