



UNIVERSITÀ DEL SALENTO



GRelC Service (TSA3.2.3 Services)

Sandro Fiore, Ph.D.

(sandro.fiore@unile.it)

SPACI and University of Salento, Lecce, Italy

GRelC Service (TSA3.2.3)

Introduction about the GRelC Project

- GRelC Project and GRelC service
- Architecture and main features
- Security support

SA3.2.3 Activity in EGI

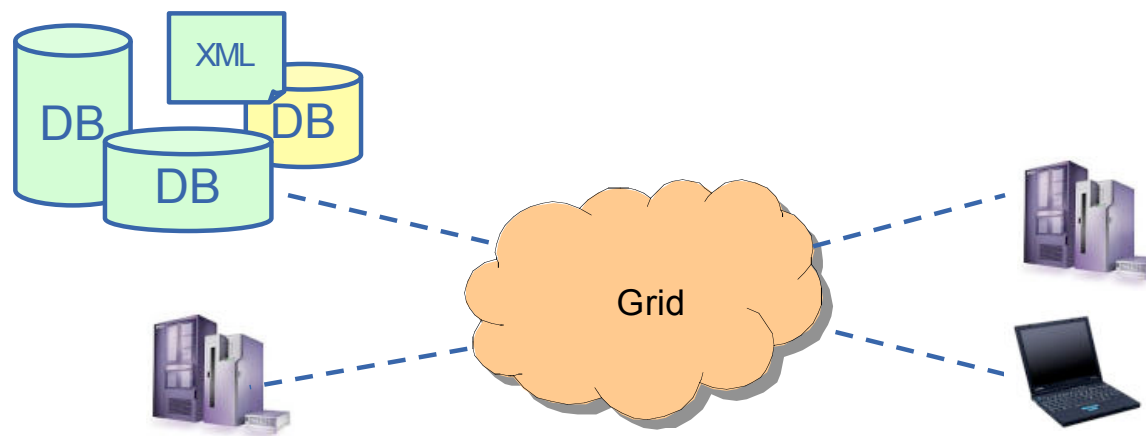
- DoW
- Activity in Q1
- Foreseen activity in Year1

Website, tutorials and portal

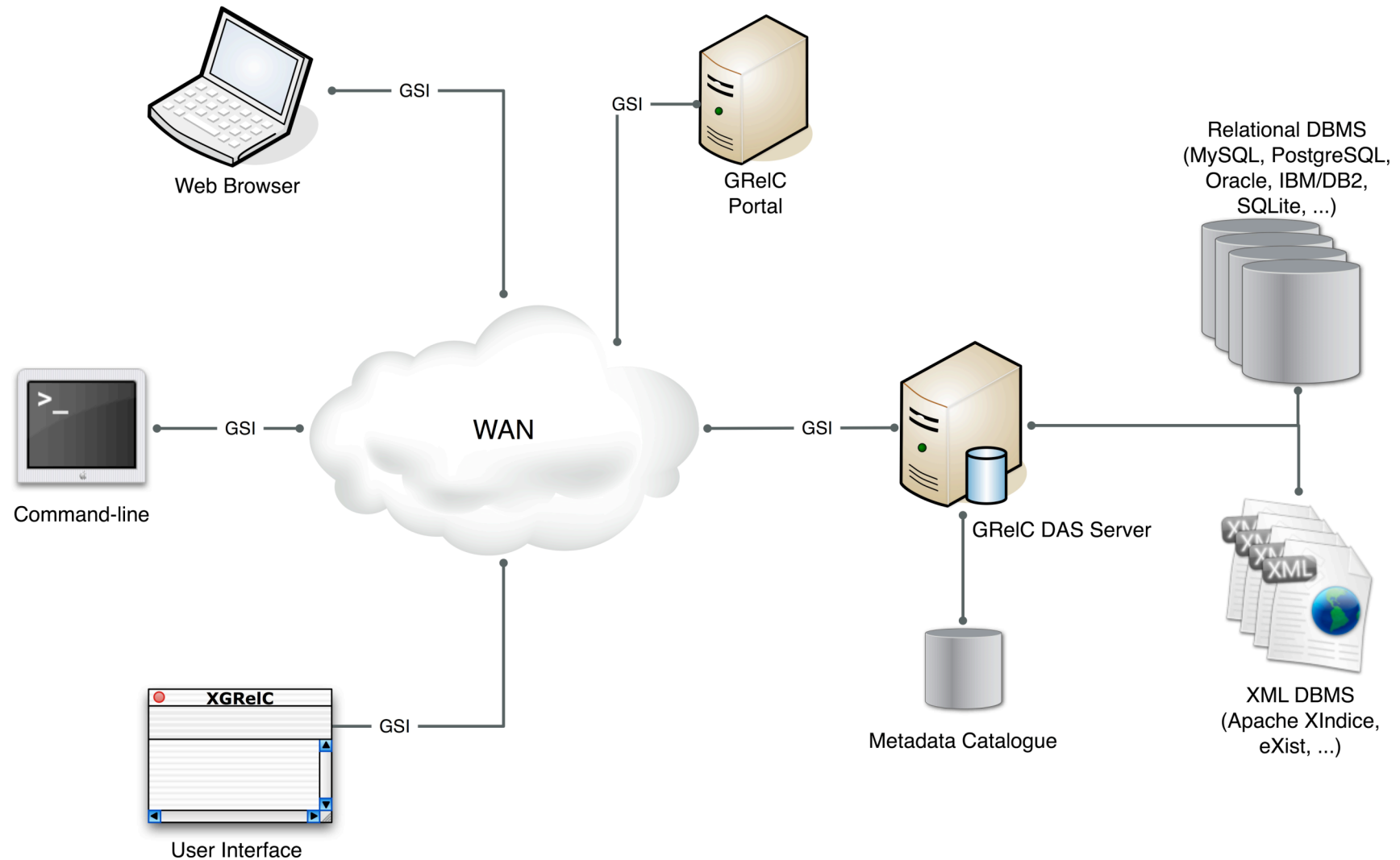
- User support through GILDA
- The GRelC website
- The GRelC Portal

