State of Interoperability

Laurence Field
CERN

OSG

- **November and December 2004**
  - Initial meeting with OSG to discuss interoperability
    - A common information schema was the key
  - Proposal for version 1.2 of the Glue Schema was discussed
    - Include new attributes required by OSG, Marco Mambelli
- **January 2005**
  - Proof of concept was tried, Leigh Grundhoefer (Indiana)
    - Installed Generic Information Provider (GIP) on an OSG CE
    - OSG CE was configured to support the dteam VO
    - “Hello world” job, submitted through the LCG RB and ran on an OSG CE
    - Installed the LCG clients available on OSG from a tarball
      - *Oliver Keeble (CERN)*
    - Submitted test job that did basic data management operations

OSG

- **Modifications to the OSG and LCG software releases**
  - Updated the GIP to publish version 1.2 of the Glue Schema
    - The GridFTP server on the OSG CE advertised as an LCG SE
  - Automatically configure the GIP in the OSG release
    - Information scavenger script, Shaowen Wang (Iowa)
- **August 2005 (month of focussed activity)**
  - Included first OSG sites into the LCG operational framework
  - Set up a BDII that represented these OSG sites
  - Included this BDII to the LCG information system
  - All OSG sites found in this BDII were automatically tested
    - Using the Site Functional Tests (SFT) framework
  - Created a script to install the LCG clients on OSG CEs
- **November 2005**
  - First user jobs from GEANT4 arrived on OSG
  - GIP validator for OSG operations. Shaowen Wang (Iowa)

OSG

- **March 2006, Operations Progress**
  - Information system bootstrapping.
    - Dynamic web page from OSG GOC DB.
  - Routing of trouble tickets.
  - Joint operations VO
    - For running tests.
    - Deployment of client libraries.
  - OSG joined the Monday WLCG operations meeting to report on WLCG issues
- **Summer 2006**
  - CMS successfully taking advantage interoperations

Future Work

- **Maintain Interoperation**
  - As grids evolver, ensure we maintain interoperability
- **Use Case Testing for Authorization**
  - Same functionality for identical roles
- **Continuous SRM Testing**
  - Between different versions and implementations
- **Accounting Discussions**
  - How do we do VO accounting across grids

ARC

- **Four official meetings.**
  - 31 August 2005
  - 31 October 2005
  - 18 January 2006
  - 23 March 2006

- **Short Term Goals.**
  - ARC2Glue Mapping Document
    - Create a translator gateway.
  - Enable the RB to submit to ARC CEs
  - Enable the ARC clients to submit to the RB.
- **Long Term Goals**
  - Use a common(standard) schema.
  - Use a common(standard) CE interface.

ARC Status

- **Prototype Translator Ready**
  - Mapping ARC2Glue
  - Translation document ready
    - Translator created
      - ***For both ways***
    - BDIIs setup
      - *Using translators*
- **Tested Condor to ARC CE submitter**
  - Shown that it works with later Condor versions.
- **Currently adding this functionality to the RB.**
  - Will also required condor upgrade
- **Started discussions on operations.**
- **Long term goals**
  - The long term goals are out of scope for this activity
  - “Both EGEE and NDGF are commit to standards for grid computing and will both work together with other grids to try and achieve this goal”

Naregi

- **Initial meeting in March**
  - To discuss willingness and feasibly.
  - Initial plan drafted.
- **More discussions privately and at GGF**
  - Especially with respect to information systems
- **Information translators**
  - Nargei2Glue ready
    - BDII set up
  - Glue2Naregi well underway
- **Job submission in progress**
  - Recent discussion at GGF to explain details

Unicore

- **Official EGEEII activity**
  - 1st F2F Meeting August 2006
- **Analyzed both architectures**
  - Worked out a plan
- **Need to identify pilot VO**
  - Fussion?
- **Need to identify infrastructure**
  - DEISA
- **Investigate condor-u component for job submission**
  - Same solution for as for ARC

