EUChinaGRID and EUMEDGRID Projects

Federico Ruggieri – INFN

YUINFO2007
Kopaonik 14 March 2007
EGEE and collaborating projects

European Commission co-funded projects
Projects with other funding
### Specific Support Actions

<table>
<thead>
<tr>
<th>Project</th>
<th>euchina\textsuperscript{grid}</th>
<th>EU MED Grid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start Date</td>
<td>1 Jan 2006</td>
<td>1 Jan 2006</td>
</tr>
<tr>
<td>Duration</td>
<td>24 Months</td>
<td>24 Months</td>
</tr>
<tr>
<td>EU contribution</td>
<td>(\sim 1.3 \text{ M}\text{\euro})</td>
<td>1.645 M\text{\euro})</td>
</tr>
<tr>
<td>Partners</td>
<td>Third Parties</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>--------------</td>
<td></td>
</tr>
<tr>
<td>INFN (IT) Coordinator</td>
<td>Università di Roma Tre – Dipartimento di Fisica (IT)</td>
<td></td>
</tr>
<tr>
<td>CERN (CH)</td>
<td>Academia Sinica Grid Computing Centre (ASGC), Taipei</td>
<td></td>
</tr>
<tr>
<td>Università di Roma Tre - Dipartimento di Biologia (IT)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consortium GARR (IT)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRNET (GR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jagiellonian University – Medical College (PL)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beihang University (CN)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computer Network Information Centre - Chinese Academy of Sciences (CN)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Institute of High Energy Physics – Chinese Academy of Sciences (CN)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peking University (CN)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
EUChinaGRID main objectives

- Support the interconnection and interoperability of Grids between Europe and China.

- Dissemination of advanced knowledge in Grid technology is also a relevant part of the activity.

- Strengthening the collaboration between scientific groups in both regions, supporting existing and project specific Grid applications.
TEIN2 & ORIENT
Monitoring (1)

GridICE >> Geo

- Poland
- GRNet
- INFN Catania, CNAF, Roma3
- CNIC
- IHEP
Monitoring (2)

GridICE >> Site::ALL

<table>
<thead>
<tr>
<th>Site</th>
<th>Region</th>
<th>GK #</th>
<th>O #</th>
<th>RunLab</th>
<th>WaitLab</th>
<th>JobLoad</th>
<th>Power</th>
<th>WN #</th>
<th>CPU #</th>
<th>CPULoad</th>
<th>Available</th>
<th>Total</th>
<th>%</th>
<th>NH #</th>
</tr>
</thead>
<tbody>
<tr>
<td>BEIJING-CNIC-LCG2-IA64</td>
<td>CERN</td>
<td>1</td>
<td>6</td>
<td>17</td>
<td>0</td>
<td>32</td>
<td>8</td>
<td>5</td>
<td>4</td>
<td>2%</td>
<td>21.3 GB</td>
<td>62.8 GB</td>
<td>66%</td>
<td>11</td>
</tr>
<tr>
<td>BEIJING-LCG2</td>
<td>CERN</td>
<td>1</td>
<td>0</td>
<td>30</td>
<td>93</td>
<td>35</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>2%</td>
<td>1.3 TB</td>
<td>2.1 TB</td>
<td>65%</td>
<td>13</td>
</tr>
<tr>
<td>CERN-PPOD</td>
<td>CERN</td>
<td>2</td>
<td>28</td>
<td>3544</td>
<td>1104</td>
<td>3544</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>2%</td>
<td>654.7 GB</td>
<td>1.7 TB</td>
<td>65%</td>
<td></td>
</tr>
<tr>
<td>CYFRONET-IA64</td>
<td>CentralEu</td>
<td>1</td>
<td>12</td>
<td>21</td>
<td>40</td>
<td>12</td>
<td>0</td>
<td>17</td>
<td>24</td>
<td>65%</td>
<td>335.2 GB</td>
<td>2 TB</td>
<td>64%</td>
<td>1</td>
</tr>
<tr>
<td>CYFRONET-LCG2</td>
<td>CentralEu</td>
<td>1</td>
<td>16</td>
<td>231</td>
<td>346</td>
<td>231</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>15.6 TB</td>
<td>20.6 TB</td>
<td>64%</td>
<td>2</td>
</tr>
<tr>
<td>GR-01-AUTH</td>
<td>SEE</td>
<td>1</td>
<td>12</td>
<td>2</td>
<td>317</td>
<td>2</td>
<td>0</td>
<td>8</td>
<td>12</td>
<td>0%</td>
<td>165.9 GB</td>
<td>217.6 GB</td>
<td>24%</td>
<td>12</td>
</tr>
<tr>
<td>INFN-CATANIA</td>
<td>Italy</td>
<td>1</td>
<td>0</td>
<td>107</td>
<td>60</td>
<td>107</td>
<td>98K</td>
<td>37</td>
<td>70</td>
<td>0%</td>
<td>19.1 TB</td>
<td>21.2 TB</td>
<td>10%</td>
<td>97</td>
</tr>
<tr>
<td>INFN-CNAF</td>
<td>Italy</td>
<td>2</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>10K</td>
<td>5</td>
<td>10</td>
<td>0%</td>
<td>591.1 GB</td>
<td>1.7 TB</td>
<td>0%</td>
<td>10</td>
</tr>
<tr>
<td>INFN-ROMA3</td>
<td>Italy</td>
<td>1</td>
<td>5</td>
<td>29</td>
<td>0</td>
<td>5</td>
<td>62K</td>
<td>48</td>
<td>24</td>
<td>0%</td>
<td>632.5 GB</td>
<td>699.9 GB</td>
<td>2%</td>
<td>27</td>
</tr>
<tr>
<td><strong>TOTAL: 9</strong></td>
<td></td>
<td>5</td>
<td>4</td>
<td>11</td>
<td>101</td>
<td>4181</td>
<td>1909</td>
<td>38.4 TB</td>
<td>50.4 TB</td>
<td>2%</td>
<td>173</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Main Achievements