Introducing the GRelC Project

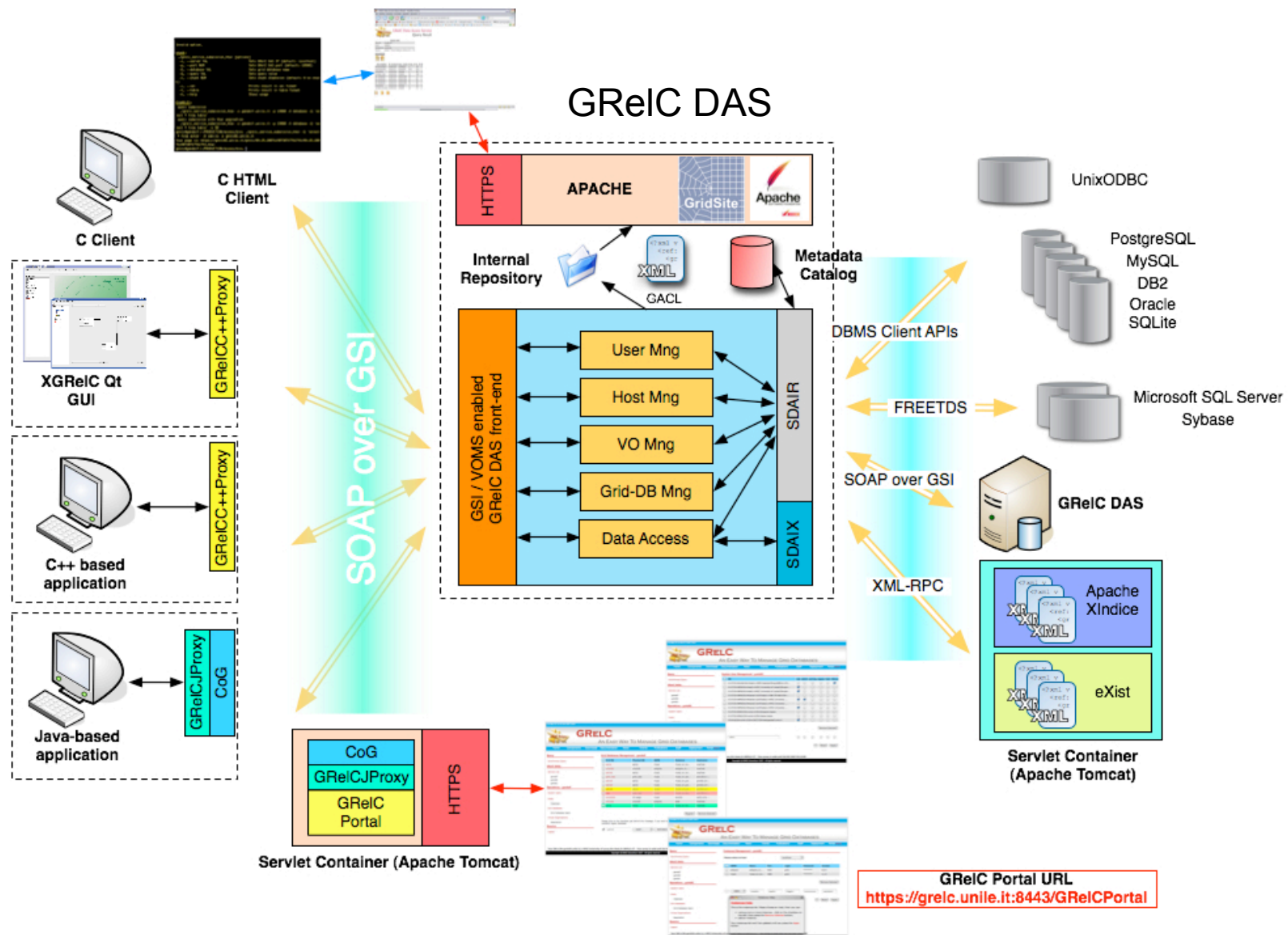
- Grid Relational Catalog is a project which aims at designing and developing a set of efficient, secure and transparent Data Grid Services (Starting date, January 2001).
- **GRelC Data Access Service** aims at providing a large set of functionalities to access both relational and non relational DataBases in a grid environment.



GReIC DAS Architecture (in the large)



Architecture in the small



GReIC DAS: Main Features

- Entirely based on C programming language (server side)
- Multithreaded web service
 - Next releases will be pre-threaded
- It exposes the **web service interface GSI enabled** and WS-I compliant
 - Compatibility with XML, SOAP, WSDL standards
- **Mutual authentication** based on GSI (X.509v3 digital certificates)
- Combined Authorization mode
 - GReIC DAS **local authorization** based on ACL
 - Wide set of data access control policyies
 - VOMS Support for **global authorization**
 - Wide set of roles on the VOMS server side
 - *Mapping role <-> set of privileges*
- Information System Support (**BDII** compliant schema extensions)
- **Full GSI support**: data encryption, data integrity, protection against replay attacks and detection of out of sequence packets