OGF-GIN Activity

- **Ad-hoc discussion at SC2005**
  - First official meeting at GGF 16 in Athens
  - Trying to bridge the grid islands
- **Split into four groups**
  - gin-auth –security
  - gin-data – data management
  - gin-info – information systems
  - gin-jobs – job submission
- **Building on bi-lateral and previous work**
  - OSG/EGEE activity
  - ARC/EGEE activity
  - ITGF
- **Central place for discussion and show cases**

gin-auth

- **Supported CAs defined by the ITGF**
  - All infrastructures must support this set.
- **VO naming convention must be used**
  - To avoid conflicts between grids
- **GSI compliant x.509 proxy certificates**
  - Or OGSA Basic profile authentication
- **Transport of supported authorization attributes**
  - via VOMS extensions
- **Work needed on policy management**
  - For VOMS roles and groups
- **Are proxies working?**
  - What alternative security models could we try?

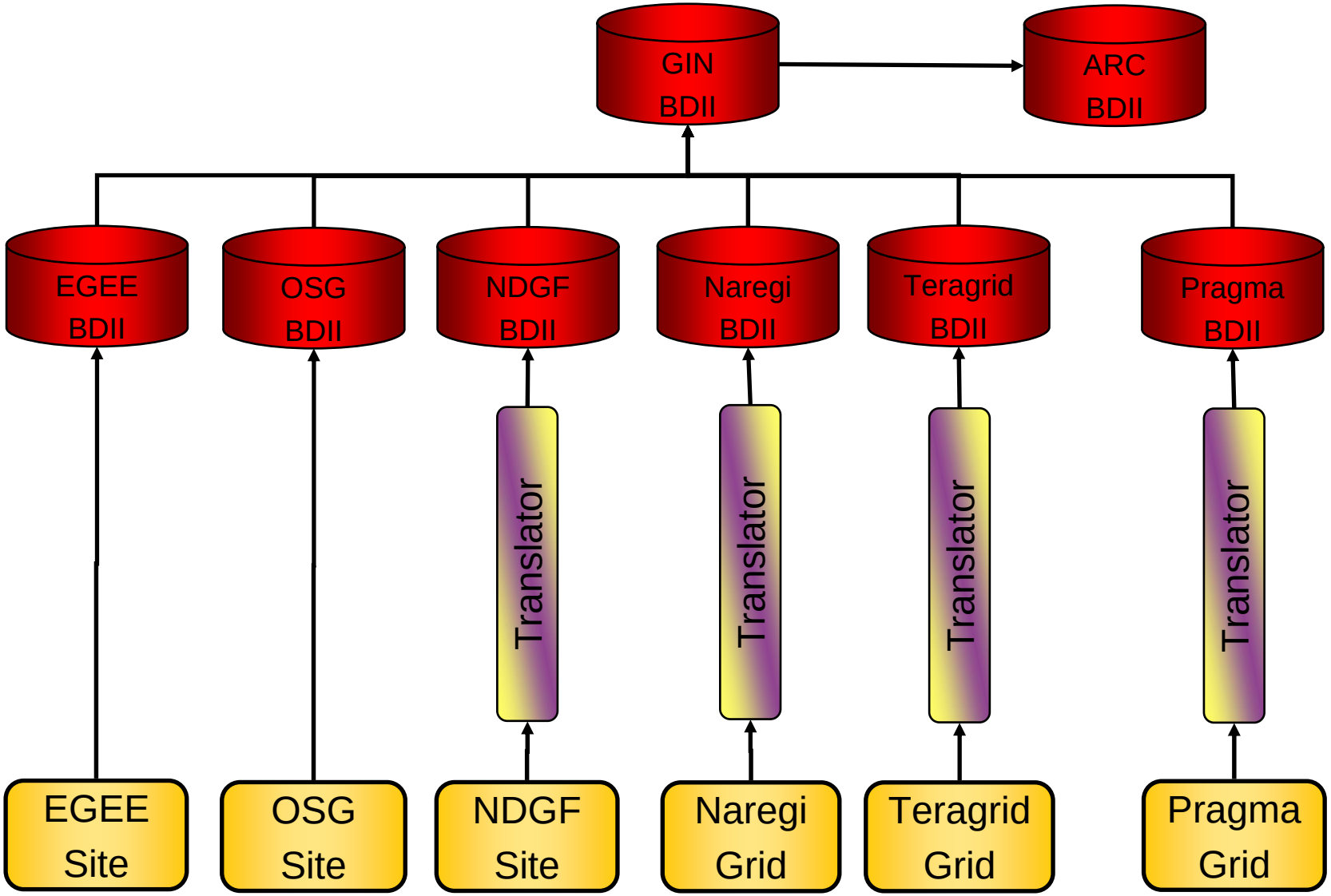
gin-data

- **Point to point movement of data**
  - between storage in different grids
- **Usage of managed resources**
  - and their more sophisticated APIs (e.g. SRM, SRB)
- **GIN-Data is sponsoring three distinct activities:**
  - grid-ftp interoperability
  - SRM interoperability
  - SRB interoperability
- **On going testing of the above activies**
  - Between different implementations
    - And installations

gin-info

- **Recognized the need for common information**
  - Define the minimal set of common attributes
- **Translators between the island**
  - Schema mapping
  - Implementation of translators
- **Set up a BDII per infrastructure**
  - EGEE, NDGF, OSG, Teragrid, Naregi Pragma
  - Using the Glue Schema
- **Translate between this BDII and the native system**
  - Eg gin-bdii -> Naergi “cell domain”
  - Avoids the n*n problem

Initial Architecture



What Now?

- **gin-bdii contains information from all grids**
  - Information of varying quality
- **What do we do with it?**
  - Need a use case
  - What is always the first use case?

- **“Sites on a map”**
  - Easy to do manually with a small number of sites
  - Problems with a large number of sites
    - Time consuming
    - Bad geographical knowledge
      - *Can lead to huge political errors ☺*
  - Using different maps
    - Need to automate

“Site on a map” use case

- **Mandatory attributes**
  - Site Location