- A pilot infrastructure using gLite is already working (9 sites).
- A gateway between gLite (EGEE) and GOS (CNGrid) has been studied and a first implementation is available.
- IPv6 middleware compatibility study was started and first results are available (web site, code checker).
- First batch of applications is running (LHC, WISDOM, Late/Early Stage) and others were tested and are about to be deployed (ARGO Data Mover, Rosetta).
gLite - GOS gateway

Extended GOS forwards batch job to gateway

Scheduler

Batch job

GOS

Thread Pool

Scheduler executes stage in pipeline using idle thread from thread pool

Pipeline for GOS to GLite

Extended LCG-CE forwards batch jobs to gateway

Batch job

Pipeline for GLite to GOS

Different colors in pipeline stand for different stages (StageIn, StageOut, ...)

Each pipeline has corresponding thread pool and scheduler

Extended GOS forwards batch job to gateway

Thread Pool

Idle thread pool used in scheduler

Batch job

WMProxy
Grid Middleware and IPv6

- China is deploying the largest production IPv6 network in the world.
- IPv6 is the natural choice for new generation IP telephony in convergence with Wireless Networking.
- Grid services should be able to run on the future IPv6 enabled PDA’s, portable phones, etc.
- A study is under way on Middleware and high level services compatibility with IPv6 -> operational tests of functionality.
- CNGrid MW (GOS) is mostly based on Java and has now an IPv6 compliant version.
Welcome to the EUCHinaGRID IPv6 web site.

Here you can find all the information and the documentation about our IPv6 study on the available and foreseen network connectivity to promote new high bandwidth links between Europe and China or Asia in general and to study the available Grid Middleware for an IPv6 network and the interaction between Grid Services and IPv4-IPv6 communication.
IPv6 compatibility of Grid middleware

- A specific task was started in order to investigate if libraries, software and other third-party elements used in gLite and GOS are IPv6 compliant.
- A compatibility list is available and constantly enlarged and updated.

<table>
<thead>
<tr>
<th>Component</th>
<th>Middleware</th>
<th>Internal related documents</th>
<th>Official compatibility</th>
<th>Compatibility tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Globus-toolkit</td>
<td>IPv6 compatibility of GLOBUS</td>
<td>Yes (note1)</td>
<td></td>
<td>TBD</td>
</tr>
<tr>
<td>Condor</td>
<td>IPv6 compatibility of Condor</td>
<td>No (note 2)</td>
<td></td>
<td>TBD</td>
</tr>
<tr>
<td>LSF</td>
<td>IPv6 compatibility of LSF</td>
<td>Yes (note 3)</td>
<td></td>
<td>TBD</td>
</tr>
<tr>
<td>PBS</td>
<td>IPv6 compatibility of PBS</td>
<td>No</td>
<td></td>
<td>TBD</td>
</tr>
<tr>
<td>Tomcat</td>
<td>IPv6 compatibility of Tomcat</td>
<td>Yes</td>
<td></td>
<td>Not needed</td>
</tr>
<tr>
<td>Java</td>
<td>IPv6 compatibility of Java</td>
<td>Yes</td>
<td></td>
<td>Not needed</td>
</tr>
</tbody>
</table>
Applications

