ENEA-GRID: a production quality grid infrastructure

S. Migliori, G. Bracco, R. Guadagni, S. Taglienti

ENEA INFO [Servizio Centralizzato Informatica e Reti], Roma

Outline

• ENEA-GRID Infrastructure and Resources
• Case studies: numerical simulations and support to experimental facilities
• GRID paradigm & ENEA-GRID
• Connection with other GRID projects

12 research centres in Italy

An ITC and Network Department (INFO)

6 Computer Centres
Casaccia, Frascati, Bologna, Trisaia, Portici, Brindisi

Multiplatform resources for serial & parallel computation and graphical post-processing.

Other computer resources in ENEA: departments & individuals
**ENEA GRID**

**ENEA-GRID** has been in development since 1998 and is now an integrated infrastructure:

- offering a *production quality, service oriented* system
- covering most of the computational resources belonging to ENEA INFO Scientific Computation Service, connected in a *Wide Area Network*.
- providing a *unified user environment and an homogeneous access method* for all ENEA researchers irrespective of their location
- implementing tools to facilitate the integration of departments and individual resources and support to experimental facilities
ENEA GRID Services

Services for research activity in ENEA

- **Multiplatform Parallel systems:**
  - AIX, Linux (Alpha,x86), IRIX, [Unicos (Cray)]

- **Graphical simulation & code result post-processing:**
  - SGI + 3D immersive facilities

- **Software resources:**
  - Commercial Codes: Fluent, Gambit, Abacus, Catia, Ansys
  - Research codes: mcpn/x, eranos, fluka,…
  - Elaboration environments: IDL, Matlab, Mathematica, SAS
  - Windows applications
<table>
<thead>
<tr>
<th>Platform OS</th>
<th>#cpu</th>
<th>Gflops</th>
<th>Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIX</td>
<td>92</td>
<td>275</td>
<td>Frascati(66), Bologna(8), Portici(18)</td>
</tr>
<tr>
<td>Linux x86</td>
<td>54</td>
<td>100</td>
<td>Frascati(28), Casaccia(14), Portici(4), Trisaia(4), Brindisi(4)</td>
</tr>
<tr>
<td>Linux Alpha</td>
<td>80</td>
<td>100</td>
<td>Casaccia</td>
</tr>
<tr>
<td>IRIX</td>
<td>26</td>
<td>40</td>
<td>Frascati(8), Casaccia(4), Portici(1), Trisaia(8), Brindisi(1), Bologna(5)</td>
</tr>
<tr>
<td>Solaris</td>
<td>8</td>
<td>10</td>
<td>Trisaia(4), Casaccia(2), Bologna(2)</td>
</tr>
<tr>
<td>Windows</td>
<td>18</td>
<td>50</td>
<td>Frascati(6), Portici(4), Trisaia(4), Brindisi(4)</td>
</tr>
<tr>
<td>Unicos[Cray]</td>
<td>16</td>
<td>19</td>
<td>Casaccia (SV1/1A)</td>
</tr>
</tbody>
</table>

Portici, Brindisi & Trisaia resources were recently increased in the framework of TELEGRID Project [PON]
ENEA GRID Graphics and 3D resources

SGI systems are coupled with 3D immersive facilities together with 3D modelling and rendering software as CAVELIB, DV-Mockup, AVS

C.R. Frascati

C.R. Casaccia

C.R. Bologna
ENEA Network connection

ENEA computational resources are distributed over WAN, connected by GARR, the Italian Academic & Research Network

ENEA-GARR
9 PoP, 4-32 Mbps

Brindisi
Bologna
Casaccia
Frascati
Portici
Trisaia
Palermo
Pisa
Roma Sede
ENEA GRID ARCHITECTURE

A choice of **mature components** for reliability and easiness of support and maintenance:

- Distributed File System: **AFS**
- Job and resources manager: **LSF Multicluster**
- Unified GUI access: **Java and Citrix Technologies**
- Quality monitoring system: **Patrol**
- Licence Servers

Integration with department and individual resources

- Distributed File System: AFS for software and data
- Licence pool sharing

Possible integration with other institutions
AFS: enea.it cell


A mature Client/Server Architecture [scalable, multiplatform].

Optimized for WAN operation: native data replication on server, data cache transparent to the user on client.

Security: Kerberos authentication, native authorization management.

UNIFORM USER ENVIRONMENT

The user HOME directory is the same on all UNIX/Linux platform and it’s also accessible by Windows system.

UNIFORM SOFTWARE DISTRIBUTION

Easy multiplatform management of WAN software installation.
LSF

Load Sharing Facility MultiCluster version
[www.platform.com]

Mature and supported multiplatform environment.

High scalability (ENEA GRID > 200 cpu)

WAN enabled: multiclient.

