IL NUOVO CIMENTO Vol. 32 C, N. 2 Marzo-Aprile 2009
DOI 10.1393/ncc/i2009-10381-0

Colloquia: CSFI 2008

Present and future of one of the largest Grid infrastructures in Europe

L. Carota([1]), T. Ferrari([1]), L. Gaido([2]) and F. Nebiolo([2])
([1]) *INFN CNAF - Bologna, Italy*
([2]) *INFN, Sezione di Torino - Torino, Italy*

(ricevuto il 22 Giugno 2009; pubblicato online l'1 Settembre 2009)

Summary. — This summary paper illustrates the Italian production Grid infrastructure, its mission and its evolution strategy.

PACS 29.85.-c – Computer data analysis.
PACS 07.05.-t – Computers in experimental physics.

1. – The Italian production Grid

The Italian Grid infrastructure aims at supporting national and international research communities by enabling the uniform and standard access to heterogeneous distributed ICT resources that are procured, owned and administered by various Italian research organizations.

It is one of the largest in Europe and currently comprises more than forty resource centers, which are hosted and operated by various research institutes and by the academy, including: INFN, CNR, ENEA, SPACI, INAF, Scuola Normale di Pisa, and other Universities. It is fully integrated with the pan-European Grid infrastructure EGEE [1,2], built and operated under the umbrella of the project EGEE-III, and is currently strengthening its integration with various Italian Grid regional infrastructures.

The Italian Grid infrastructure provides a distributed computing platform based on the principle of resource sharing by users from many scientific disciplines, such as Physics, Astrophysics, Biology, Health, Chemistry, Geophysics, Economy, and Finance. It actively contributes to the promotion of Grid development and deployment activities in Italy.

It currently comprises more than 5900 CPUS plus 9.5 PB of disk and tape storage from about fourty certified sites. The uniform access is made possible by the *INFN Grid Release* [3], a customization of the EGEE gLite middleware distribution.

2. – The backbone

The core multi-VO (Virtual Organization) Grid services constitute the *backbone* of the national Grid, enabling the transparent access to the resources hosted and operated

by the individual resource centers. Such shared infrastructure comprises—at the time of writing:

- twenty-two Workload Management System and eleven Logging and Bookkeeping services [4] (for the distribution of computational jobs to the resource centers according to the user requirement and the matching available computing farms);
- one Local File Catalogue (LFC) server for regional VOs;
- five top-level information services (BDII) operated in load-balancing mode;
- four VO Management Services for authentication of members from regional VOs, one myproxy server;
- thirteen usage record repositories for collection of job accounting information from the various sites, and one File Transfer Service for data management [5].

Moreover, the operation of a production Grid requires the availability of additional infrastructures for beta-testing of new middleware releases and for pre-deployment of certified middleware components.

The national Grid backbone is constituted by various auxiliary services for the monitoring and management of the overall infrastructure, including:

- Repositories of middleware distributions.
- Monitoring and alarm tools to proactively keep the functionality of the production Grid under control.
- Accounting services (currently just limited to computing resources)—namely the Distributed Grid Accounting Service (DGAS) [6]. Based on a client/server architecture, DGAS stores accounting information in a set of distributed repositories (the Home Location Register—HLR) organized in a two-tier hierarchy. The usage records are then aggregated and displayed via a web portal—HRLmon [7], and centrally aggregated and displayed by the EGEE accounting portal.
- The Grid helpdesk to assist Grid users and Grid site managers: it is fully integrated with the EGEE Global Grid User Support system (GGUS) [8] via a Web service interface, which allows the bidirectional exchange of trouble tickets. The EGEE central helpdesk is mainly meant for end-users, while the regional helpdesk is the primary tool for communication between the Regional Operations Center and the Italian site managers. It is based on the XOOPS tool and provides a number of support units called *Departments*—typically one for each group of relevant release components, and one for each production site.

The Italian Regional Operations Center is a distributed team of people responsible of the production and testing of the INFN Grid release, of the deployment and management of the Grid infrastructure backbone, of testbeds and pilot infrastructures, of infrastructure monitoring, of the support and, finally, of the infrastructure planning in agreement with the users' and resource providers' requirements.

3. – The Italian Grid Initiative

Grids in Europe need sustainability and restructuring to ensure the long-term running of infrastructures supporting international research collaborations. The Italian Grid Initiative [9] will foster the evolution of the current production infrastructure into a National Grid Initiative (NGI), fully integrated with the European Grid Initiative (EGI). The NGIs are the legal entities which ensure the operation of the grid infrastructures in each country, as well as a transparent representation of the requirements of all their scientific communities together with resource providers and all e-Infrastructure–related institutions.

* * *

The authors would like to thank the colleagues working for the Italian Regional Operations Center, whose joint effort made this work possible.

REFERENCES

[1] Laure E. and Jones B., *Enabling Grids for e-Science: The EGEE Project*, EGEE-PUB-2009-001, 2009 (`http://www-eu-egee.org`).
[2] *Release Notes 3: Corresponing to the full Production Grid Infrastructure Operational*, EGEE II Deliverable DSA1.9 (`https://edms.cern.ch/document/489465`) Feb. 2006.
[3] *INFN Grid Release*, `http://igrelease.forge.cnaf.infn.it/doku.php?id=start`.
[4] Andreetto P., Borgia S., Dorigo A. *et al.*, *Practical Approaches to Grid Workload and Resource Management in the EGEE Project*, *Proceedings of the Conference on Computing in High Energy and Nuclear Physics - CHEP 2004, Interlaken (CH), Sept. 2004*. Vol. **2**, pp. 899-902.
[5] Badino P., Brito da Rocha R., Casey J. *et al.*, *The gLite File Transfer Service*, EGEE-PUB-2006-023, 2006.
[6] Piro R. M., Guarise A. and Werbrouck A., *An Economy-based Accounting Infrastructure for the DataGrid*, *Proceedings of the 4th International Workshop on Grid Computing, GRID 2003, USA, Nov 2003* (IEEE Computer Society Washington, DC, USA) 2003. ISBN: 0-7695-2026-X.
[7] Dal Pra S., Fattibene E., Misurelli G. *et al.*, *HLRmon: a Role-based Grid Accounting Report Web Tool*, *J. Phys. Conf. Ser.*, **119** (2008) 052012.
[8] *Global Grid User Support*, `https://gus.fzk.de/pages/home.php`.
[9] *The Italian Grid Initiative*, `http://www.italiangrid.org/`.