



The Grid Relational Catalog Project

GRelC

Sandro Fiore, Ph.D.

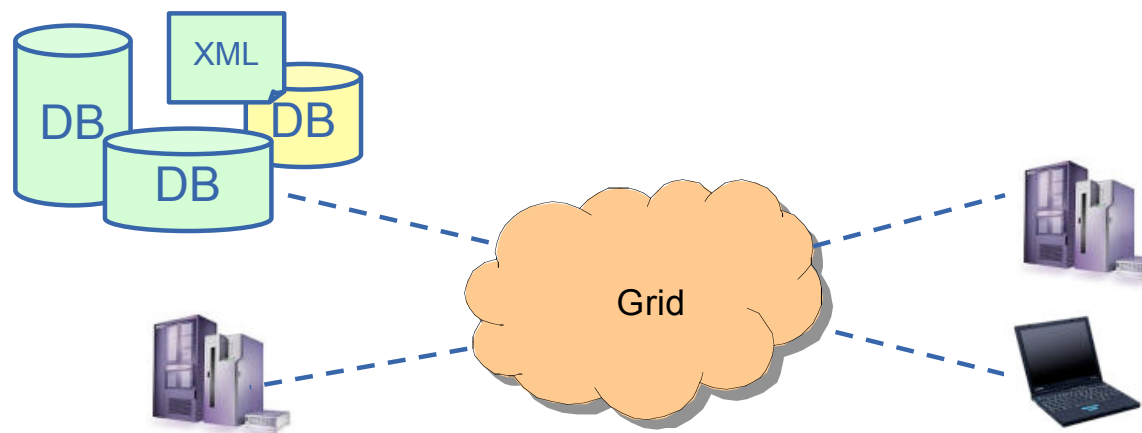
University of Salento (Lecce) and SPACI Consortium - Italy

- Motivations
- GReIC Project
- GReIC DAS
 - *Architecture*
 - *Internal Components*
 - *GUI*
 - *Queries*
 - *Experimental Results*
 - *Porting on gLite*
 - *On Line User Tutorial (GILDA)*
 - *GReIC Portal*
- GReIC DAIS
 - *Architecture extensions*
 - *GReIC Portal extensions*
- A case study: Euro-Med Centre for Climate Change
- European testbed
- Conclusions

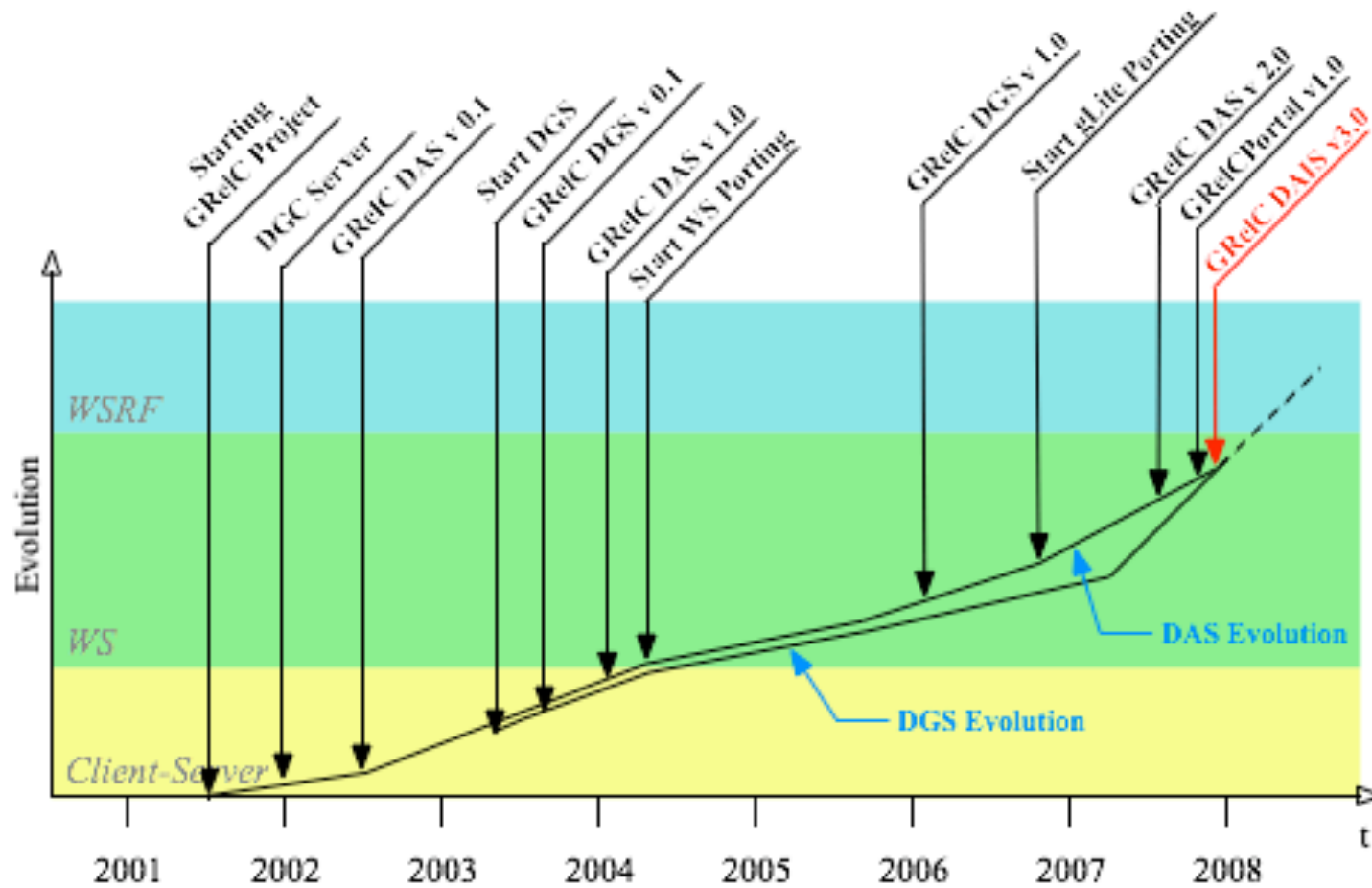
- Data Grids should provide a low level framework also for grid-database management (fine grained approach)
- Few research efforts devoted to grid-db management in 2001
- No new DBMS or new query language
- Legacy systems/databases and standard SQL
- Need to deal with heterogeneity of data sources, formats, model.
- Need for more complex and efficient *queries* in grid
- Compatibility/Integration with existing production grid environments (based on gLite, Globus, etc.)
- Main topics: Integration, access and management
- Requirements: security, transparency, interoperability, efficiency, robustness, scalability, etc.