EGEE
ATLAS and CMS support

ARGO
Data mover
MEDEA++
Corsika

NBP Application
Rosetta
Early/Late Stage
EGEE applications

- Deployment of CMS and ATLAS applications on EUChinaGRID pilot infrastructure
ARGO – YBJ Laboratory

- Unique High Altitude Cosmic Ray Laboratory (4300 m) Tibet, 90 km North to Lhasa. Chinese-Italian collaboration.
- The Experiment data rate to be transferred is 250 TB/Year requiring a steady transfer rate of the order of 100 Mbps to Beijing and from there to Italy.
ARGO-YBJ Data Mover

- Reliable Data Transfer from Yangbajing
- 2 Data Repositories (CNAF & IHEP) synchronized via LFC and FTS + custom scripts.
The “never born proteins”

- The number of natural proteins is just an infinitesimal fraction of the theoretically possible ones (Ex., a hydrogen atom vs. the total mass of the universe).
- There exists an astronomical number of “never born proteins”.
  - Their study can teach us lessons on the principles of protein folding.
  - Have natural proteins special physico-chemical properties?
  - Are there “never born proteins” potentially useful for biotechnological or biomedical purposes?
Biological applications

- Generate “never born protein” sequences database
- Port protein structure prediction software on grid
- Develop web tools for non grid trained users

KWCWPFAHNDLKVSQ
WYVEPPDTIPPYNKYGTN
FIKHLCQYIAHMQGTHFF
NVRMHLQWKIIVDCAA

Rosetta
Early/Late Stage
Future work

- Continue in the collaboration with other projects, e.g.:
  - Joint conference with ECHOGRID (Beijing 24-25 April).
  - Cooperation with EGEE and ETICS on IPv6 compliance and interoperability of middleware.

- Extend the actual basic functionality of the gLite – GOS gateway.

- Expand the pilot infrastructure:
  - Peking University is going to join in these days;
  - CNGrid nodes will be accessible through the gateway.
Empowering eScience across the Mediterranean

EUMEDGRID

FP6-2004-Infrastructures-6-SSA-026024

www.eumedgrid.org
<table>
<thead>
<tr>
<th>Partners</th>
<th>Third Parties</th>
</tr>
</thead>
<tbody>
<tr>
<td>INFN (Italy) Coordinator</td>
<td>Università di Messina – Dipartimento di Matematica (IT)</td>
</tr>
<tr>
<td>CERN (Switzerland)</td>
<td>Dipartimento di Fisica – Università di RomaTre (IT)</td>
</tr>
<tr>
<td>CYNET (Cyprus)</td>
<td>ICTP – (IT)</td>
</tr>
<tr>
<td>DANTE (UK)</td>
<td>CRS4 – (IT)</td>
</tr>
<tr>
<td>GARR (Italy)</td>
<td>IUCC (Inter University Computing Center) – ISRAEL</td>
</tr>
<tr>
<td>GRNET (Greece)</td>
<td>Institute of Communication and Computer Systems, Athens, Greece</td>
</tr>
<tr>
<td>RED.ES (Spain)</td>
<td>JUNet, Jordan</td>
</tr>
<tr>
<td>University of Malta (Malta)</td>
<td>PADI2, Palestine</td>
</tr>
<tr>
<td>CERIST (Algeria)</td>
<td>Laboratoire CRISTAL - Ecole Nationale des Sciences de l'Informatique (TU)</td>
</tr>
<tr>
<td>CNRST (Morocco)</td>
<td>Research Unit of Technologies of Information and Communication - University of</td>
</tr>
<tr>
<td>EUN (Egypt)</td>
<td>Tunis (TU)</td>
</tr>
<tr>
<td>HIAST (Syria)</td>
<td>Bilkent University (TK)</td>
</tr>
<tr>
<td>MRSTDC (Tunisia)</td>
<td></td>
</tr>
<tr>
<td>TUBITAK-ULAKBIM (Turkey)</td>
<td></td>
</tr>
</tbody>
</table>
Follows the positive experience of SEE-GRID and applies it to the MED area.

It’s about People: The core of the EUMEDGRID approach is to establish a human network in the eScience area, enlarge and train this community, and establish a pilot Grid infrastructure supporting proof of concept regional applications.