Certified by main vendors (IBM, SGI,..) and integrated with commercial software.

Integration with AFS.

**ENEA GRID 6 LSF clusters:**

Frascati, Bologna, Casaccia, Trisaia, Portici, Brindisi
CITRIX (1)

Citrix Metaframe, now Citrix Presentation Center, has been used in ENEA-GRID to provide a unified GUI to all resources.

Citrix is a leader company in providing multiplatform access to remote computer resources using a client/server approach.

Citrix technology assures WAN optimized performances and secure communication methods.

**ENEA GRID access** is managed by 6 Citrix UNIX servers, one for each of the main computer centers: Bologna, Casaccia, Frascati, Trisaia, Portici, Brindisi. 4 Windows Citrix servers provide access to Windows applications (Frascati, Portici, Trisaia, Brindisi).

Citrix technology is also used to provide Web access to ENEA GRID offering a user/application oriented environment.
ENEA GRID makes use of Citrix Metaframe to publish an application providing all the available resources with a unified GUI interface.

Application components:
- Java (GUI)
- shell scripts
- EnginFrame
CITRIX: ENEA GRID Web Access

http://www.afs.enea.it
ENEA GRID Architecture

- AFS Geographical multiplatform File System
- AIX
- SGI
- SUN
- HP
- Win
- Linux
- Linux Alpha
- ...

- LSF
- LSF multi-cluster as integrator
- User programs and commercial codes
- Graphical User Interface
- Web Portal Citrix

- Monitoring service quality
- Load Leveler
- RMS
GRID usage, examples and some case studies

650 registered users
2 TB stored software/data out of 4 TB available in enea.it AFS cell

Examples and case studies

• Running FLUENT on ENEA GRID
• Support to department experimental research:
  • Biotechnology
  • Electron microscopy
  • Nuclear fusion research
FLUENT (CFD) code batch submission

A FLUENT case can be run from a dedicated GUI for submission

LSF Options
-o filename.%J - output file
-w "done(idjob)" - start after idjob
-u e-mail userid - output by E-mail
-b begintime - job start time
....
Monitoring job status by xlsbatch (LSF)

xlsbatch information
Jobid
Users
Job status
Queues
Submission host
Run host
Submission time
Submission command

Job status:
Monitoring host status by xlsmon (LSF)

Status of sp3-1 host using xlsmon utility
3D CFD Simulation on Aircraft Axial Turbin Stage with Wake & Vortex Shedding Analysis on Turbine Blades [FLUENT]

Temperature contours on Stator blades & Cp experimental data comparison
Secondary flow prevision and 3D Path Lines analysis

Mesh: 1 million cells
CPU Time: 6000 hours
Elapsed Time: 525 hours on 12 IBM SP3 platforms

Large Eddy Simulation (LES) provides high resonable results on wake analysis and turbine stage performance.

Temperature & wakes interactions
CFD Analysis of a fluid dumper [FLUENT]

Mesh: 600,000 cells
CPU Time: 960 hours
Elapsed Time: 130 hours
on 8 Linux platforms

Punto di progetto (apertura valvola)

Caratteristica parabolica dello smorzatore con valvola aperta fino a 3 mm

Primary Flux
Secondary flux
Vortex stagnation

Pressure losses

Vortex stagnation
Secondary flux
Primary Flux

Mesh: 600,000 cells
CPU Time: 960 hours
Elapsed Time: 130 hours
on 8 Linux platforms

CFD Analysis of a fluid dumper [FLUENT]
FLUENT performance on different platforms in ENEA GRID

FLUENT Performance Rating on ENEA’s Grid Computing

<table>
<thead>
<tr>
<th>Platform</th>
<th>#CPU</th>
<th>GHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIX</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sp3-1</td>
<td>16</td>
<td>.375</td>
</tr>
<tr>
<td>Sp4-1</td>
<td>32</td>
<td>1.1</td>
</tr>
<tr>
<td>Ostro</td>
<td>16</td>
<td>1.3</td>
</tr>
<tr>
<td>Pace</td>
<td>8</td>
<td>1.3</td>
</tr>
<tr>
<td>Linux</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C..03</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>bw</td>
<td>8</td>
<td>1.8</td>
</tr>
<tr>
<td>IRIX</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Onyx</td>
<td>8</td>
<td>0.4</td>
</tr>
</tbody>
</table>

Cpu time normalized to the case: Sp4-1, 2 CPUs

Scaling with #CPU & platform type
ENEA GRID and experimental facilities

DNA Sequence system
(ABI Prism 3700)
Trisaia

DB_1
ENEA GRID
CPUS
WEB
ICA
SSH

DB_2

DB_3

Electronic Microscope
(Brindisi)
300 Kev (sept. 2004)

Controlled Nuclear Fusion:
FTU
Frascati Tokamak Upgrade
Video Acquisition

ENEA-GRID, Incontro GARR, Roma 15/7/2004
ENEA GRID: FTU video acquisition data
ENEA GRID & GRID PARADIGM

“GRID concept: coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations (VO)”


ENEA GRID infrastructure has been designed to provide a multisite resource sharing inside ENEA, both to optimize the utilization of Central Computer Service (ENEA INFO) resources and to integrate the other resources available inside the organization.

