g-Eclipse

A Framework for Accessing Grid Infrastructures

Nicholas Loulloudes
Trainer, University of Cyprus
(loulloudes.n_AT_cs.ucy.ac.cy)
Outline

• Grid Reality – The Problem
• g-Eclipse - The Idea
• g-Eclipse Framework
• Technical Overview
• Abstracting Grid Middlewares
• gLite Support
• Conclusions
Grid Infrastructures: a huge collection of computational and storage resources.

The EGEE infrastructure in Europe alone has available 24/7:

- 267 Sites (in 54 countries)
- ~114,000 CPUs
- 20 Petabytes of disk space.
- Supporting ~ 15 application domains
Grid Reality - The Problem II

Grid Technology is **complex**:

- Different systems are used:
  - Middleware (gLite, GRIA, Globus, Unicore...)
  - Installation (rpm, tar, Quattor,...)

- Different Programming paradigms exist:
  - Batch Type systems Vs. Service Oriented Systems
  - Many programming Languages.

- Interaction with the Grid is mostly done via Command Line Interface (CLI).

The Threshold is too “high” for the Standard user.

Users want easy access to Grids without knowing the details.

Nicholas Loulloudes

Nicosia, May 6th 2009
g-Eclipse – The Idea

• Provide a friendly UI for accessing Grids.

• Provide an **extensible, middleware-independent**, framework for accessing Grids.

• Support the roles of Grid **users**, **developers** and **operators**.

• Provide the necessary tooling to hide the complexity (**wizards, editors, views ...**).

• **Substitute** CLI with GUI.

• Conform to **Grid Standards**.
Single UI for all Middlewares
Perspective - An Eclipse term for the set of tools (views, editors wizards)

**Grid Developer**
- Code Applications
- Compile Apps.
- Debug Apps.
- etc.

**Grid User**
- Data Management
- Job Management
- Visualization
- etc.

**Grid Site Operator**
- Site Administration
- User Administration
- Job Management
- etc.
The g-Eclipse Project

Funded by the g-Eclipse (FP6) EU project

- Duration: July 2006 – December 2008
- Funding: 2 Million Euro
- Consortium members:

Forschungszentrum Karlsruhe in der Helmholtz-Gemeinschaft
Johannes Kepler Universität Linz
INNOOPRACT
POZNAŃ
IT Innovation
NEC

The University of Reading

Nicholas Loulloudes
Nicosia, May 6th 2009
g-Eclipse - An official Eclipse Technology Project
- Around 20 developers among the 8 partners
- 14 members with committer status at Eclipse.org
- Community established and started to grow.

Source code released under Eclipse Public License
- Version 1.0 released in January 2009
- Roughly 70 plug-ins, 4000 classes and interfaces aka 350,000 LOC
- Available for free download
- Supported platforms: Linux, Windows, Mac OS X
**g-Eclipse - Technical Overview**

- **Based on the Eclipse platform**
  - It is modular (OSGi) and extensible (Extension point mechanism of Eclipse).

- **Provide a middleware-independent architecture that:**
  - Abstracts common grid concepts.
  - Provides abstract core functionalities.
  - Is extensible by middleware-specific plugins.

- **Provide a graphical user interface that:**
  - Is based on the abstract core and therefore.
  - Looks and behaves (at least) the same for any middleware.
Abstraction Layer
- Core functionalities, e.g.
  - Grid Authentication / Authorization
  - Virtual Organization management
  - Data Management
  - Job submission
- Common User Interface
  - Views
  - Wizards
  - Editors
  - Dialogs
  - Preference Pages

Implementation Layer
- Extended core functionalities
- Middleware specific functionalities

g-Eclipse - Technical Overview II

Eclipse Platform
- Eclipse Extension Point
Four Pillars of Modern Grid Middleware

1. Security
   - Authentication
   - Authorization
   - Delegation
   - Encryption

2. Data Management
   - Transfer Protocol
   - Replicas
   - Access Control
   - File Catalogue

3. Job Management
   - Job Submission
   - Job Scheduling
   - Resource Broker
   - Job Status

4. Information Service
   - Discovery
   - Monitoring
   - Caching
   - Query

Nicholas Loulloudes
Nicosia, May 6th 2009
Abstracting Grid Middleware

Core Architecture / Abstraction Layer

Security Framework
- Authentication
- Authorization
- Access Control
- Token-Management
- Certificate-Management

Grid Model
- Local Resources
- Remote Resources
- Computing
- Storage
- Services
- Job-Management
- Data-Management
- VO-Management

Nicholas Loulloudes
Nicosia, May 6th 2009
Supporting the gLite Middleware
• **AAI support**
  - Globus proxy creation and management
  - VOMS proxy creation and management
  - CA certificate management

• **Data Management**
  - GSIFTP
  - SRM
  - LFC
  - Transfer manager (resume, restart, ... transfers)
  - 3rd party transfers (server-2-server)
• **Information Services**
  - Support for BDII
  - Query available resources for a dedicated VO
    • Build the personalized Grid
  - Uses the GLUE schema to store information.
  - Provides methods to:
    • Fetch all available information or
    • Query the server side for specific information
• **Job Description Languages**
  – JSDL and JDL supported
  – Parametric jobs supported
  – Workflows supported