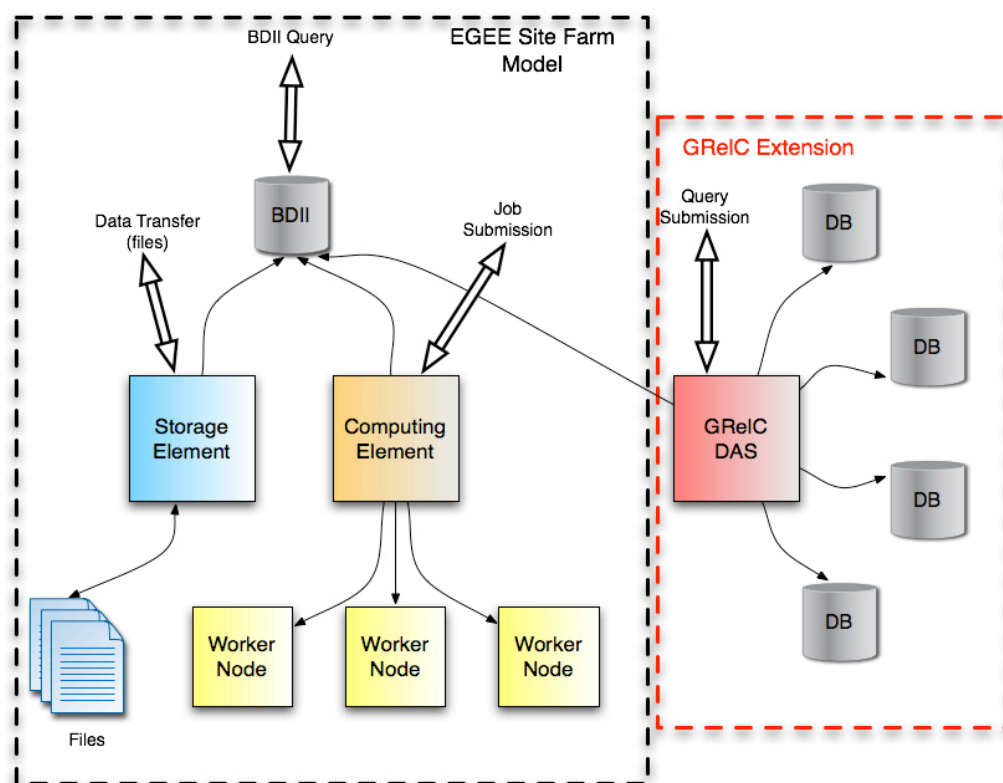
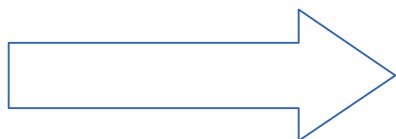
GReIC DAS: Main Features

- Support for heterogeneous **data models** (relational and hierarchical)
- Support for **synchronous** and **asynchronous** queries
- Dynamic binding to **heterogeneous RDBMSs** and **XML-DB engines**
- **Two levels logging** for users, connections, queries, etc.
 - **Server Level**
 - Management activities performed on the server side
 - **Grid-Database Level**
 - Queries, connections, etc.
- GSI enabled remote administration tools and remote log
- **Compression, chunking, prefetching** and **streaming** to enhance performance on a WAN
- Wide **SDK** for developers (mainly C, and Java)
- **No dependencies** concerning other middleware (only GSI)
 - Globus and gLite compatibility

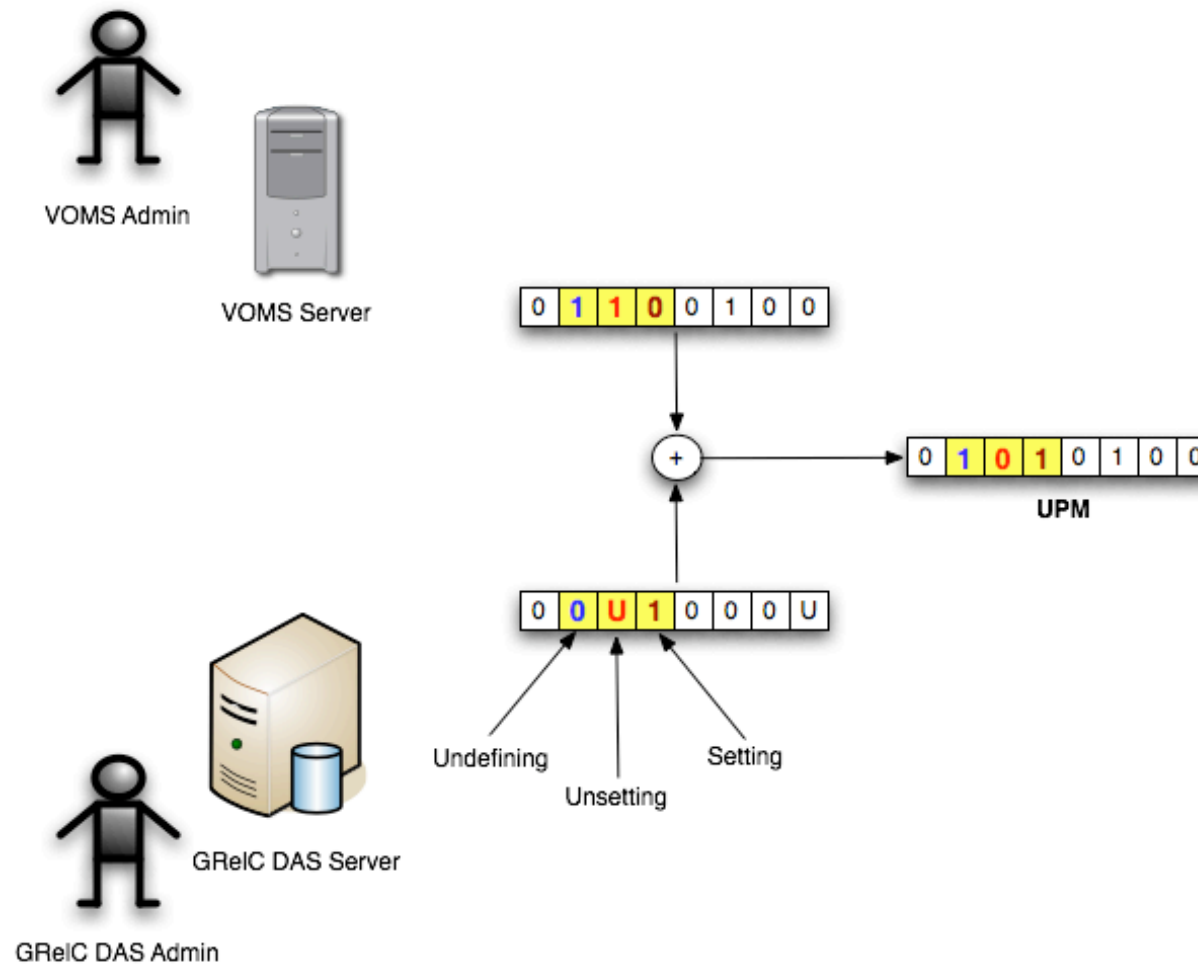
GRelC on gLite: A New Service

- Straightforward integration within the EGEE farm model
- GRelC DAS provides fine grained data mng service
- This service can be used both as farm service and as VO service depending on the context, the database policies/constraints