Introducing the GReIC Project

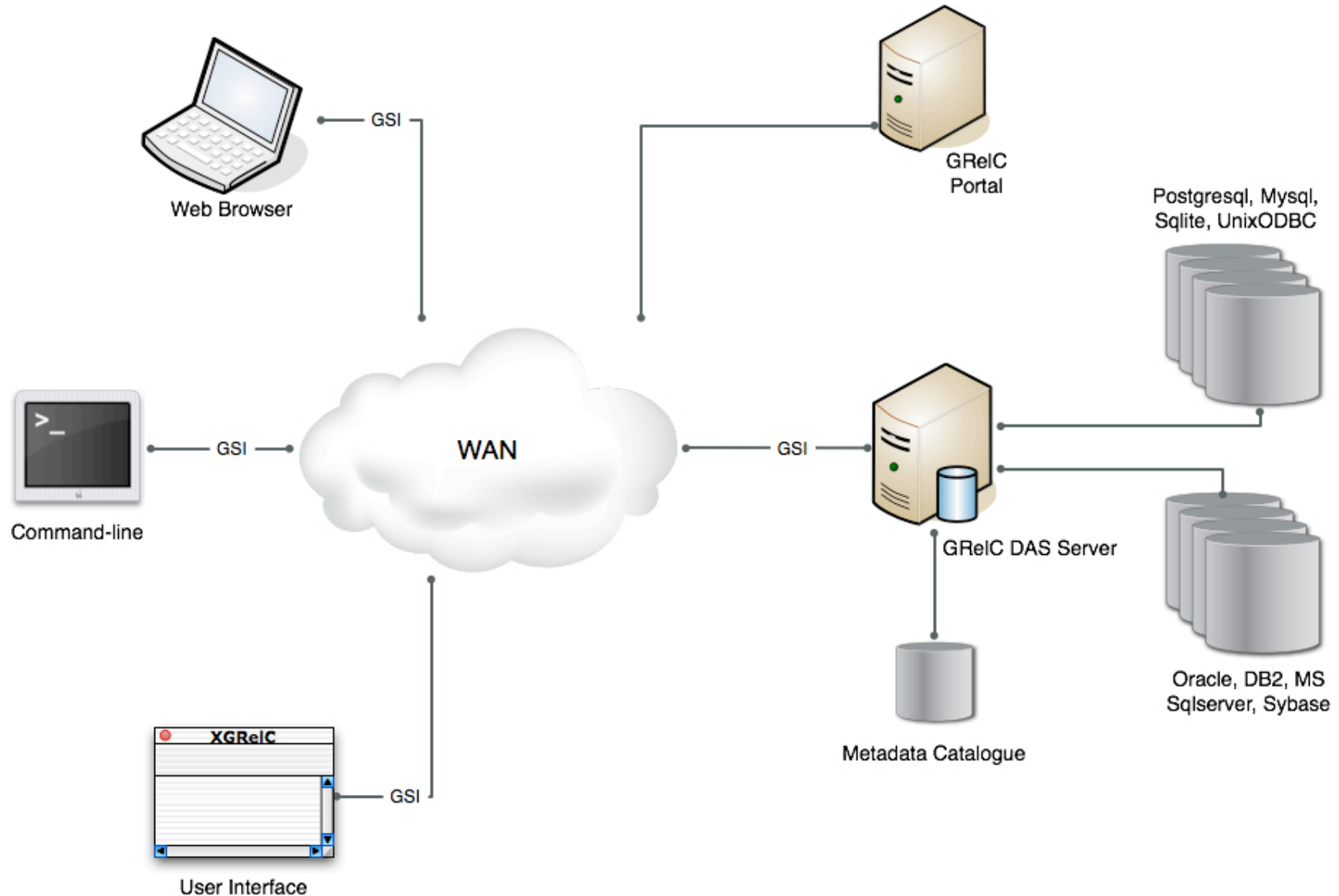
- Grid Relational Catalog is a project which aims at designing and developing a set of efficient, secure and transparent Data Grid Services (Starting date, January 2001).
- **GReIC Data Access Service** aims at providing a large set of functionalities to access both relational and non relational DataBases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