• **Job Submission services**
  – WMS
  – Cream

• **Job Monitoring**
  – Logging & Bookeeping (L&B)
gLite - User Role Mappings

g-Eclipse

User Perspective

Operator Perspective

Developer Perspective

Nicholas Loulloudes

Nicosia, May 6th 2009
The User Perspective

<table>
<thead>
<tr>
<th>Workbench</th>
</tr>
</thead>
</table>

### Information View

- g-Eclipse (User) - testGridProjectJob Descriptions/submitSDL.jSDL - Eclipse Platform

### Grid Project View

- Worlds
- LMsystems
- jobs
- Workflows
- gcclipse
- Computing
  - aol-104-128.gridka.de:2
  - ce-fzggrid.de:2119job
  - ce.egee.man.poznan.pl:
  - cel01.grid.ucy.ac.cy:2119
  - egee-cel1.gup.uni-linz.ac.at:

### Authentication View

- Jobs
- Workflows
- gcclipse

### JSDL Editor View

<table>
<thead>
<tr>
<th>Workbench</th>
</tr>
</thead>
</table>

### Job Definition Page

- Application Name: gnuplot
- Executable: /usr/local/bin/gnuplot
- Standard input file: input.dat
- Standard output file: output1.png
- Standard error file: output1.png

### Resources Page

- Resources: This element describes the resource requirements of the job.

### Application Page

- Application: This element describes the Application and its requirements.
  - Application: This element describes a POSIX style Application and its requirements.

### Data Staging Page

- DataStaging: This element defines the files that should be moved to the execution host (stage in) and the files that should be moved from the execution host (stage out).

### Overview

<table>
<thead>
<tr>
<th>Job Definition</th>
<th>Application</th>
<th>Data Staging</th>
<th>Resources</th>
<th>SubmitSDL.jSDL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connections</td>
<td>Job Details</td>
<td>Jobs</td>
<td>Auth Tokens</td>
<td>Properties</td>
</tr>
<tr>
<td>ID</td>
<td>State</td>
<td></td>
<td></td>
<td>Time Left</td>
</tr>
<tr>
<td>VOMS Proxy</td>
<td>VOMS Proxy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>01: 31.59.59s</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Nicosia, May 6th 2009
• Everything revolves around the Grid Project.

• A Grid Project is created as any other project in Eclipse.

• The Grid Project provides a view of:
  • Mounted Remote File Systems
  • Job Description
  • Submitted Jobs
  • Workflows
  • VO-specific Resource Browser.
The User Perspective III

- The Job Description Wizard helps users to prepare a job description quickly.

- The Wizard can be enhanced with application defined pages

- Other job description languages can also be implemented (e.g. an RSL editor is available)

- The JSDL Editor allows users to edit JSDL files with a user-friendly, multipage, form-based editor
The Operator Perspective contains tools for Grid operator daily activities

- Management of Computing Elements and Worker Nodes.
- Queues management.
- Virtual Organizations management.
- Monitoring / Testing / Benchmarking various services.
- CLI is also available.
The Operator Perspective II
The Operator Perspective III

- **The Batch UI Framework**
  - Support for the Portable Batch System (PBS)
  - Start / Stop / Drain queues.
  - Hold / Un-hold / Move Jobs

- **Queue Configuration**
  - MPE for editing queue configuration files (QDL).
  - Wizard for applying QDL files to Batch services

- **Tests Framework**
  - A Framework to test physical Grid Resources and investigate causes of failure.
The Developer Perspective contains tools for grid developer activities:

- Develop C/C++ / Java applications
- Debug applications locally or remotely on the Grid.
- Application Deployment.
- Application Monitoring.
Other Middleware support

• **GRIA (Grid Computing, Industry and Commerce)**
  - Fully integrated

• **Globus (Grid Computing, Scientific)**
  - GT2 partly integrated
  - GT4 integration initiated.

• **AWS (Cloud Computing, Amazon)**
  - S3 and EC2 integrated
Other Middleware support

- Support of gLite, GRIA, Condor.

Prove of g-Eclipse Middleware Independence.

- With Amazon EC support:

Prove of g-Eclipse Grid Independence.
• g-Eclipse is and will stay an official Eclipse Project!

• Self-sustainability by the Eclipse Eco-system.

• Most developers are and will stay Committers on the Project!

• Contributions from other projects started (e.g DORII).
EVERYBODY can contribute ...

YOU can contribute ...

- Report bugs or request features:
  - http://bugs.eclipse.org

- Contact the team:
  - geclipse-user@eclipse.org
  - geclipse-dev@eclipse.org
  - contact@geclipse.eu

- Contribute source code:
  - Bug fixes
  - Improvements
  - New features
!!! Thank you for your attention !!!

For more information visit:

http://www.geclipse.eu

http://www.eclipse.org/geclipse