Using iRODS with the EnginFrame Grid Portal into the GRIDA3 project

Francesco Locunto
< francesco.locunto@nice-software.com >
Marco Piras
< marco.piras@nice-software.com >
Matteo Vocale
< matteo@crs4.it >

The GRIDA3 Project

- GRIDA3 is an interdisciplinary project founded by the Italian Research Ministry in order to provide advanced problem-solving tools for the integration, through a computing portal, of human know-how, instrumentation and resources for data communication, storage, virtualization and computation.

- GRIDA3 will provide a GRID infrastructure for the sharing of data and resources across multiple federated domains, both public and private.

- Project Partners:
  - CRS4
  - NICE
  - University of Cagliari

- People working in: about 25 researchers and technicians.

About CRS4

- Interdisciplinary research center focused on computational sciences:
    - No-profit consortium
    - Operational since 1992
- Research Areas:
    - Visual Computing, Biomedical Computing
    - Distributed Computing, ICT
    - Fuel Cell – Solar Energy
    - Energy and Environment, Bioinformatics
- Technical staff of ~ 140 people.
- Turnover of ~ 10M Euros.
- ~ 50% from external funding:
    - EU /National research project
    - Industrial contracts

NICE

CRS4

Ministero
dell'Università e della Ricerca

eng*in*frame

CRS4 Computing facilities

- Total computing capability: about 40 Tflops (34,7 Tflops on the new cluster only).

- Storage: 480 Tbytes real

- LAN infrastructure based on 60 Gbps backbone links
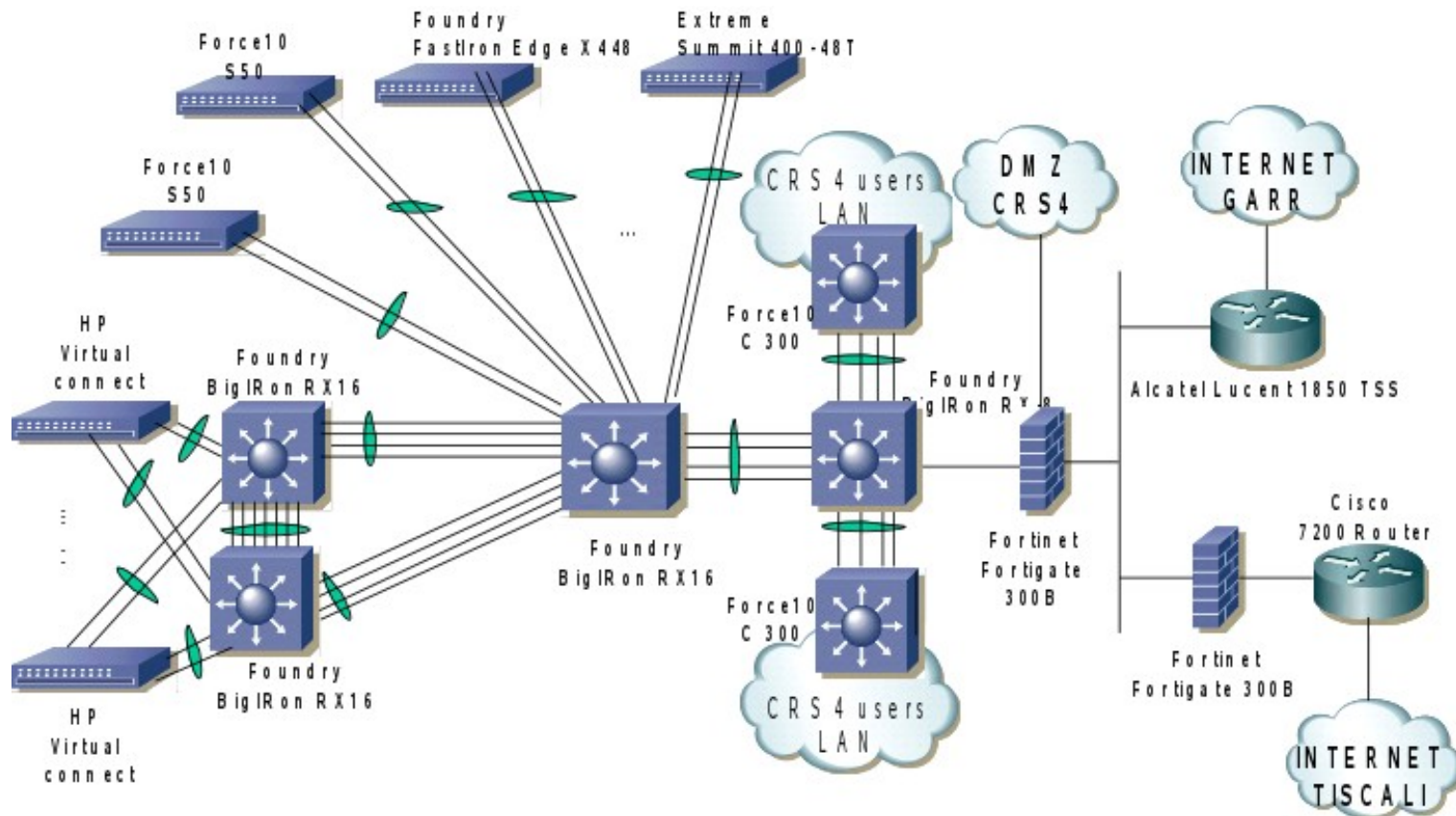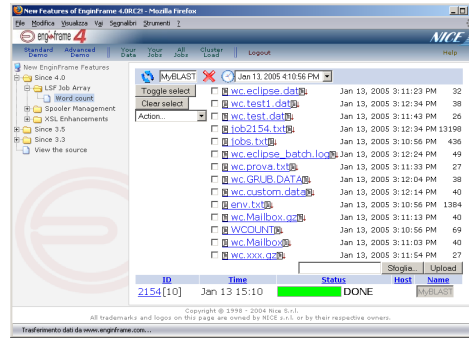
About NICE

