



The new EGRID infrastructure

An update on the status of the
EGRID project

S.C. on behalf of the EGRID team



The new EGRID infrastructure

The EGRID project:

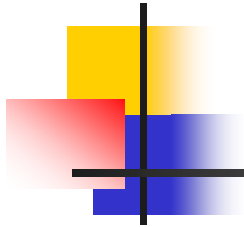
- To implement Italian national grid facility for processing Economic and Financial data.
- Underlying fabric on top of which *partner projects* develop Economic and Financial applications.



The new EGRID infrastructure

Summary:

- I. The first EGRID release
- II. Redesigning EGRID
- III. A web portal to access EGRID
- IV. applications within EGRID
- V. Dissemination/training activities



I. the first EGRID infrastructure



The first EGRID release (10/04)

- HW infrastructure to store+manage 2TB Stock Exchange Data: NYSE, LSE, Borsa di Milano, etc.
- Bulk computing power access and bulk storage rented from INFN Padova within GRID.IT
- Employed same EDG middleware INFN uses
- Facility organised for raw data pre-processing and end-user applications.



The first EGRID release

data privacy: EDG's data access mechanism implied critical and fragile fine-tuning.

- Classic SE: local files exposed through GridFTP.
- GridFTP allows file manipulation *compatible* with underlying Unix filesystem permissions.
- The underlying filesystem must be carefully managed (for details see www.egrid.it)

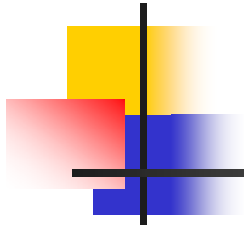


Main problem: privacy

Privacy:

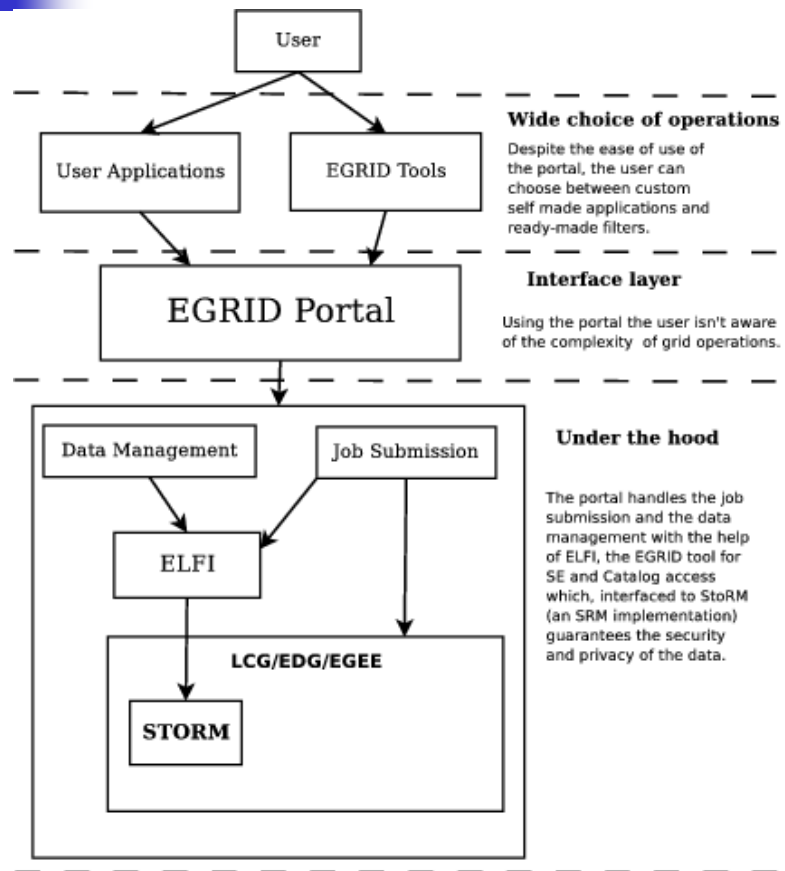
- EDG and successor middleware LCG still lacked data access mechanism strong enough for EGRID.
- Implemented solution is complex and does not scale: real account for each user in each SE, permissions on filesystem make tree replication tricky, etc...

The middleware did not allow a solution in line with a pervasive grid view.