GRelC Project: a bit of history



GReIC DAS Architecture (in the large)





GReIC DAS: Main Features

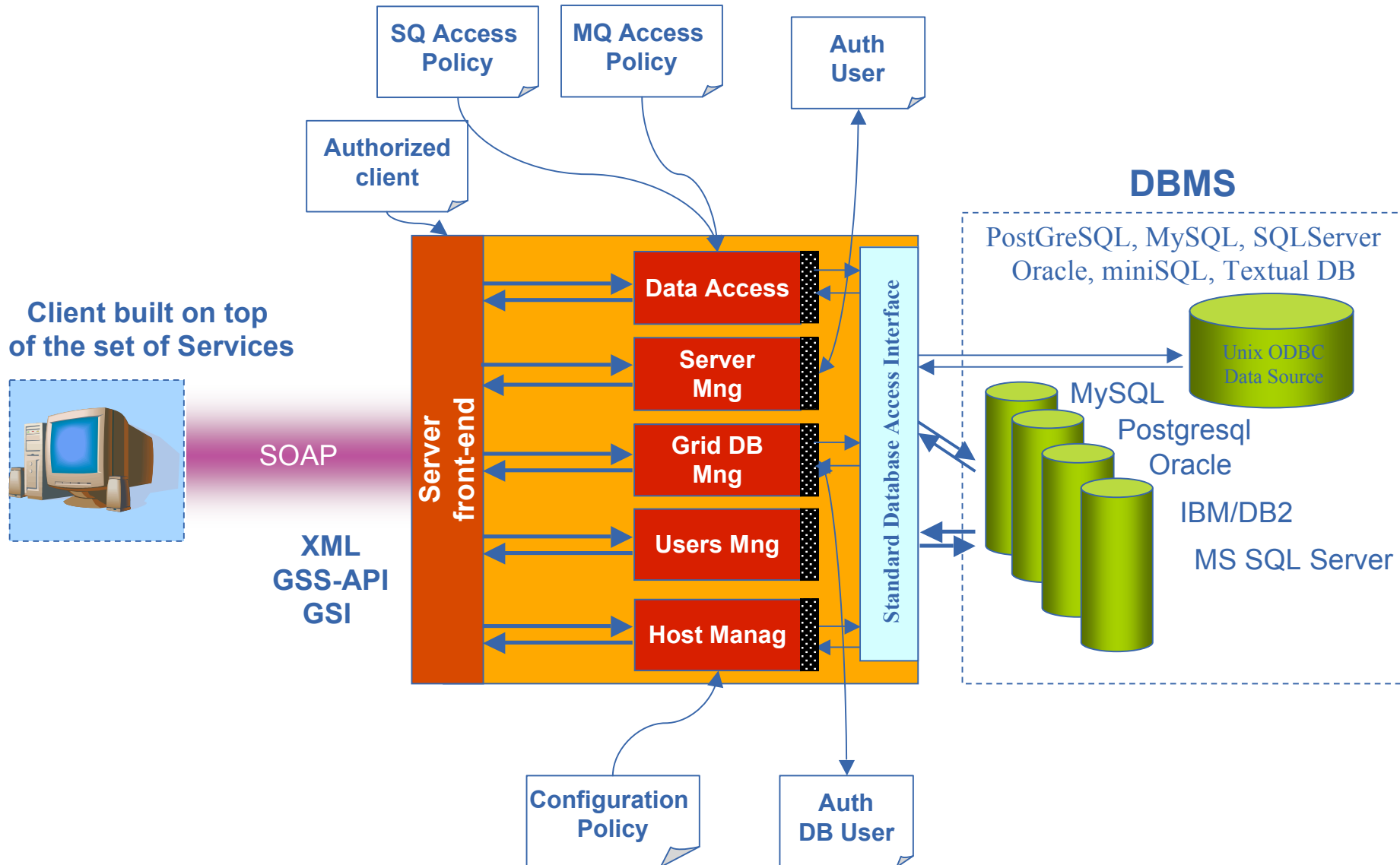
- Entirely based on C programming language (server side)
- Multithreaded web service
 - Next releases will be pre-threaded
- It exposes the **web service interface GSI enabled** and WS-I compliant
 - Compatibility with XML, SOAP, WSDL standards
- **Mutual authentication** based on GSI (X.509v3 digital 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