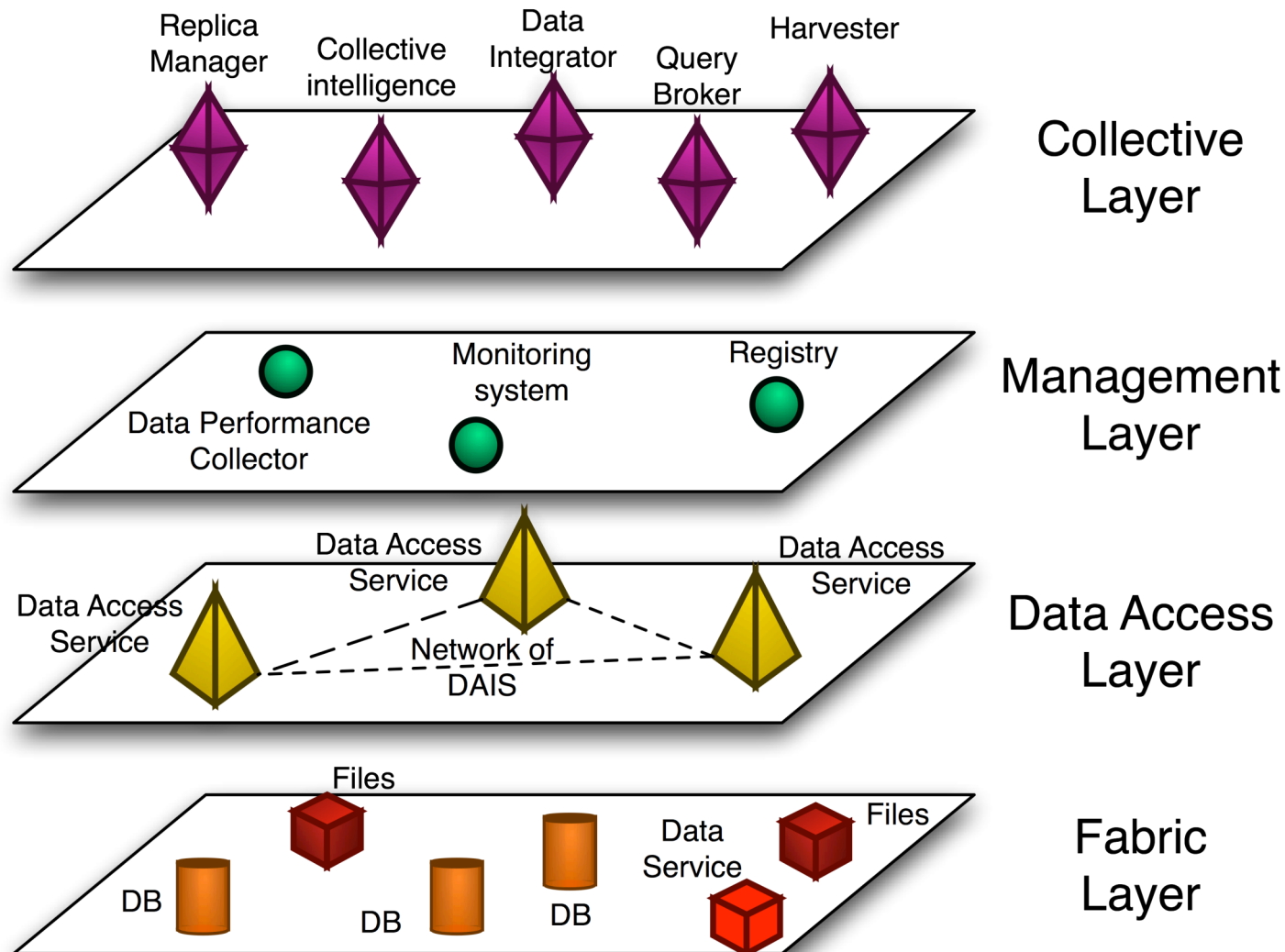
Extended EGEE
Farm Model



Security support



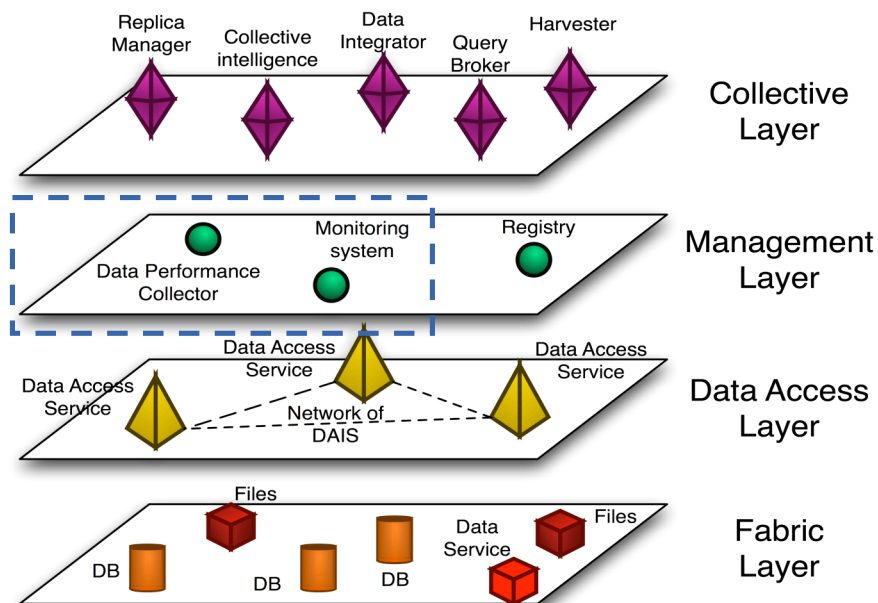
GRelC Data Management System Stack



S. Fiore, A. Negro, G. Aloisio, "The Data Access Layer in the GRelC System Architecture", Future Generation Computer System, Special Section Data Management for eScience.<http://dx.doi.org/10.1016/j.future.2010.07.006>