- Expertise in HPC/Grid deployment
  - 12 years experience with enterprise Grid solutions throughout all industries
  - Strong relationship with Research and Academia
- Core business: Access to Grid/HPC solutions
  - EnginFrame & GENIUS Grid Portal product line
  - Integration with global ISVs
- Other relevant competencies:
  - Distributed Resource Management
  - Enterprise Grids
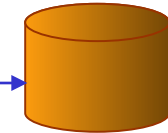 - Visualization Farms
  - Grid Intelligence

What is EnginFrame?

Internal Users

Affiliates

Administrators

Managers

Standard Protocols

Applications
(Interactive & Batch)

License keys

Storage and Data

NICE EnginFrame
Grid Portal / Gateway

Enterprise Portal

Grid/Compute/Visualization Farm
(Linux, Windows, …)

EnginFrame – iRODS integration

Possible solutions:

- EnginFrame plug-in that interacts with iRODS through the *iCommands*
  - light integration
  - it quickly provides valuable results for a first prototype implementation
- Extension to the EnginFrame spooler concept in order to seamlessly and directly embrace iRODS resources
  - more robust approach
  - it implies more development efforts to write new Java modules to be plugged into the EnginFrame core architecture

IRODS plug-in features

Implemented features:

- **IRODS-enabled Remote File Browsing *(RFB)***

- **iRODS-enabled *Remote Spooler***

- **IRODS Authentication system fully integrated with EnginFrame Authentication schema**

- **Bidirectional data transfer**

iRODS plug-in in GRIDA3 Project

- iRODS system can be really useful for the storage needs of GRIDA3.

- Since GRIDA3 Project uses EnginFrame as its Grid Portal application, iRODS – EnginFrame integration is desirable.

- *EIAGRID*, one of the GRIDA3 application, was chosen as the test-bed application for this integration.

About EIAGRID

- The EIAGRID portal provides real-time subsurface characterization by on-the-fly processing of seismic field data and allows fast optimization of the data processing workflow.

- IRODS can be exploited in this context for the storage of all the seismic data (input and processed).