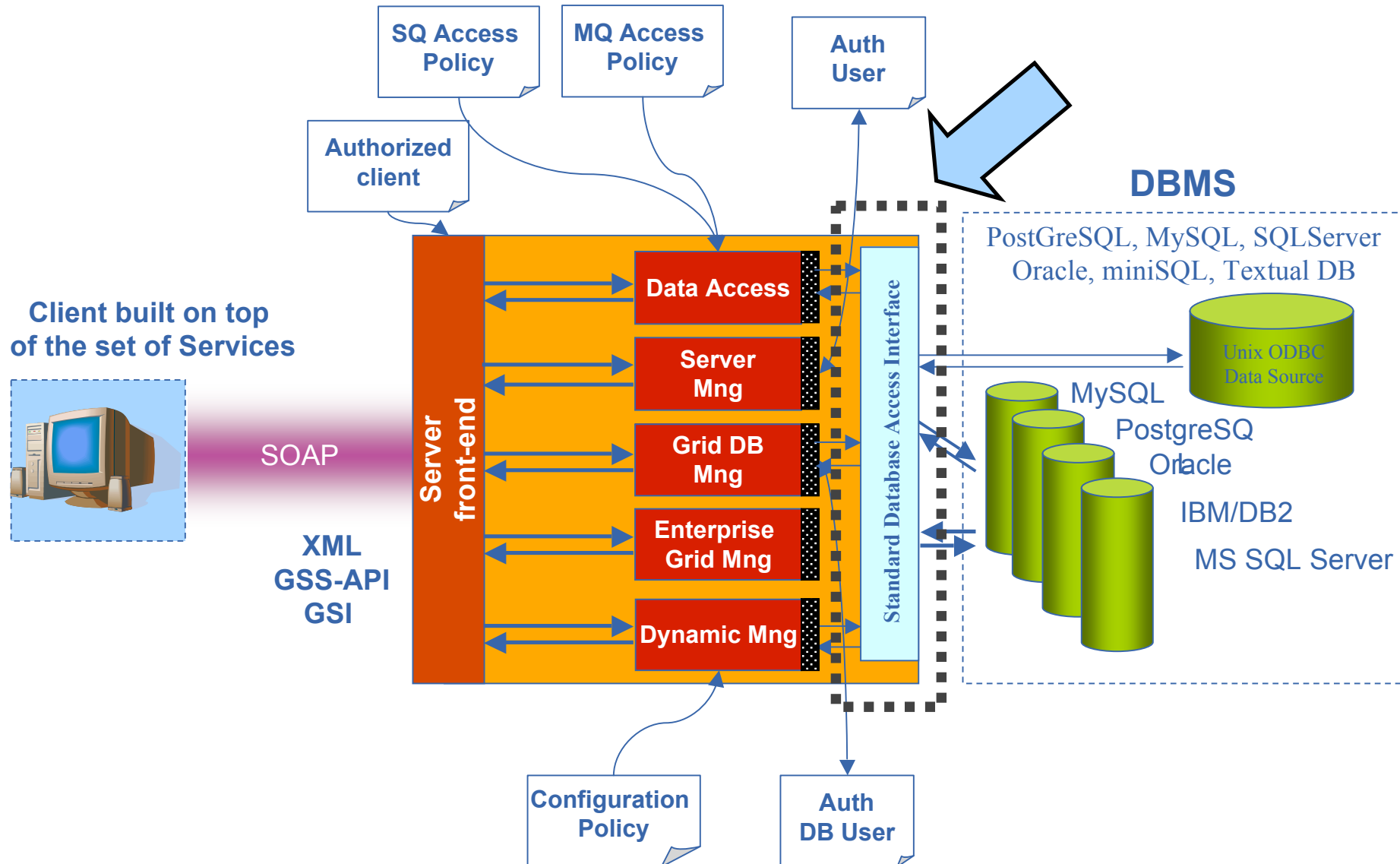


GReIC DAS: Main Features

- 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