DoW (I) - “Management Layer”

- Management layer, control and monitoring functionalities available through a dashboard-based application and including charts, diagrams, tables, etc.
- Support to port in grid existing databases. It will be available through the GRelC Portal and will help users in avoiding to use the CLI



The screenshot shows the GRelC Portal interface, titled "AN EASY WAY TO MANAGE GRID DATABASES". The interface includes a navigation menu with links: Home, Components, Downloads, Documentation, News, Events, Publications, Staff, Deployment, and Portal. The main content area displays a "Synchronous Query Result" for a query executed on the "gandalf" server, using the "sakila" database. The query is "select * from actor". The result is a table with 20 rows, showing actor details.

actor_id	first_name	last_name	last_update
1	PENELOPE	GUINNESS	2006-02-15 04:34:33
2	NICK	WAHLBERG	2006-02-15 04:34:33
3	ED	CHASE	2006-02-15 04:34:33
4	JENNIFER	DAVIS	2006-02-15 04:34:33
5	JOHNNY	LOLLOBRIGIDA	2006-02-15 04:34:33
6	BETTE	NICHOLSON	2006-02-15 04:34:33
7	GRACE	MOSTEL	2006-02-15 04:34:33
8	MATTHEW	JOHANSSON	2006-02-15 04:34:33
9	JOE	SWANK	2006-02-15 04:34:33
10	CHRISTIAN	GABLE	2006-02-15 04:34:33
11	ZERO	CAGE	2006-02-15 04:34:33
12	KARL	BERRY	2006-02-15 04:34:33
13	UMA	WOOD	2006-02-15 04:34:33
14	VIVIAN	BERGEN	2006-02-15 04:34:33
15	CUBA	OLIVER	2006-02-15 04:34:33
16	FRED	COSTNER	2006-02-15 04:34:33
17	HELEN	VOIGHT	2006-02-15 04:34:33
18	DAN	TORN	2006-02-15 04:34:33
19	BOB	FAWCETT	2006-02-15 04:34:33
20	LUCILLE	TRACY	2006-02-15 04:34:33

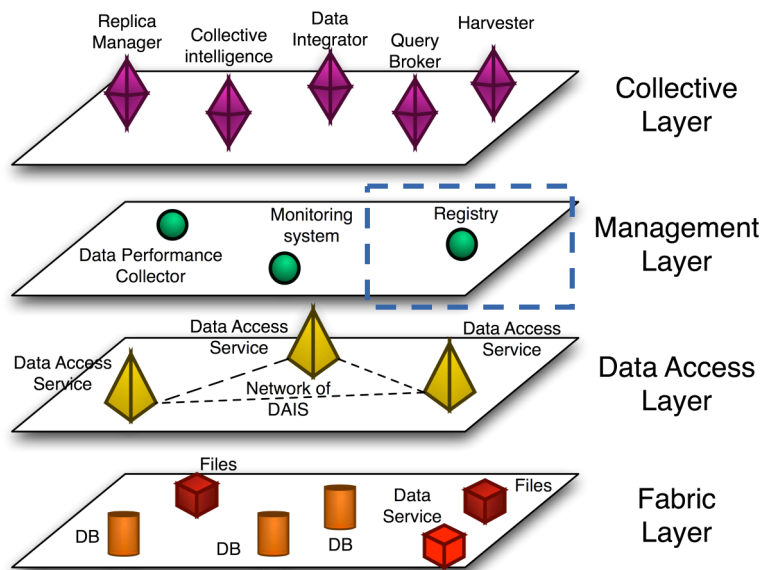
The interface also includes a "Logout" button and a footer with copyright information: "Copyright © SINAC Consortium 2007 - All rights reserved".

DoW (II) - “Management Layer”

- Registry database will store information about the available grid databases
- Cross-VO support that will help people in discovering existing data sources, join the database, search by tag, VO, etc.

and also

- post comments, rate resources, etc., that is creating a “community” around the available resources following Web2.0 approach

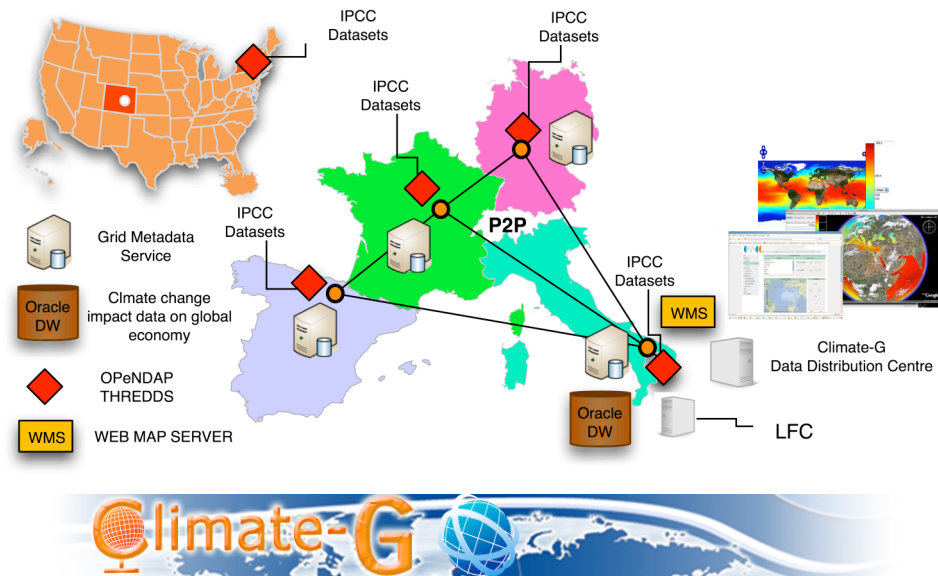
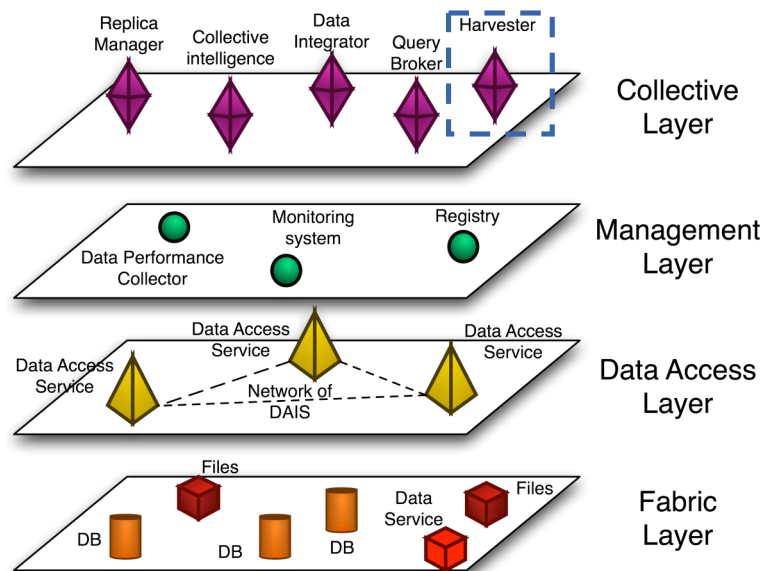


