GOCDB
A repository for a worldwide grid infrastructure

G. Mathieu, A. Richards, J. Gordon, C. Del Cano Novales, P. Colclough, M. Viljoen

CHEP09, Prague, March 2009
Outline

• What is GOCDB
  - Context and generalities
  - Purpose and history
  - GOCDB interactions with other EGEE tools

• GOCDB architecture
  - Components, services and interfaces
  - Development and procedures
  - Availability and failover

• Evolution
  - GOCDB in a distributed environment
What is GOCDB? (1)
Context and generalities

- What’s in a name?
  - GOCDB = Grid Operations Centre DataBase

- A central (static) information repository

- Key component of EGEE and WLCG, used as an authoritative data source
What is GOCDB? (2)

Context and generalities

- Stores information about (and links together):
  - Regions and countries
  - Sites, nodes and services
  - Users

- Consists of 2 parts:
  - A Database
  - A web portal (https://goc.gridops.org)
What is GOCDB? (3)
Handled data

- Administrative info
  - Contacts, names, groupings (region, country...)
  - *Used for official procedures, reports, listings*

- Resources and services
  - Service endpoints and types
  - *Used for monitoring and operations*

- Maintenance plans
  - *Used for monitoring and availability reports*
<table>
<thead>
<tr>
<th>Year</th>
<th>Era</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>Stone Age</td>
<td>Static list of sites and their contacts</td>
</tr>
<tr>
<td>2004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>Iron Age</td>
<td>Simple MySQL DB + scripts</td>
</tr>
<tr>
<td>2006</td>
<td>Medieval times</td>
<td>MySQL DB + proper web interface</td>
</tr>
<tr>
<td>2007</td>
<td>Industrial Revolution</td>
<td>Oracle DB + user friendly web portal</td>
</tr>
<tr>
<td>2008</td>
<td>Modern times</td>
<td>Oracle DB + Web portal + web services</td>
</tr>
<tr>
<td>2009</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
What is GOCDB? (5) interactions

- GridView Accounting Portal
- Site Fabric Monitoring
- APEL, Accounting Enforcement Portal
- SAM, GStat
- Operations Dashboard
- GGUS
- GOCDB, Operations Portal
Architecture overview

- Web portal
  - Based on a Model-view-Controller (MVC) principle
  - Coded in PHP, object fashion
- Database
  - Hosted on Oracle 11g cluster
  - 3 schemas: Production, test, development
- Authentication
  - Based on X509 certificates
  - Role management
Services and interfaces

- GOCDB Programmatic interface (GOCDB-PI)
  - REST based interface over https
  - Business logic put in stored functions at DB level
  - Use of Oracle XML DB
  - Different methods with different security levels

- SOAP web services under study
Processes and procedures

- Code synchronisation
  - GOCDB repository in SVN
- Release procedure
  - RPM deployment
  - IP switch between production and test server
- Development cycle
  - Savannah task/bug tracker
  - Well established validation process
Development workflow

1. **USER** submits a request.
2. The request is sent to GGUS and UKI Helpdesk.
3. If the request needs discussion, it is sent to Mail/other.
4. If the request needs further validation, it is sent to GOCDB AG.
5. If the request is acceptable, it is added to the Official Development List.
6. If the request is not acceptable, it is marked as refused.

- **Savannah Wish List**
- **GOCDB Admin**
- **OAT**

**Decision Points:**
- Does the request need discussion? (yes/no)
- Does it need further validation? (yes/no)
- Is the request acceptable? (yes/no)
GOCDB Failover (1)

- Web portal
  - Master instance at RAL, UK
  - Replica instance at ITWM, Germany

- Database
  - Master instance at RAL, UK
  - Replica instance at CNAF, Italy
  - Local backup instance at RAL, UK
  - Oracle 11g
  - Oracle 10g
  - Oracle Xpress
GOCDB Failover (2)

- End users
- 3rd party tools
- Master web portal
  - GOCDB-PI
  - conf
- Replica web portal
  - conf
  - GOCDB-PI
- Master DB
- Backup DB
- Replica DB

DNS alias: https://goc.gridops.org
GOCDB evolution

• Model evolution in EGEE
  – From central to distributed
  – Automate, simplify, standardize

• Distributing “the beast”
  – New architecture under development
  – Challenging work and short deadline
Main ideas

- Keep a central service, not necessarily a central DB
  - There is a need for a central access point, but:
  - the fact that regional DB are distributed or not must not be an issue

- Build a sustainable architecture that allows regionalisation but doesn’t force it
  - Not all regions are at the same level

- Propose an implementation where nothing exists, work with existing solutions otherwise
  - Some regions have their own solution and don’t want to be forced to use another one
High level architecture

Region 3 local portal
Region 3 homemade local DB
Publish (push)

3rd party tools
End users

WS interface
Central portal
Query interface

Data collector
GOCDB Central instance

Query (pull)

Region 1 local portal
Query interface

GOCDB regional instance

Region 1
RAL
One level down

Region 3 local portal

Region 3 homemade local DB

Publisher

WS interface

Central portal

Query interface

GOCDB internal model (regions 1, 2 and 3)

Data collector

Region 1 local portal

Query interface

GOCDB Region 1 model

RAL

GOCDB internal queries

3rd party tools

End users
Details and components

Region 3 local portal
- Region 3 homemade local DB
- Publisher

3rd party tools
End users

Region 1 local portal
- Query interface
- Custom tables
- Data collector

Data processor
- Region 3
- Region 2
- Region 1

Data
- R3
- R2
- R1

Metadata
- R3
- R2
- R1

Central portal
WS interface

Query interface

RAL
GOCDB internal queries
Current relational model

- Physical Data Tables
- Hard coded relationships and constraints

Proposed object model

Core tables (relationships)

Data tables

Collection1  Collection2  Collection3
For more details...

- GOCDB homepage
  - http://www.grid-support.ac.uk/content/view/406/290/

- GOCDB4 development wiki
  - http://goc.grid.sinica.edu.tw/gocwiki/GOCDB4_development

- “A pseudo object database model and its applications on a highly complex distributed architecture”
  - IARA/IEEE Conference on Advances in Databases (DBKDA 2009)
    March 1-6, 2009 - Cancun, Mexico
Thank you

- Questions?