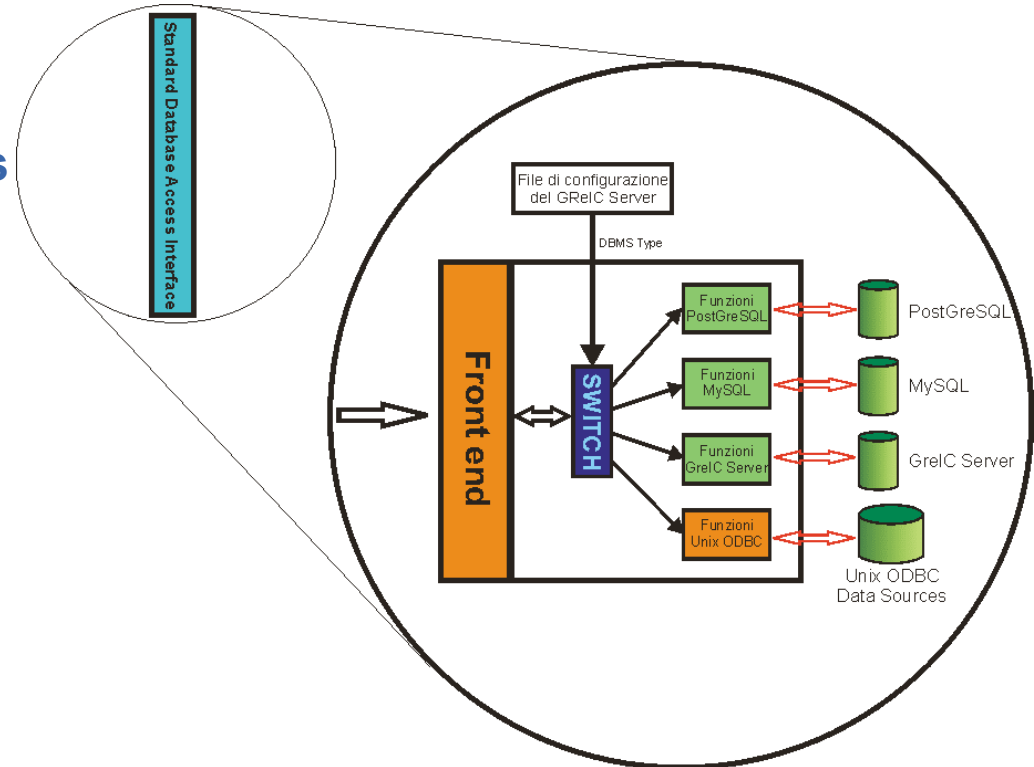
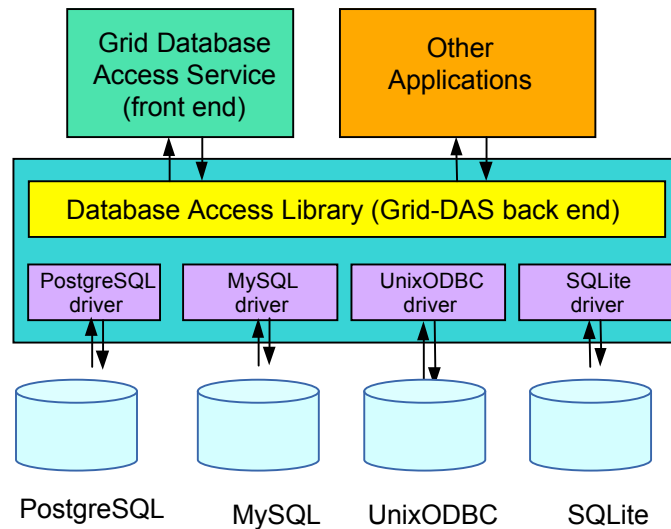
First Design fo the GReIC DAS Arch.