The screenshot shows the Global Registry web application. It includes a 'Settings' section with filters for Project Name, Host Name, V.O., Publish Date, and Tag. Below this is a 'Grid' section displaying a table of database records.

Database Name	Description	Publish Date	V.O.	Project Name	Host Name	Score
DB1	Bionformatics database	08/29/2009 16:35	Virtual Organization 1, Virtual Organization 2	GReIC	GOOGLE	★★★★★
DB2	Eath Science database	08/20/2010 16:34	Virtual Organization 1, Virtual Organization 3	GReIC	GOOGLE	★★★★★
DB3		09/04/2010 17:06	Virtual Organization 3	Priva	ESCHER	★★★★★

DoW (IV) - “User Support”

- As part of the SA3 activity will provide support to the user communities, defining some use cases and trying to support them as much as possible
- An example is the Climate-G testbed, an ongoing work started in 2008 in the context of the Earth Science Cluster Community
- A GReC service instance has been already deployed in our site to provide support also in terms of resources



Activities in the Q1 (I)

- Identification of the most important classes of information related to **management layer** (monitor and registry).
- First version of the **Entity/Relationship diagram** modelling such concepts and of the associated logical schema. First implementation (MySQL based) of the system database. Preliminary tests and bug fixing.
- Extension of the system database to include a **registry section**. New classes of information concerning the grid-databases managed by the GRelC services have been identified and modelled. An extended version of the E/R model, a new logical schema and the related MySQL implementation have been provided. This work represents a core part of the **registry database** described in the DoW.
- Design and preliminary implementation of the management **client** performing the **publication task** into the system database. The client has to take care of the databases and the GRelC services availability and status.

Activities in the Q1 (II)

- New release of the GReIC service (3.1.3)
 - bug fixing (memory leaks in the dime query have been removed)
 - log mechanism improved
- New release of the Climate-G portal addressing users' feedback and comments about the search & discovery interface (user support activity)
- A GReIC service has been deployed in our site to support LS database registration (user support activity)

Wednesday, 15th September 2010 11:16 AM +0200

GReIC

AN EASY WAY TO MANAGE GRID DATABASES

Home Components Downloads Documentation News Events lg_GReIC Portal (active)

Headlines

- 18-08-2010** GReIC DAIS version 3.1.3 is now available for gLite 4.x
- 16-11-2009** GReIC Portal v2.0 online today
- 16-11-2009** GReIC DAIS version 3.1.2 is now available for gLite 4.x
- 12-06-2009** GReIC & Climate-G
- 14-05-2009** CFP HDDM09 Workshop
- 08-04-2009** DaGreS09 Workshop Schedule
- 24-12-2008** CFP DaGreS09 Workshop
- 20-12-2008** GReIC Client API New release
- 13-10-2008** Climate Modelling and EGEE
- 29-09-2008** HPDataGrid08 Workshop Schedule

Events / Invited Lectures

- 09-11-2009** GReIC Tutorial @ ICITST09
- 09-11-2009** GReIC Talk @ ICITST09
- 22-06-2009** GReIC Demo @ HPCS09
- 08-05-2009** GReIC Tutorial @ GPC09 - Geneva, Switzerland
- 24-04-2009** GReIC Talk @ EGU - Vienna, Austria

Mission

The **Grid Relational Catalog Project (GReIC)** aims at providing a set of advanced data grid services to transparently, efficiently and securely manage Databases on the Grid.

The GReIC middleware is currently used within several grid research projects to support bioinformatics experiments on distributed and huge data banks as well as the metadata management related to Earth Observation System applications, etc.

At the moment the GReIC Data Access and Integration Service (GReIC DAIS) allows users to access and interact with different DBMSs both relational (PostgreSQL, MySQL, Oracle, DB2, SQLite, etc) and non-relational (eXist, XIndex, XML flat files). It provides an uniform access interface to heterogeneous data sources.

The GReIC DAIS is now part of the **INFN-GRID** release. Clients are installed on both UI & WN within the Italian Grid. The GReIC DAIS server is available as **lg_GReIC profile**. The installation and configuration procedure is full YAIM compliant.

The GReIC DAIS Service is fully compatible with the gLite and Globus middlewares and it has been ported on several platforms (Scientific Linux Cern 3 and Scientific Linux Cern 4, SUSE Linux, FreeBSD, Mac OS X, both IA32 and IA64).

The GReIC middleware has been included in the EGEE **RESPECT** Program (Recommended External Software Packages for EGEE Communities) since it works well in concert with the EGEE gLite software by expanding the functionality of the grid infrastructure (w.r.t. database management in grid).

