GOCDB4

Gilles Mathieu, RAL-STFC, UK

An introduction to regional operation tools
Monday 21 Sept 2009, EGEE'09, Barcelona

www.eu-egee.org

e-infrastructure

EGEE and gLite are registered trademarks

Enabling Grids for E-sciencE

- **GOCDB4 Overview**
  - Context and big picture
  - Architecture of a GOCDB module
  - Central and regional interactions

- **GOCDB4 in regions: review of possible use cases**
  - Region 1: Use a GOCDB4 module in region
  - Region 2: Keep using central GOCDB
  - Region 3: Use an other tool in region
  - What is the best solution for me?

- **Outstanding issues, timeline and future plans**
- **Questions**

Enabling Grids for E-sciencE

- **Why a new GOCDB?**
  - To transform the central operation tool it is now into a tool available both for central and regional purposes.
  - To ensure the new version can work in an EGI/NGI context

- **Key principles**
  - Keep a central service and propose regional services
  - Build a sustainable regionalised architecture
  - Propose an implementation where nothing exists, work with existing solutions otherwise

- **Solution**
  - A module, deployable and configurable, to be used as the central service and distributed to the regions that want it

Standard interfaces

- Giving access to the tool to end users and tools

-Allowing modules to communicate with one another

3rd party tools

End users

XML I/O

- Standard and configurable ways to access the data

- Transparent and independent from actual data schema

Web services / interfaces

| SOAP | GOCDB-PI | ActiveMQ |

GUI

| Portal |

XML output

XML input

PROM database

- A relational model implemented the object way

- Relations are stored as meta-data in the DB itself

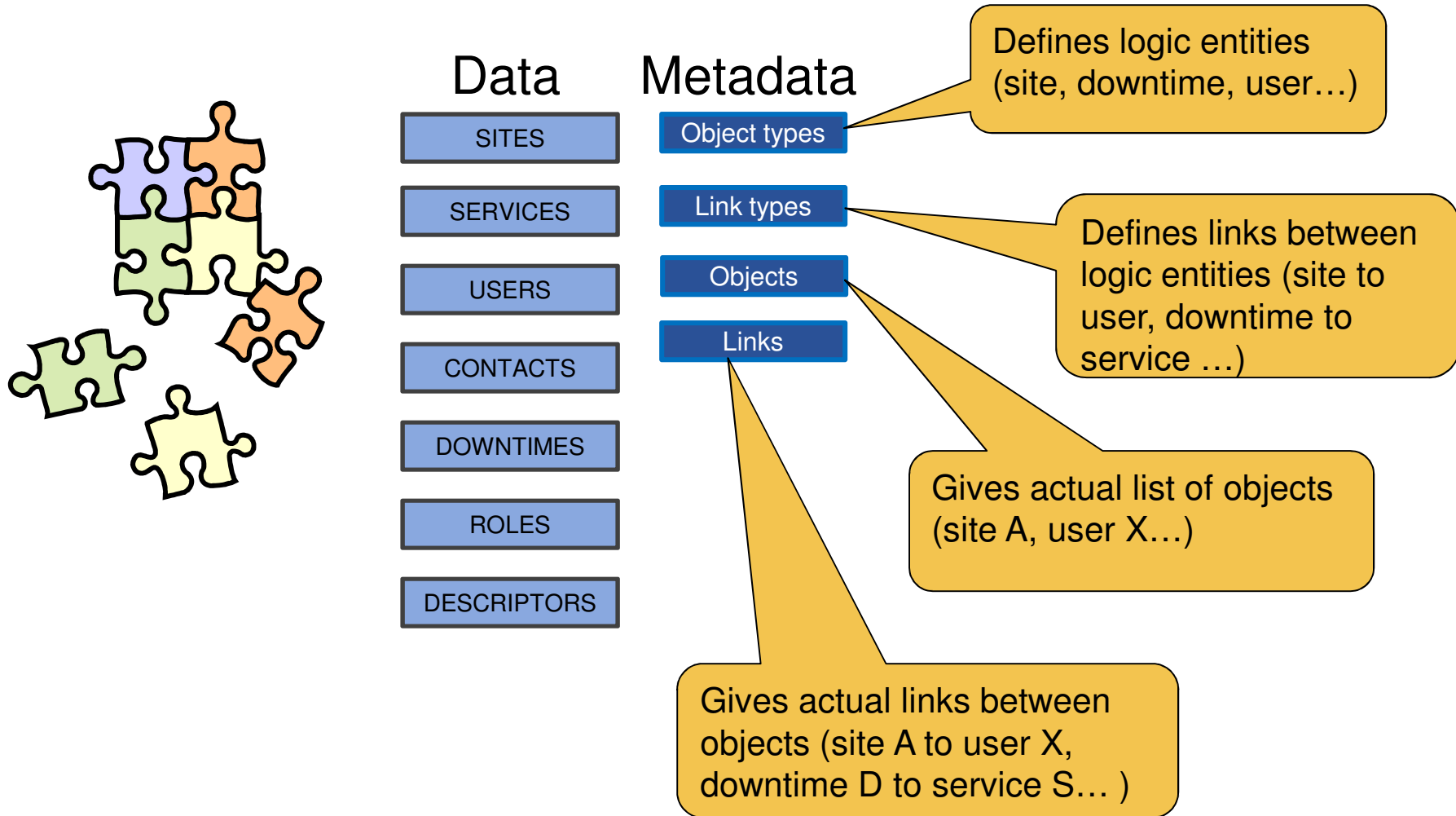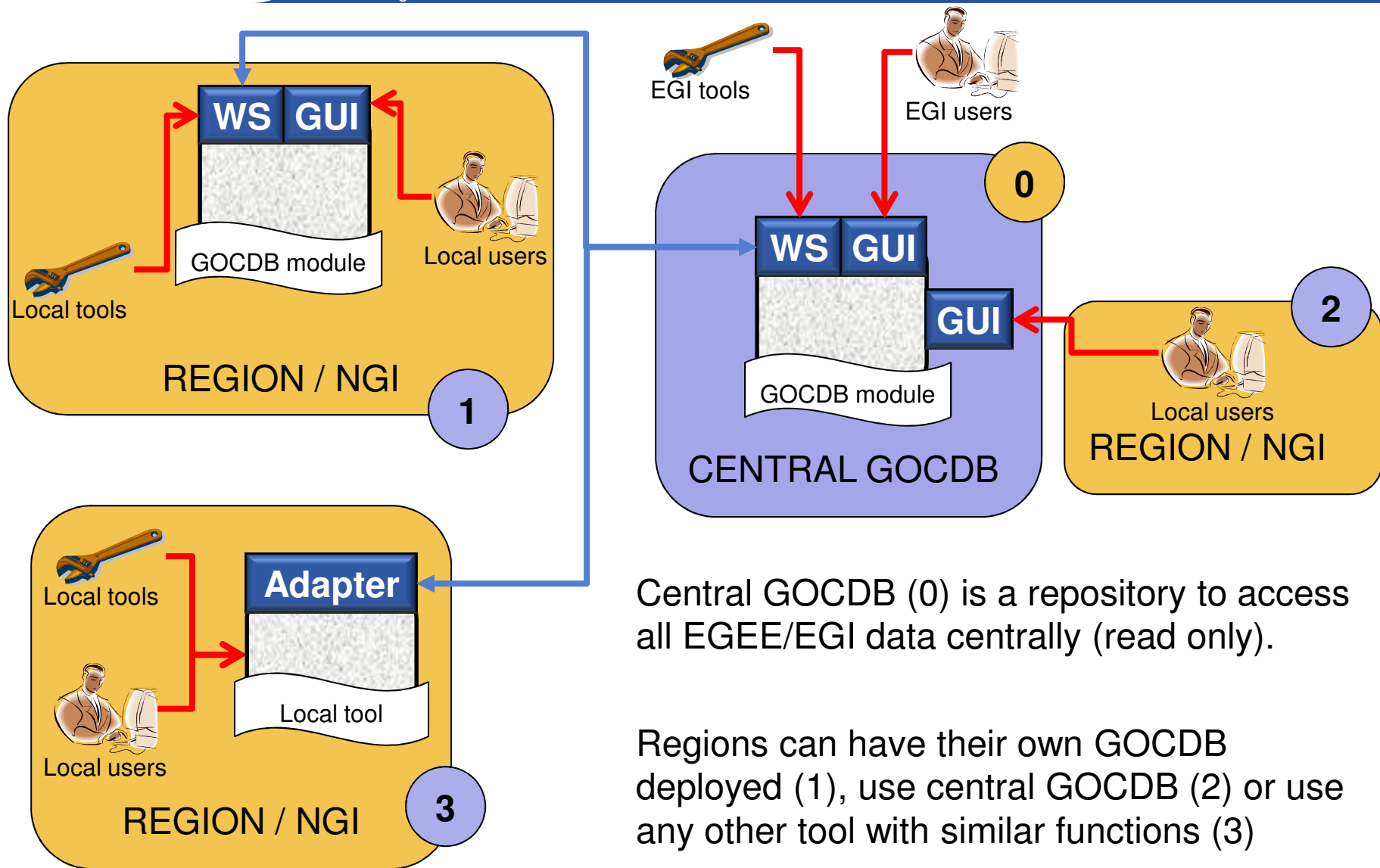- The schema can easily evolve and is completely configurable
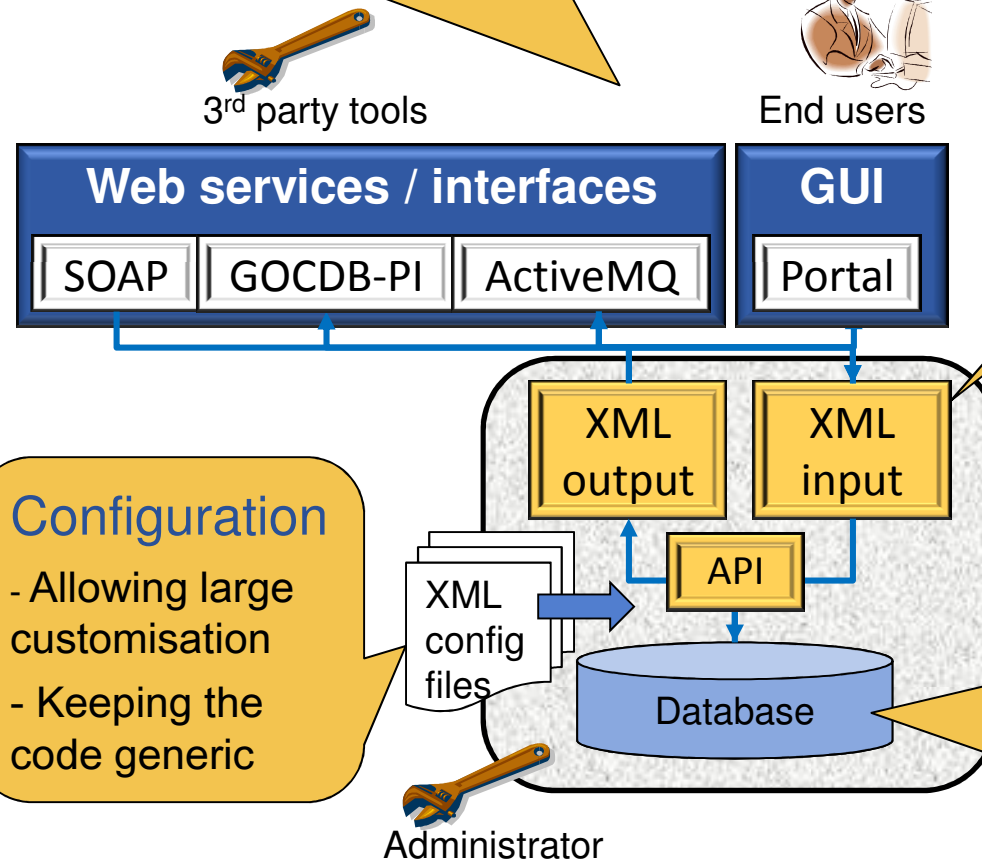
Configuration