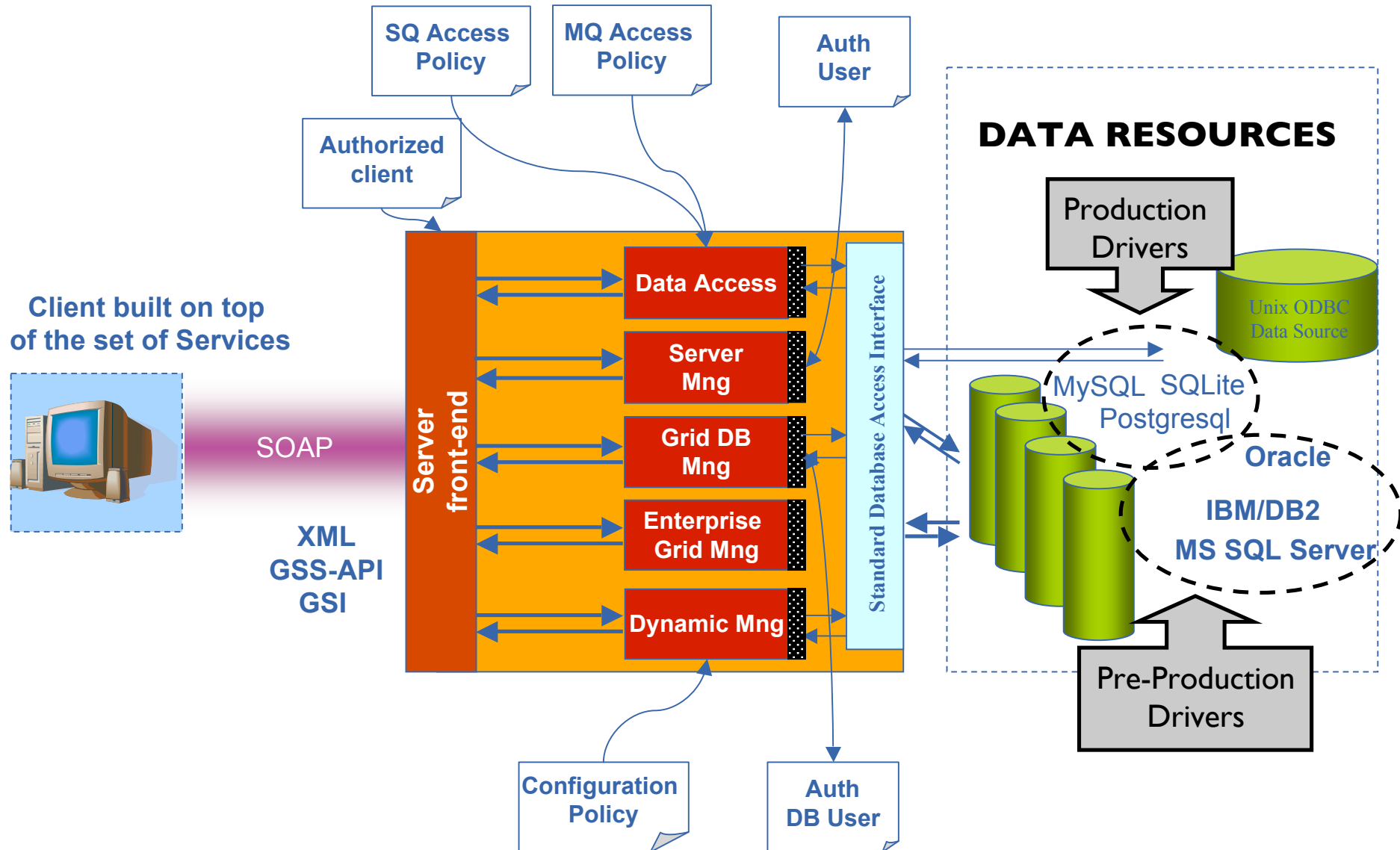
Features:

- Standard access to data sources
- Types uniformity
- Error uniformity
- Plug-in architecture based on dynamic libraries



- Dynamic binding to:
PostgreSQL MySQL SQLite
IBM/DB2, Oracle9.i, MS-SQL Server,
UnixODBC, Textual DBs, etc.

New drivers: IBM/DB2, Oracle, MS-SQL



- `int grelc_sdai_handle_set_grelc_dbname (grelc_sdai_handle *handle, const char *value);`
- `int grelc_sdai_bind(grelc_sdai_handle* handle);`
- `int grelc_sdai_unbind(grelc_sdai_handle* handle);`
- `int grelc_sdai_init();`
- `int grelc_sdai_exit();`
- `int grelc_sdai_query_submission(grelc_sdai_handle* handle, char* query);`
- `int grelc_sdai_ntuples(grelc_sdai_handle* handle);`
- `int grelc_sdai_nfields(grelc_sdai_handle* handle);`
- `int grelc_sdai_field_name(char** field, grelc_sdai_handle* handle, int i);`
- `int grelc_sdai_field_type(grelc_sdai_handle* handle, int i);`



GReIC SDAI Library: APIs (II)