    - Latitude and Longitude
  - Site Name
  - Unique identifier for the site
- **Optional attributes**
  - Site Description
  - Site Location, human readable form
  - Site email contact
  - Site web page
- **Glue schema has a site entry**
  - Provides this information

Implementation

- **OSG and EGEE**
  - Already provide the Glue Site entry
- **For the others**
  - Fudge in the translator ☺
  - Create configuration file containing mapping and information
    - Map a cluster id to site information.
- **Visualize information with Google Earth**
  - Script used to generate kml file
    - Queries gin-bdii to find information
    - Script run every 5 mins via cron
  - http://www.cern.ch/lfield/gin.kml

Information Quality

- **Quality of an Information System**
  - Dependent on the quality of information
- **Many reason for poor quality information**
  - Poor schema design
  - Poor quality information providers
  - Incorrect deployment or configuration
  - Site problems
  - Bit rot etc.
- **How do we ensure good quality information?**
  - Need to develop tests for the information
    - Based on the use cases
      - *How can we ensure the coordinates are correct for a site?*
  - Sounds like grid operations!

Experience Gained

- **Information systems are very similar**
  - Information provider to populate system
  - Query mechanism to extract data
  - Hierarchical architecture, resource -> site -> top
  - Information conforms to a schema
- **Joining information systems is easy**
  - Create and information provider for one system
    - Which queries the other
- **Translating information is tricky**
  - Moving from one model to another is straight forward.
  - Showstopper if information doesn't map, ie missing attributes
- **We can live with different information systems**
  - But we can't live with different information

Summary

- **Two terms have been highlighted**
  - Interoperation:
    - Two grid infrastructures working together
  - Interoperability:
    - Grid middleware enable to work together
- **Need to focus on Interoperation**
  - Interoperability may be needed to achieve interoperation
- **OSG/EGEE interoperation achieved**
  - Need to ensure that this is maintained
- **ARC/EGEE in progress**
- **Taking first steps with Naregi and Unicore**
- **OGF-GIN**
  - Working to link all the grid infrastructures
    - One link at a time