- Allowing large customisation

- Keeping the code generic

XML config files

API

Database

Administrator

GOCDB schema designed the "PROM" way

Data

- SITES
- SERVICES
- USERS
- CONTACTS
- DOWNTIMES
- ROLES
- DESCRIPTORS

Metadata

- Object types
- Link types
- Objects
- Links

Defines logic entities (site, downtime, user…)

Defines links between logic entities (site to user, downtime to service …)

Gives actual list of objects (site A, user X…)

Gives actual links between objects (site A to user X, downtime D to service S…)

WS GUI

GOCDB module

Local tools

Local users

REGION / NGI **1**

EGI tools

EGI users

WS GUI

GUI

GOCDB module

0

CENTRAL GOCDB

Local users

REGION / NGI **2**

Adapter

Local tools

Local users

Local tool

REGION / NGI **3**

Central GOCDB (0) is a repository to access all EGEE/EGI data centrally (read only).

Regions can have their own GOCDB deployed (1), use central GOCDB (2) or use any other tool with similar functions (3)

Enabling Grids for E-sciencE

- **Use case description**
  - A GOCDB module is deployed locally
  - Code and support provided by GOCDB team
  - Operation, maintenance and customisation done locally
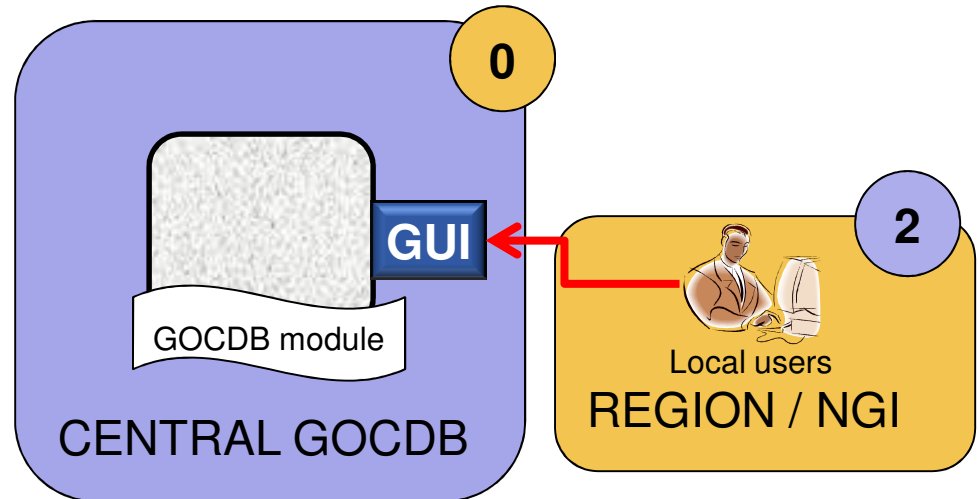
- **Main features**
  - Local users and tools access regional data locally
  - The module can be customised to suit regional needs
  - Data synchronised with central GOCDB through standard web services interfaces

To Central GOCDB

WS GUI

GOCDB module

Local tools

Local users

REGION / NGI

1

Enabling Grids for E-sciencE
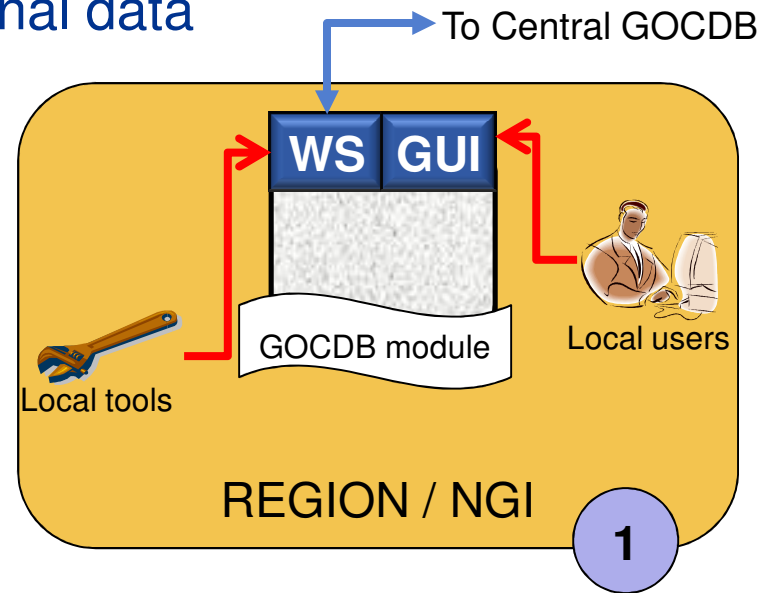
- **Hardware/software requirements**
    - A web machine with Apache2, PHP5, oci8 and SSL support
    - An Oracle database (10g or higher) – OracleXE is fine

- **Deployment process**
    - Download and install GOCDB4 RPM
    - Customise configuration files
    - Run web-based administration tool to:
        - Deploy database schema
        - Configure web portal
    - Contact us to import data from central GOCDB to your instance
    - Online documentation is at
      http://goc.grid.sinica.edu.tw/gocwiki/GOCDB_Regional_Module_Technical_Documentation