The GReIC DAIS is currently adopted as grid metadata management service in the **Climate-G testbed** to enable geographical data sharing, search and discovery activities. Moreover it is currently used at the Euro-Mediterranean Centre for Climate Change to manage climate metadata across the CMCC data grid infrastructure through the CMCC Data Distribution Centre portal

[Top]

Climate-G

HOME

ABOUT

LOGIN (CERT NEEDED)

SEARCH (CERT NEEDED)

- BASE SEARCH
- ADVANCED SEARCH
- DATASET SEARCH

DASHBOARDS (CERT NEEDED)

- DASHBOARD

TOOLS/SERVICES

USAGE

LOGOUT

Home

Climate-G is a distributed testbed for climate change addressing challenging data and metadata management issues at a very large scale. The main scope of Climate-G is to allow scientists to carry out geographical and cross-institutional data discovery, access, visualization and sharing of climate data.

The Climate-G testbed is the result of an open, successful and wide collaboration joining grids and P2P paradigm, OGC services, visualization tools, etc. To enable geographical data sharing, search and discovery activities (through the Climate-G data grid portal interface) we adopted a distributed CMCC metadata solution leveraging P2P and grid technologies, the GReIC Data Access and Integration Service.

The Climate-G testbed provides a proof of concept concerning the involved technologies and right now it manages about 2TB of data provided by IPSL and University of Cantabria. Other datasets come from the IPCC website (AR4). Additional data from CMCC will be soon added to the digital library.

The Climate-G Data Distribution Centre is the data grid portal of the testbed and it is intended for scientists and researchers that want to carry out search and discovery activities on the available large scale digital library. It provides a ubiquitous and pervasive way to ease data publishing, metadata search & discovery, metadata annotation and validation, data access, etc.

The Climate-G data portal security model includes the use of HTTPS protocol for secure communication with the client (based on X509v3 certificates that must be loaded into the browser), secure cookies to establish and maintain user sessions as well as a complete role-based authorization system.

Foreseen Activities in Year1

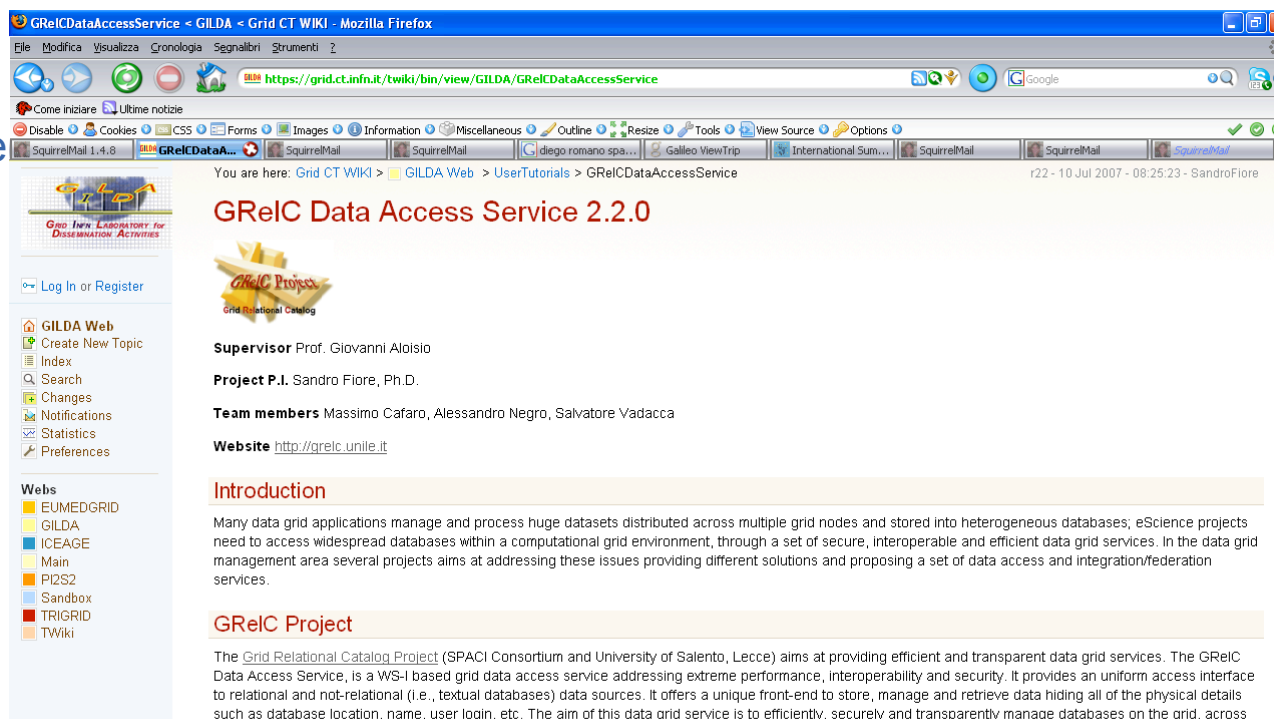
- New release of the GRelC service supporting the registration functionality (at catalog, service and CLI level)
- New release of the GRelC Portal:
 - supporting the “porting in grid” of existing data sources
- Design and preliminary implementation of the monitoring and registry interface to implement the web application related to the “Management layer”
- Implementation of the client devoted to the “Management layer”.
- User support
 - map-enabled distributed search and discovery interface in the Climate-G Portal
 - search & discovery by model and scenario added to the Climate-G Portal
 - upgrade of the GRelC service devoted to user support activities
 - participation in “user community oriented” activities (i.e. ES, LS), initiatives and conferences (AGU2010, EGU2011, etc.)