IRODS-enabled Remote File Browsing *(RFB)*

- Extension of the standard RFB feature to allow iRODS users to browse collections in iRODS space and select files from within a standard Web browser.
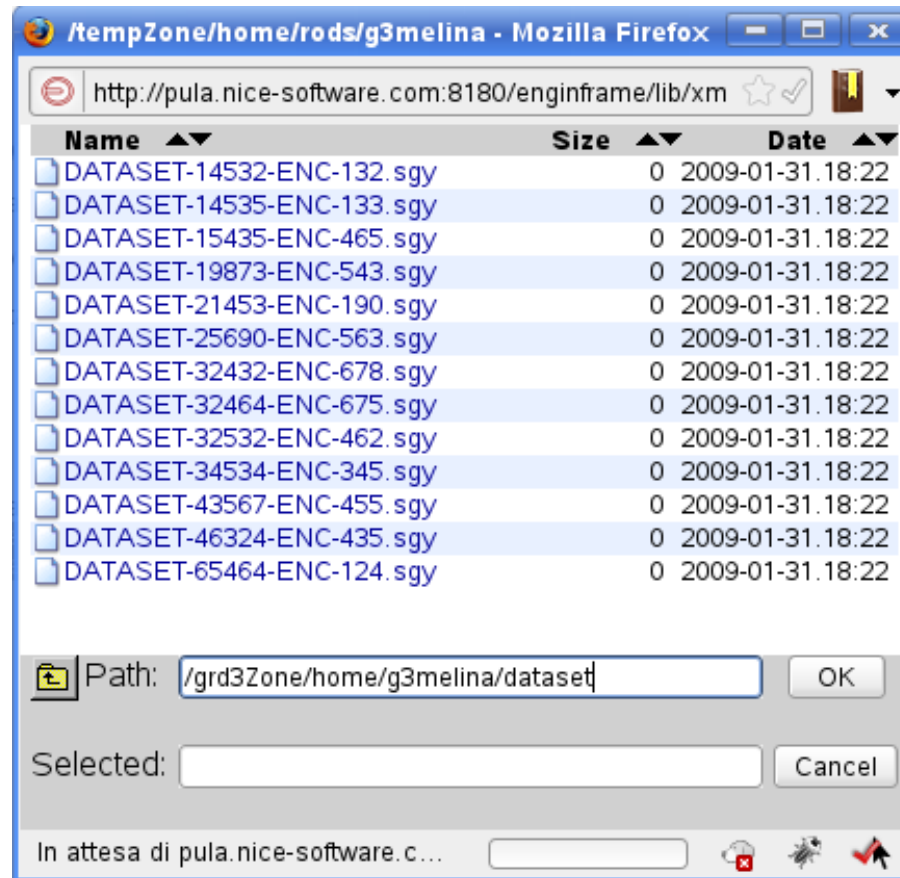
XML representation of an iRODS Collection

```xml
<ef:rfb xmlns:ef="http://www.enginframe.com/2000/EnginFrame">
  <spooler path="/grd3Zone/home/g3melina/dataset">
    <file name="DATASET-14532-ENC-132.sgy" size="4407"
        timestamp="2008-09-18-14.33" />
    <file name="DATASET-14535-ENC-133.sgy" size="6245"
        timestamp="2008-09-19-16.51" />
    <directory alias="InputData" name="InputData" />
  </spooler>
</ef:rfb>
```

iRODS-enabled *Remote Spooler*

- New protocol added to the EnginFrame Remote Spooler system, in order to seamlessly manage files and collections stored in an iRODS virtual file system.

Actions in an iRODS Remote Spooler

- Using the *actions* available in Remote Spooler is possible to directly interact with iRODS resources.

- The user can seamlessly download, upload, delete and also archive in various formats any iRODS file and Collection.

iRODS – EnginFrame Authentication integration

- iRODS Authentication system is perfectly integrated into the EnginFrame Authentication mechanism.

- In order to use any iRODS resource, an EnginFrame user must be previously associated with an iRODS user.

- By this way, logging into the EnginFrame Portal automatically logs the user into the iRODS system with its credentials.

- The plug-in includes scripts for easily managing these associations.

Conclusions

- We have proposed a method to integrate iRODS data Grid and compute Grid components via the EnginFrame Grid Portal.

- EIAGRID portal application have been described to show how it is possible to exploit this type of integration.

- The iRODS plug-in is still a prototype, though it is fully functional.

- We expect the iRODS plug-in will be useful in the construction of general data exchange infrastructures for wide range of different application areas.

Thank you for your attention!

- Questions?

- If you are interested in an online demo, please contact us personally.