VO concept can not be fully developed inside an unique organization but similarities can be found with Project Groups sharing specific resources and ENEA GRID components can provide support to this (AFS groups, LSF resources).

There are cases where ENEA GRID infrastructure has been used to support the collaboration with other institutions (+ other GRID projects)
• DATA sharing
• Applications in ENEA GRID
• Sharing Working Sessions (Citrix)
ENEA GRID & GRID PARADIGM

ENEA GRID architecture answers to many of the challenges posed by the “GRID Problem” : “unique authentication, authorization, resource access and resource discovery”, Foster & Kesselman 2001

The choice of mature components (LSF & Citrix - proprietary, AFS - opensource project since 2000) reduce the generality of ENEA GRID as a GRID model but improves greatly the reliability and the easiness of update & management: production quality infrastructure.
ENEA GRID EXPERIENCE

Some remarks can be drawn from the analysis of the operation of ENEA GRID infrastructure in the last years.

User acceptance of GRID working mode requires a new cultural approach to the access to computational resources and depends strongly on the service quality:

- “Value” of the accessed resources
- Efficient networking and system faults monitoring
- Support of user customization for GRID environment

The administration of a distributed & integrated resource system is a technical, cultural and organizational challenge in a complex structure such as ENEA.
ENEA GRID improvements in progress

Increasing resources: AIX +32 SP nodes/Frascati, Cluster 16 AMD 64bit/Casaccia, Cluster 8 G5/Trisaia, 2 CPU AIX/Brindisi (TELEGRID Project).

Extending Windows services.

AFS optimization.

Studies in collaboration with CASPUR to manage the replication of large amount of data over WAN with better performances than AFS.

Participation in CASPUR storage technology tracking activity (SAN systems and file systems testing and evaluation, AFS implementation for parallel systems [IPP Garching] => DEISA Grid Project)
Connection with other GRID Projects

Each GRID project is characterized by its own architectural design and middleware.

The focus of the participation of ENEA INFO in other GRID projects is **GRID interoperability**

Access to ENEA GRID through other GRID models should be compatible with ENEA GRID architecture: **gateway implementation**.

Current GRID projects:

- **EGEE** [Datagrid]
- **FIRB/Grid.it** : WP13, Computational Chemistry (CHEMGRID), Perugia Univ., ENEA Casaccia, et al. ENEA activity starting July 2004 at Frascati
ENEA GRID and DATAGRID PROJECT

ENEA participated (unfunded) in DATAGRID project (ended 3/2004) through the collaboration with ESA, ESRIN Laboratory in Frascati.

A gateway between DATAGRID infrastructure and ENEA-GRID has been set-up at ESRIN enabling the submission of jobs from DATAGRID to ENEA-GRID.

The gateway has been tested running the analysis of data from the GOME satellite (Ozone level) using IDL codes which can be easily run in a multiplatform environment.

The gateway provided a limited integration between the different GRID environment: EGEE should attain a full integration.
ENEA GRID/DATAGRID Test case: GOME data analysis
## ENEA GRID/DATAGRID: GOME run cases in ENEA

<table>
<thead>
<tr>
<th>User</th>
<th>Job</th>
<th>Queue</th>
<th>Parameter</th>
<th>Options</th>
<th>Start Time</th>
<th>Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
<tr>
<td>maglioci</td>
<td>normal</td>
<td>grid0007</td>
<td>esrin</td>
<td>enea</td>
<td>Jun 9</td>
<td>22:19:02</td>
</tr>
</tbody>
</table>
ENEA GRID/DATAGRID: GOME results on Web
ENEA-GRID and EGEE Project

EGEE (Enabling GRID for e-science in Europe) is a project in the 6th EU framework program, with the mission to deliver production level GRID services [www.eu-egee.org].

ENEA is one of the funded partners in the project, among the Italian participation which is coordinated by INFN.

EGEE project started recently (April 2004); a kick-off Meeting was held in Cork, April 2004.

The installation of a small testing environment (7 Linux boxes) with standard EGEE middleware is currently in progress at ENEA Frascati.

The development of a gateway to ENEA-GRID will follow.

Enea commitment: ~100 cpu (at 20%), second half 2005.