It’s about Technology: EUMEDGRID aims to provide specific support actions to assist the participation of the states of the Mediterranean region in the pan-European and worldwide Grid initiatives, thus expanding and supporting the European Research Area (ERA) in the region.
Digital Divide

http://maps.maplecroft.com/
Social Impact

- e-Infrastructures support wide geographically distributed communities → enhance international collaboration of scientists → promote collaboration in other fields.
- Grids and networks allow the access of many researchers to Science (laboratories and data) → the brain drain can be reduced.
- The European Research Area program wants to set Europe as the most advanced region in e-Infrastructures and promote the development of less advanced countries to alleviate as much as possible the digital divide.
EUMEDCONNECT Network

FP6-2004-Infrastructures-6-SSA-026024

EUMEDgrid

Backbone Topology
October 2006

Creating a research and education network for the Mediterranean
Strategies

(TOP-DOWN)
- Foster the creation of National Grid Initiatives (NGIs) in the Med Countries → promote the creation of sustainable national grid infrastructures with production quality level.
- Deploy first those applications which are already grid-enabled and/or easy to be ported (LHC, WISDOM).

(BOTTOM-UP)
- Discover and support New User Communities which will be the driving force of constant expansion and evolution of the Grid Infrastructures → stimulate new demanding application to be deployed and support NGIs creation.
- Discover and support new applications which are relevant for the Region (Earth Science, Archaeology, etc.).
This assessment questionnaire must be filled in for each application. It aims to gather relevant information to evaluate the suitability of the application to the EUMEDGRID environment and allow for a better infrastructure support.

**Users and Institutions**

1. First Name and Last Name
2. Institutions

[https://secure.um.edu.mt/eumedgrid/questionnaire/wp4/](https://secure.um.edu.mt/eumedgrid/questionnaire/wp4/)
First EUMEDGRID School for Application Porting (EGSAP-1)

- Date: 16-27 April 2007
- Venue: Cairo (Egypt)
- based both on lectures and hands-on practices.
- the number of applications to be ported will be limited.
- Applications proposals had to be submitted before the 11th of March 2007.
- Selection is under
Already Deployed Applications

- **CODESA-3D (Earth Science)**
  - It is a three-dimensional finite element simulator for coupled flow and density dependent transport in variably saturated porous media.
  - Requires data from: Geology, Topography, Meteorology, Water extraction, Aquifer properties, Soil maps, Land use.

- **ArchaeoGrid (Multi-disciplinar)**
  - Paleoclimate simulation (with modern tools like MM5)
  - Aims at studying the origin of the City in the Mediterranean Region between X-VIII Centuries B.C.
Grid-enabled Subsurface Hydrology application in the Mediterranean area

- Sustainable management of groundwater exploitation using Monte Carlo simulation of seawater intrusion in the Korba aquifer (Tunisia).
- Seawater intrusion in the aquifer is a very severe problem in some of the coastal zones of the Med Area.
CODESA-3D is a three-dimensional finite element simulator for coupled flow and density dependent transport in variably saturated porous media.
WHERE
Archaeological data are geospatial data
Archaeological GIS

Best resolution
by Differential GPS
X, Y < 1 m
Z < 0.1 m (SAR)

Aerial photo or satellite image

-best resolution
X, Y, Z ~ few cm

DEM

Infrastructures

Land use

Archaeological sites

How many layers?
Number of layers large, depending from the complexity of the archaeological problem
WHEN

Archaeological data must be ordered by time

Archaeological GIS + time

1000 B.C.

500 B.C.

500 A.D.

~20 years, most precise resolution in time
EUMEDGRID supported the WISDOM data challenge on its infrastructure.

WISDOM is an initiative for Grid-enabled drug discovery against neglected and emergent diseases, using GRIDs for in Silico Drug discovery.

Simulation of docking of compounds on proteins.
Perspectives

- Increase site availability and usability.
- Strengthen and enlarge the collaboration with the Universities and Research Institutes of the Med Countries.
- Increase the number of NGIs.
- Discover new communities and interesting (for the area concerned) applications.
Conclusions

- EUChinaGRID and EUMEDGRID are aiming to promote the extension and interoperation of the Grid Infrastructures outside EU.
- EUChinaGRID is more “Infrastructural and Technical”.
- EUMEDGRID is focused on Policy, promotion of Grids creation of NGIs and operations startup.
- Both are exploiting the existing and foreseen EU co-funded network infrastructures (EUMEDCONNECT, TEIN2/ORIENT, Geant2).
- Several applications have been selected and deployed together with the EGEE supported ones.
Thank You!