II. Redesigning EGRID

Redesigning EGRID



- Driving factors:
 - Leaner and more general infrastructure
 - Robust privacy
 - Thoroughly re-examined grid usability



Redesigning EGRID: privacy via SRM ad hoc implementation

- Classic SE replaced with specific implementation of Storage Resource Manager (SRM) protocol currently being completed.
- Implementation is result of StoRM collaboration with INFN-CNAF.
- Not a proprietary solution – SRM becoming standard for grid disk access: security solution compatible with mainstream grid trends.
- EGRID in charge to implement security/privacy issues within StoRM



Redesigning EGRID:elfi

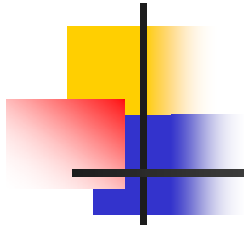
- ELFI tool developed to allow classic POSIX I/O software interface access to grid files.
- ELFI is FUSE filesystem implementation: grid resources are seen through local mount points.
- All existing file management tools work automatically with grid files:
 - Text tools: cat, grep, etc.
 - Graphical tools: Konqueror, etc.
- ELFI speaks SRM protocol: there is lack of production quality SRM clients.
- More info: www.egrid.it/elfi



Redesigning EGRID :web portal

Grid usability:

- Web portal key solution: portals long proved to be effective ways to allow user interaction with organisation's information system.
- Old command line tools will remain:
 - For backwards compatibility.
 - For few users that eagerly adopted them.
 - New development will concentrate on web portal.



III. A web portal to access EGRID



A web portal to access EGRID

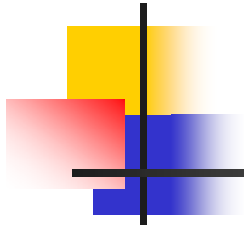
- Main entrance to new EGRID infrastructure.
- All tools in one place + Graphical UI:
 - Closer to users' way of working.
 - Lowers resistance to new technology.
- No need to install grid SW on users' workstation:
 - Interaction through portal as displayed in web browser.
- P-grade chosen as portal technology:
 - Sufficiently sophisticated as starting point to meet EGRID requirements.
 - Does not fully meet EGRID requirements: extra development needed.



A web portal to access EGRID

- P-grade portal adds new functionality: Workflow manager
 - Graphically specify several jobs.
 - Define connections among them showing data flow.
 - Portal takes care of retrieving job output and feeding it to linked jobs.
 - Monitoring of workflow done graphically showing data flow.
- Extra development needed by EGRID:
 - Improved proxy management (done)
 - SRM data management (done +via elfi)
 - SRM support in Workflow Support for special workflow jobs: *swarm jobs*

(On going collaboration with P-grade developers to better define requirement and study feasibility)



IV. Applications within EGRID



Different kinds..

- Financial group already ported several ones:
 - Palermo/UCLA/Columbia University/Roma “La Sapienza” / St.Gallen University/ University of Firenze..
 - Many other to come
- Different fields involved as well:
 - BIOCDV (IRST/Trento)
 - Computational physics (QMC and classical MD) coming from Democritos National Simulation Center



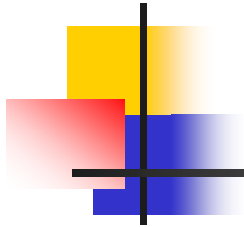
One Example from financial applications:

- A risk management application based on Genetic Algorithms (GA) and Kalman Filter (KF).
- Developed by University of Columbia (Stefano D'Addona)
- The application takes the history of a set of assets and produces a forecast.
- This forecast should be done in less than few minutes
- We therefore need **REAL TIME COMPUTING**



Our implementation: Job reservation...

- We submit many requests in advance in order to have resources ready when needed.
- Once each job is running, it waits until the user has some data to process.
- On the UI there's a server program that accepts connections from the WN and sends computational requests to them
- More details: R. di Meo's presentation in session 2C (special types of jobs)



V. dissemination/training activities



Grid in finance 2006

- Palermo 3-4 February 2006
- 50 people mostly from commercial companies and banks
- This community expressed precise requirements:
 - Standards
 - Security
 - Interoperability ..
- We felt a clear risk that private sector will go its own way in financial grid implementations.

www.gridinfinance.org



Training activities

- 3 training events (1-day) for [Grid@Trieste](#) projects
- 2 week workshop in “Porting scientific applications on computational grids”
 - 6-17 february 2006 (see www.ictp.it)
 - Mainly devoted to development countries
- More to come as net result of the previous one.
- Strong cooperation with GILDA team



Conclusions:

- Egrid community is using the GRID.IT production grid and has already ported a few applications even coming from different fields (bioinformatic/ computational physics..)
- Egrid project is actively involved in training and dissemination activities (expecially toward developing countries)
- Egrid is giving important software contributions to grid middleware (Storm/Elfi/LiveCD)
- Egrid project is addressing important issues
 - privacy and security
 - Real time approach
 - Accounting and grid on demand
- Egrid project is now acting as aggregation point (at least in Italy) for financial/economical institutions interested in Grid technologies



EGRID within EGEE

- We try to follow mainstream grid developments within EGEE project but
 - Our privacy and security issues are still to be fulfilled by the production grids
 - We did not receive enough support from EGEE developers community