- `int grelc_sdai_get_value(char** value, grelc_sdai_handle* handle, int i, int j);`
- `int grelc_sdai_clear_result(grelc_sdai_handle* handle);`
- `int grelc_sdai_lock(grelc_sdai_handle* handle, int mode, char* table);`
- `int grelc_sdai_unlock(grelc_sdai_handle* handle);`
- `int grelc_sdai_begin_transaction(grelc_sdai_handle *handle);`
- `int grelc_sdai_commit_transaction(grelc_sdai_handle *handle);`
- `int grelc_sdai_rollback_transaction(grelc_sdai_handle *handle);`
- `int grelc_sdai_get_tables (grelc_sdai_handle *handle);`
- `int grelc_sdai_get_fields (grelc_sdai_handle *handle, char* table);`

SDAI Client

```

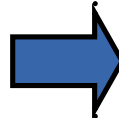
if (res = grelc_sdai_bind (&handle))
{
    fprintf (stderr, "ERROR! Database bind failed [Code: %d]\n", res);
    ...
    return -1;
}

if (grelc_sdai_query_submission (&handle, query))
{
    fprintf (stderr, "ERROR! Query submission failed!\n");
    ...
    return -2;
}

if (strcasestr (query, "SELECT"))
{
    for (outer = 0; outer < grelc_sdai_ntuples (&handle); outer++)
    {
        for (inner = 0; inner < (grelc_sdai_nfields (&handle)); inner++)
        {
            grelc_sdai_get_value (&field, &handle, outer, inner);
            printf ("%20s\t\t", field);
        }
        printf ("\n");
    }
}

grelc_sdai_clear_result(&handle);

if (grelc_sdai_unbind (&handle))
{
    fprintf (stderr, "ERROR! Database unbind failed!\n");
    ...
    return -1;
}
    
```



grelc_sdai_bind (MySQL Impl.)

```

int dbms_bind(grelc_sdai_handle *handle)
{
    MYSQL      *mysql = NULL;
    ....

    if (!(mysql = mysql_init (NULL)))
        return SDAI_MEMORY_ERR;

    ....

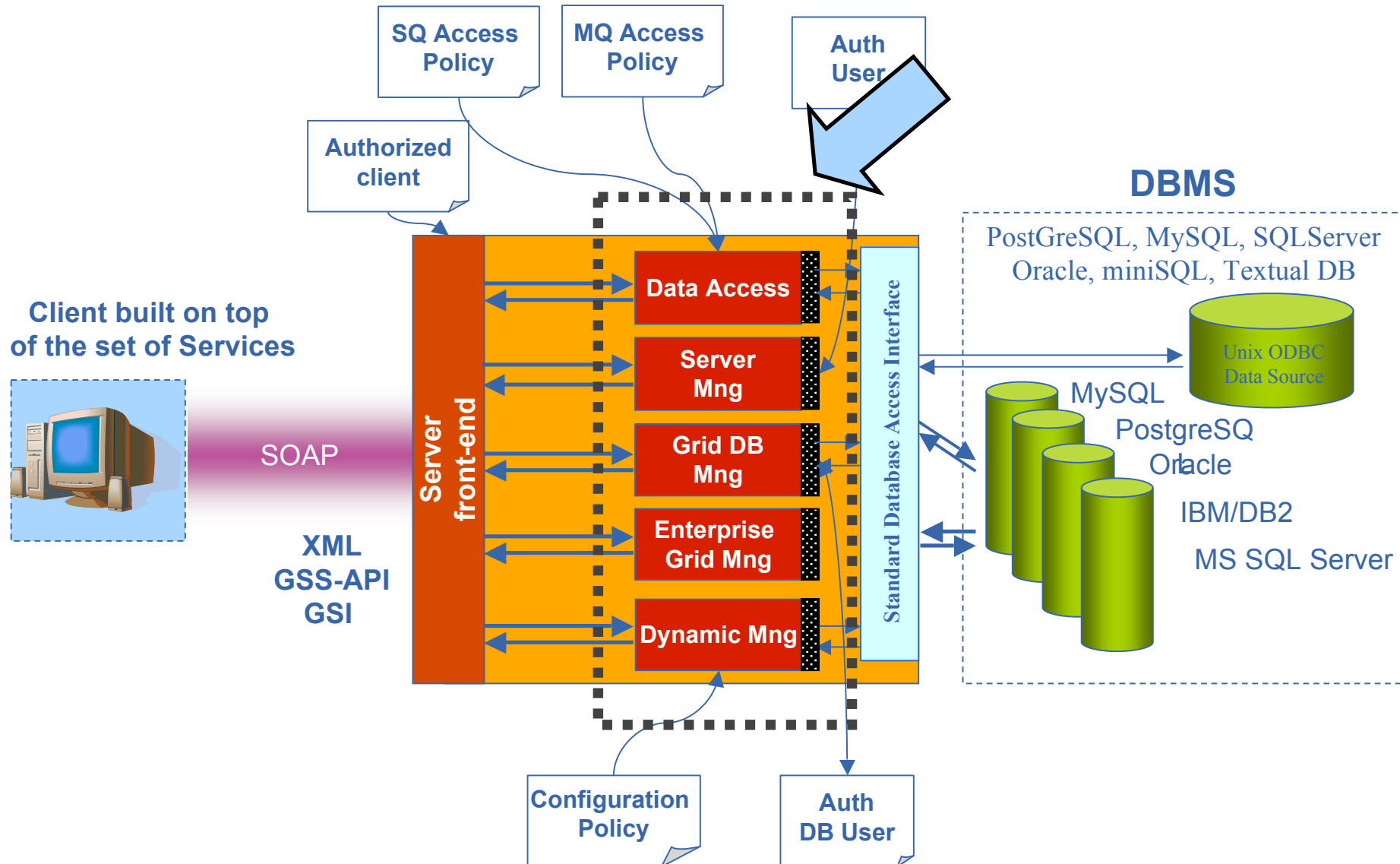
    if (!mysql_real_connect
        (mysql, handle->db_host_name, handle->db_login, handle->db_password,
         handle->db_name, (unsigned int)atoi(handle->db_port), NULL, 0)
        )
    {
        mysql_close(mysql);
        return SDAI_COMMAND_ERROR;
    }

    if (!(handle->dbms_handle = (mysql_dbms_handle*)calloc(1,
        sizeof(mysql_dbms_handle))))
    {
        mysql_close(mysql);
        return SDAI_MEMORY_ERR;
    }

    ....

    return SDAI_SUCCESS;
}
    
```