User support: tutorials on GILDA

GReIC DAS User Tutorial on GILDA Grid CT Wiki Website

Info about:

- Log in to the grid
- Query Submission



For any information about GILDA t-Infrastructure please
contact roberto.barbera@ct.infn.it & grid-prod@ct.infn.it

GReIC DAS Tutorial link:

<https://grid.ct.infn.it/twiki/bin/view/GILDA/GReICProject>

**Special thanks to the GILDA
Staff for their support**

User support: the GRelC WebSite



Main sections:

- Download (rpms available)
- News
- Publications
- Events
- Deployment
- Documentation
- Components
-

GRelC Website

URL: <http://grelec.unile.it/>

Mailing List

mail: grelec-user@sara.unile.it

User support: the GRelC Portal

Functionalities:

- Login
- GRelC DAS Registration
- Host Management
- Instance Management
- User Management
- Query submission
- Deployment Map

The screenshot displays the GRelC Portal interface. At the top, it says 'Wed Sep 26 13:53:24 CEST 2007'. The main header features the 'GRelC' logo and the tagline 'AN EASY WAY TO MANAGE GRID DATABASES'. A navigation bar includes links: Home, Components, Downloads, Documentation, News, Events, Publications, Staff, Deployment, and Portal (BETA). The left sidebar contains a 'Query' section with a 'Synchronous Query' button, a 'GRelC DASs' section with a 'Servers List' (showing gandalf, grelc02, grelc01), and an 'Operations - gandalf' section with 'System Users', 'Hosts', 'Instances', 'Grid Database Users', and 'Virtual Organizations'. The main content area shows a 'Synchronous Query Result' for the query 'select * from actor'. It includes a table with columns 'actor_id', 'first_name', 'last_name', and 'last_update', listing 20 actors. Navigation links 'Previous', '1', '2', '3', '4', '5', '6', '7', '8', '9', '10', and 'Next' are at the bottom of the table. A 'Back' button is also present. The footer contains the user's DN and a copyright notice for SPACI Consortium 2007.

actor_id	first_name	last_name	last_update
1	PENELOPE	GUINNESS	2006-02-15 04:34:33
2	NICK	WAHLBERG	2006-02-15 04:34:33
3	ED	CHASE	2006-02-15 04:34:33
4	JENNIFER	DAVIS	2006-02-15 04:34:33
5	JOHNNY	LOLLOBRIGIDA	2006-02-15 04:34:33
6	BETTE	NICHOLSON	2006-02-15 04:34:33
7	GRACE	MOSTEL	2006-02-15 04:34:33
8	MATTHEW	JOHANSSON	2006-02-15 04:34:33
9	JOE	SWANK	2006-02-15 04:34:33
10	CHRISTIAN	GABLE	2006-02-15 04:34:33
11	ZERO	CAGE	2006-02-15 04:34:33
12	KARL	BERRY	2006-02-15 04:34:33
13	UMA	WOOD	2006-02-15 04:34:33
14	VIVIEN	BERGEN	2006-02-15 04:34:33
15	CUBA	OLIVIER	2006-02-15 04:34:33
16	FRED	COSTNER	2006-02-15 04:34:33
17	HELEN	VOIGHT	2006-02-15 04:34:33
18	DAN	TORN	2006-02-15 04:34:33
19	BOB	FAWCETT	2006-02-15 04:34:33
20	LUCILLE	TRACY	2006-02-15 04:34:33

Features:

- **Seamless** and **ubiquitous** access to GRelC DAS enabled resources
- **No additional software** installation / configuration is required
- **Complete** and **user-friendly** Grid Data Portal Interface (It entirely replaces CLI)

GRelC Portal: Some Snapshots

Wed Sep 26 13:45:10 CEST 2007

Sun Sep 30 17:35:15 CEST 2007

GREL C

AN EASY WAY TO MANAGE GRID DATABASES

Home Components Downloads Documentation News Events Publications Staff Deployment Portal (BETA)

Query

Synchronous Query

GRelC DASS

Servers List

gandalf
grelc02
grelc01

Operations - gandalf

System Users

Hosts

Instances

Grid databases

Grid Database Users

Virtual Organizations

Association

Session

Logout

Your DN is CN=grelc04.unile.it,L=HPCC University of Lecce

Your DN is CN=Salvatore Vadaccaro

Synchronous Query Result

Server - gandalf

Database - sakila

Query - select * from actor

actor_id	first_name	last_name	last_update
1	PENELOPE	GUINNESS	2006-02-15 04:34:33
2	NICK	WAHLBERG	2006-02-15 04:34:33
3	ED	CHASE	2006-02-15 04:34:33
4	JENNIFER	DAVIS	2006-02-15 04:34:33
5	JOHNNY	LOLLOBRIGIDA	2006-02-15 04:34:33
6	BETTE	NICHOLSON	2006-02-15 04:34:33
7	GRACE	MOSTEL	2006-02-15 04:34:33
8	MATTHEW	JOHANSSON	2006-02-15 04:34:33
9	JOE	SWANK	2006-02-15 04:34:33
10	CHRISTIAN	GABLE	2006-02-15 04:34:33
11	ZERO	CAGE	2006-02-15 04:34:33
12	KARL	BERRY	2006-02-15 04:34:33
13	UMA	WOOD	2006-02-15 04:34:33
14	VIVIEN	BERGEN	2006-02-15 04:34:33
15	CUBA	OLIVIER	2006-02-15 04:34:33
16	FRED	COSTNER	2006-02-15 04:34:33
17	HELEN	VOIGHT	2006-02-15 04:34:33
18	DAN	TORN	2006-02-15 04:34:33
19	BOB	FAWCETT	2006-02-15 04:34:33
20	LUCILLE	TRACY	2006-02-15 04:34:33

Previous 1 2 3 4 5 6 7 8 9 10 Next

Back

Your DN is CN=grelc04.unile.it,L=HPCC University of Lecce,OU=Host,O=INFN,C=IT - Your proxy is valid until 27/09/2007 01:57:21

Copyright © SPACI Consortium 2007 - All rights reserved

For any information



Supervisor: Prof. Giovanni Aloisio (giovanni.aloisio@unile.it)

Project P. I.: Ph. D. Sandro Fiore (sandro.fiore@unile.it)

GRelC WebSite: <http://grelc.unile.it>

GRelC Portal: https://gandalf.unile.it:8443/GRelCPortal-v2.0/upload_proxy.jsp

GILDA support: <https://grid.ct.infn.it/twiki/bin/view/GILDA/GRelCProject>

Mailing lists: grelc-user@sara.unile.it