A Simplified Access to Grid Resources for Virtual Research Communities

Roberto BARBERA\(^{(1-3)}\), Marco FARGETTA\(^{(3,*)}\) and Riccardo ROTONDO\(^{(2)}\)

\(^{(1)}\) Department of Physics and Astronomy of the University of Catania, Italy
\(^{(2)}\) INFN – National Institute of Nuclear Physics, Division of Catania, Italy
\(^{(3)}\) Consorzio COMETA, Catania, Italy

\(^{(*)}\) email: marco.fargetta@ct.infn.it

ISGC2011 & OGF31
Taipei, 25.03.2011
Outline

• Science Gateway
• Enabling technologies
• The Scenario
  – DECIDE project
• Grid Security and Federation
• Shibboleth and Robot certificates integration
• Conclusions and outlook
The way users access Grid resources has continuously evolved towards simplicity and transparency:

- command line: Globus and gLite CLI - used by the enthusiastic and early adopter scientists
- GUI applications: gEclipse, Grid2Win - good to expand the communities but difficult to maintain
- web interface: GENIUS, P-GRADE - easier for new users but monolithic
- science gateways: Taipei, ISGC2011
A framework of tools that allows scientists to run applications with little concern for where the computation actually takes place. This is similar to cloud computing in which applications run as Web services on remote resources in a manner that is not visible to the end user. However, a science gateway is usually more than a collection of applications. Gateways often let users store, manage, catalogue, and share large data collections or rapidly evolving novel applications they cannot find anywhere else. Training and education are also a significant part of some Science Gateways.
The “brick” Approach

- Science Gateways need to be customised to meet the needs of the Virtual Research Community they support;

- Build them from scratch requires a lot of effort;

- Many small tasks behind the portal are the same and can be shared across different gateways;

- The development should be oriented to create modules, “bricks”, easily deployable in different application context.
• Highly-configurable, scalable, open source portal framework;
• Compatible with JSR 168/286 standards;
• Based on modern web 2.0 technologies;
• Several (>60) portlets for the e-collaboration available out-of-the-box;
• Available with both commercial and free open source licenses;
• Liferay is presently the most used framework to build Science Gateways.
One Liferay... many views

Websites that require the presentation of different pages depending on a user’s login status (i.e., whether the user is logged in).

Websites that require presentation of different pages depending on a user’s role.

Websites that require the integration of multiple existing web applications

Websites that allow groups of individuals to collaborate through applications, on content, or with documents.

SOURCE: www.liferay.com
• Portlets can interact with the Grid e-Infrastructure

• Different approaches available:
  – Execute the Command Line behind the portal
  – Using API where available
    ▪ Must be in Java or other languages supported by Liferay
  – Invoke REST services from javascript code in the browser

• Additional layers between liferay and the Grid can be necessary for some services
• Each portlet can follow its own communication method
• Objectives:
  – Create a support service for the early diagnosis of the Alzheimer and other brain diseases;
  – Build a service accessible via web by the clinicians:
    ▪ Based on a grid e-Infrastructure;
  – Validate the service through applications to real patients cases

• Strategy:
  – Promoting the use by clinicians of specialised applications:
    ▪ CIVET/ADABoost (RMI images);
    ▪ GridSPM (Pet/SPECT images);
    ▪ EEG patterns;
  – Building a pilot European reference e-Service linking the database of images of the European clinical centres;
  – Supporting the clinical community with other performing applications, currently available just to a few researchers.
Different Actors

- People accessing DECIDE services can have different roles and privileges on the available resources
  - Normal User (Neurologist)
    - Upload input data
    - Retrieve the analysis results
  - Expert/External (Collaborator)
    - Normal User privileges but can run analysis on data
  - Data Manager (Scientist)
    - Expert user but can verify the data and update the main DataBase (DB of normal cases)
- Roles and privileges are defined on an application basis
Too Strong Security

- The distributed nature of Grid requires strong security mechanisms;
- Users struggle to comply with complex security rules:
  - Create certificates, create proxy, update credentials and so on;
- Some institutions want to maintain the control of their users’ authentication and the service available:
  - Science Gateways have to be able to interact with other services.
In the web technology arena many approaches are available to federate the authentication among different entities;

A standard provided by OASIS defines the Security Assertion Markup Language (SAML);

Shibboleth is one of the most famous SAML-based tools:

- Implement the SAML standard;
- Allows different approaches to manage users:
  - LDAP, CAS, Plain text, etc.;
- Deployed in many universities and research institutes;
- Free and Open Source;
- Easy to integrate with Liferay;

Shibboleth has been selected for the integration.
1. Access a Service

2. Login

Science Gateway

Authorisation

OpenLDAP™
http://www.OpenLDAP.org

GrIDP (WAYF)

Authentication

Shibboleth.

IDPCT

IDP_1

IDP_n

LDAP

CAS
Usage workflow

1. Portal Login
2. Operation Request
3. Verify ACL
4. Robot Proxy
5. Perform Operations
6. Results

Science Gateway

OpenLDAP

Liferay

enginframe

GLibrary
Role Mapping

- Authorisation is centralised into the LDAP portal;
- Robot proxy may have VOMS attributes corresponding to the roles in LDAP:
  - For each application and user profile a LDAP role and a VOMS attribute is defined;
- Users have to explicitly request the authorisation for the roles they need:
  - A group of experts evaluates the requests;
- If users try to access Grid resources with other tools they do not gain more privileges;
- Roles coming from the federation are currently not accepted:
  - For other projects they could be granted.
Activity Tracking

- All Grid activities performed with robot certificates
  - Impossible to distinguish the users from the proxy (non-repudiability violation)
- The Science Gateway is responsible for the users
- No generic operations on the resources are allowed
  - Only a set of well defined applications and data are accessible through the portal
- The portal and the services made available have to track the user
  - The administrator should be able to identify a user in case of malicious operation on the Grid resources
1. ask for a service
2. create a proxy with the robot certificate
3. execute action
4. get output
5. get the results
Conclusions and outlook

• Conclusions
  – Science Gateways can simplify the use of Grid resources to Virtual Research Communities
  – Integrate Shibboleth with robot certificates allows an easier access to Grid resources for the users
  – User access verified at many levels from the own institution to the gateway

• Future Work
  – Integrate different federations in the same portal
  – Test the new Science Gateway in a production environment
    ▪ Goal of DECIDE is to provide a production service
Any questions?