Enabling Grids for E-sciencE

- ## Use case description

  - There is no tool deployed locally

  - Operation and maintenance provided by GOCDB team

- ## Main features

  - Local users and tools access regional data centrally

  - Instance cannot be customised to suit regional needs

  - Data stored directly in Central GOCDB

Enabling Grids for E-sciencE

- **Use case description**
  - A tool (not GOCDB) is deployed locally
  - Code and support provided locally
  - Operation, maintenance and customisation done locally

- **Main features**
  - Local users and tools access regional data locally
  - Technical possibilities depend on chosen tool
  - Data synchronised with central GOCDB through an adapter implemented locally following defined specifications

To Central GOCDB

Adapter

Local tools

Local tool

Local users

REGION / NGI

3
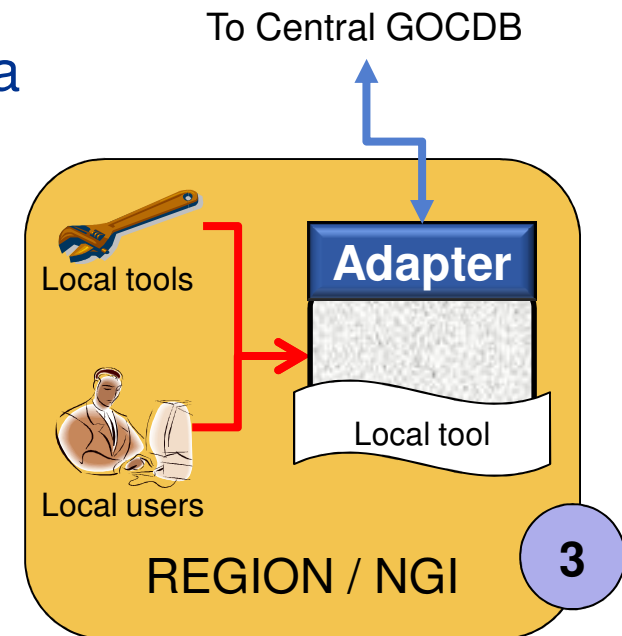
- **Region 1**
    - Schema and code already provided   **+**
    - Customisable with minimum effort   **+**
    - Oracle needed   **-**
- **Region 2**
    - No need to care about hosting   **+**
    - No need to care about any development   **+**
    - Not customisable   **-**
    - Central schema not ideal for local use   **-**
- **Region 3**
    - All latitude in the choice of a solution   **+**
    - Possibly reuse existing infrastructure in region   **+**
    - Need to develop publishing adapter   **-**
    - Development and maintenance effort   **-**

Enabling Grids for E-sciencE

- **GOCDB4 Release schedule**
  - Regional package test versions – ***Summer 2009***
  - Central GOCDB4 in production – ***early November 2009***
  - Central GOCDB3 decommissioned – ***November 2009***
  - Regional package beta version – ***December 2009***
  - Regional package production version – ***early 2010***

- **Regionalisation plan**
  - Regions currently involved in testing: UK, Germany, Serbia
  - Pilot regions for beta deployment: UK, Ireland, Italy, Germany
  - All schedules to be agreed with corresponding regions

Enabling Grids for E-sciencE

- **Migration from GOCDB3 to GOCDB4**
  - Keep interruption of service to a minimum during migration
  - Ensure data consistency
  - Ensure no workflow is broken along the way

- **Regional package release and deployment**
  - Large scale testing relies upon availability in regions
  - Prepare release and distribution plan for future versions

- **Future plans and evolutions**
  - Build a MySQL version
  - Interactions with other operation tools
  - GOCDB/GlueSchema convergence

Enabling Grids for E-sciencE

- **GOCDB4 architecture**
  - http://goc.grid.sinica.edu.tw/gocwiki/GOCDB4_architecture

- **GOCDB4 development and plans**
  - http://goc.grid.sinica.edu.tw/gocwiki/GOCDB4_development

- **PROM (DB model behind GOCDB4)**
  - "A pseudo object database model and its applications on a highly complex distributed architecture" *by P.Colclough*

    IARA/IEEE 1st Conference on Advances in Databases (DBKDA 2009)

    March 1-6, 2009 – Cancun, Mexico

- **Questions?**