Eclipse Webinar:

Access a worldwide Grid Infrastructure

Dr. Harald Kornmayer (NEC Laboratories Europe)
Mathias Stuempert (Forschungszentrum Karlsruhe)

on behalf of the g-Eclipse consortium
Which Grid do we talk about..

• The “Grid Vision” is
  – “To solve together a (scientific) problem, by connecting the distributed resources (of scientists) within different administrative domains dynamically and coordinately with the help of fast networks to build a „virtual computing center/organization”
    (According to Foster/Kesselman)

• When we talk here about Grid infrastructure,
  – we expect that the Grid infrastructure exists
  – This is not about building a Grid infrastructure
Existing Grid infrastructures

And many, many more.....
Grid architecture

<table>
<thead>
<tr>
<th>User Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain specific Services</td>
</tr>
<tr>
<td>High Level Grid Services (Resource Broker, Replica Services)</td>
</tr>
<tr>
<td>Grid Middleware</td>
</tr>
<tr>
<td>Fabrics (computing, storage, network)</td>
</tr>
</tbody>
</table>

i.e. European part of the EGEE infrastructure

g-Eclipse is about the User interface
Grid middleware basics

• Basic functionalities:
  – Security mechanisms
  – Execution of computations
  – Seamless access to data

Analogy: “Operating System” of the Grid
  • Grid resources must run Grid middleware services

• Implementations:
  - gLite
  - GRIA
  - the globus toolkit
  - GEMSS

• Standards are emerging!
Grid components
Grid application

Examples:
• Bio-Med
  – new drugs against malaria

• Financial sector
  – Risk management

• Engineering
  – Automotive
    • Integration of the product design process chain (CAE/CAD/CAT) including external engineering companies, developers and suppliers
  – Aerospace
    • collaborative configuration design of complex products
• Many, many others…
The reality for the Grid user

• Infrastructure for scientists were built in the past years
• Many application domains start using Grid infrastructures
• But…
  – Grid technology is complex
    • Different middleware systems are used
      – gLite, Globus, GRIA, UNICORE, …
  – Different programming paradigms
    • Batch type systems vs. service oriented systems
    • Many programming languages

→ The threshold is too high for the “standard” user!
Grid application life cycle

• In most cases, e-Users (e-Scientists, e-Engineers, e-Stock Traders) have their application(s)
  – Legacy code written in different languages (FORTRAN, C, C++, …)

• e-Users want to collaborate
  – A Virtual Organisation is build around a Virtual Computing Center on existing (and new) infrastructure

• e-Users create Grid projects

• e-Users want to interact with the Grid
  – without knowing all details!!
    (development, deployment, testing, management, …)

• → Tooling is necessary!!
  – Wizards, Editors, …
  – Hide the complexity!!
g-Eclipse – the idea

• Users want easy access to the system

• Users act within different roles
  – Grid applications users
  – Grid resource providers and operators
  – Grid application developers

• Users are middleware agnostic
  → Build a middleware independent framework

• Provide a general UI framework/eco system for the different Grid actors based on a reliable platform
  → (re-)use Eclipse and contribute!
  → gain OS independence (by using JAVA!)
g-Eclipse – the project

- www.geclipse.eu

- Project funded by the European Commission (INFSO-32347)

- 7 partners

- Until autumn of 2008

- www.eclipse.org/geclipse

- Technology project at Eclipse Foundation

- Release 0.5.0 finished at 28th of September 2007

- Release 1.0.0 with stable API scheduled for autumn of 2008
Architecture - Overview

- Abstraction Layer
  - Core functionalities, e.g.
    - Authentication/Authorization
    - VO management
    - Data management
    - Job submission
  - Common user interface, e.g.
    - Views
    - Wizards
    - Dialogs
    - Preference pages
- Implementation Layer
  - Extended core functionalities
  - Middleware specific functionalities
Demo

- Set-up
- VO Declarations
- Create a Grid project
- Data management
- Job management
- Remote Queue management
- Application development
- Visualization

![Diagram of g-Eclipse features]

- submit jobs
- monitor jobs
- check jobs
- organize workflow
- replicate data
- create/delete data
- visualize data
- build workflow
- check resources
- monitor resources
- configure resources
- benchmark resources
- manage VO
- debug/test application
- deploy application
- monitor application
Virtual Organization

- Authentication
- Authorization
- Membership
- Access rights
- Collaboration
- ...

- Based on X.509 certificates

1. has a valid certificate
2. is member of
3. trusts
4. provides resources
5. gets access to the distributed Grid resources
Grid components (II)
Grid Resource Provider

• How can a site A support a new VO with computing resources?
  Set up a queuing system:
  – Old: Know all the details of the queuing system
  – New: Configure the batch system on site with g-Eclipse
    • Set up a VO specific queue
    • Drain queues
    • Manage Cluster nodes
How to “gridify” a Legacy application?

1. Develop them on your local computer as a separate JDT/CDT project
2. After a code change, compile them locally and on a remote Grid resource
3. (if needed debug them locally or on a remote Grid site)
4. Deploy the application

Usage of gLogin introduce some firewall issues!!!
Visualization

• Scientific computing often creates big data files, which must be visualized for further studies
• Usage of Open Source VTK library / GVK library
• 2 possible solutions:
  – VTK: Download the data from the Grid and start the visualization locally on your computer!
  – GVK: Leave the data on the Grid and start a remote visualization service and connect to it!
Manage Complexity

• By providing solution to common problems on Grid infrastructures
  – g-Eclipse provides an extended problem reporting mechanism based on the Eclipse core exception
  – Problems have associated solutions
  – Solutions may be
    • passive: just a descriptive text
    • active: provide an action that helps the user to solve the problem, e.g. open an associated preference page
Roadmap

- g-Eclipse release cycle
  - for assuring the quality a monthly milestone release cycle is applied
Summary

• g-Eclipse can be used to access Grid infrastructures
  – first Release g-Eclipse 0.5.0 supporting the gLite middleware is available since September
  – g-Eclipse will support a second Grid middleware in 2008
    – the middleware independent approach will be proved with the GRIA middleware (www.gria.org)
• g-Eclipse is open for contributions
  – more middleware implementation
  – based on the “Eclipse way”
  – integration of existing tools
How to contribute

• Use our tool and send us feedback!
  – We do it the Eclipse way!
  – Webpage www.eclipse.org/geclipse or www.geclipse.eu
  – Newsgroup
    • http://dev.eclipse.org/newslists/news.eclipse.technology.g-eclipse/
  – Developer mailing list
    • https://dev.eclipse.org/mailman/listinfo/geclipse-dev
  – Bugzilla
    • https://bugs.eclipse.org/bugs

• Collaborate and provide patches for other middleware systems
  – Contact {at} geclipse.eu