GReIC DAS: Internal Components



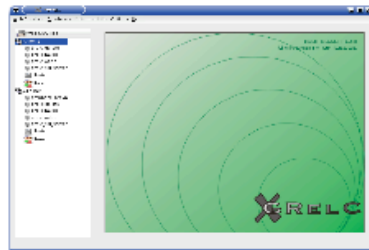
The GReIC Library: APIs Classification

- **Database access and query services**
 - bind
 - unbind
 - query submission
- **Remote manipulation services**
 - get_value
 - get_current_tuple
- **Resultset store and retrieving services**
 - store_result_disk
 - fetch_stored_recordset
- **User management services**
 - add_user
 - remove_user
 - set_user_policy
- **Enterprise Grid management services**
 - add_host
 - add_dbms
- **Virtual space management services**
 - create_virtual_database
 - register_database
- **QoS services**
 - relocate_database

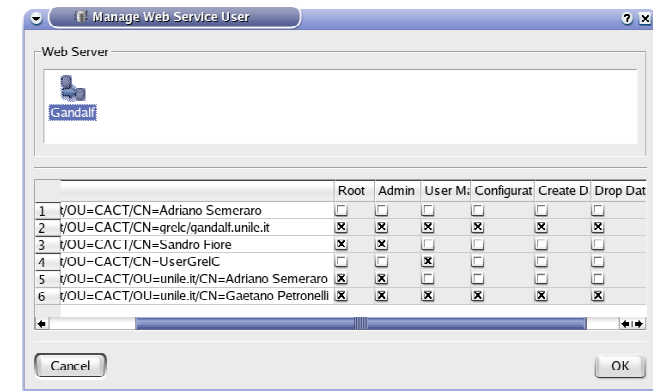
Wide SDK both for C
and Java developers

GReIC-Proxy: a C++ Module

A C++ module was created in order to allow an easy development of new web services client with this language.



This module hides the communication level with web services.



We easily developed a graphical user interface based on the Proxy class and Qt libraries.

```
class CppProxy
```

```
{
```

```
public:
```

```
int bind (string grelc_db_name);
```

```
int unbind ();
```

```
int query_submission (string query);
```

```
....
```

```
int create_database (string grelc_db_name, string identity, string dbms, string host, string istance, int log_type);
```

```
int create_physical_database_and_register(string grelc_db_name, string dbms, string host, string istance, int log_type);
```

```
int drop_database (string grelc_db_name);
```

```
int get_log(int num_lines, string &log);
```

```
int get_log_database(string grelc_dbname, int num_lines, string &log);
```

```
int get_host_position_info(HostInfoResponse &response);
```

```
int get_value (int row, int column, string &value);
```

```
private:
```

```
struct soap *soap;
```

```
struct gsi_plugin_data *data;
```

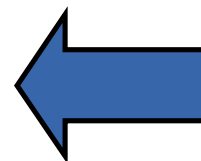
```
char* connection;
```

```
bool connected;
```

```
bool enable_credential;
```

```
string dn;
```

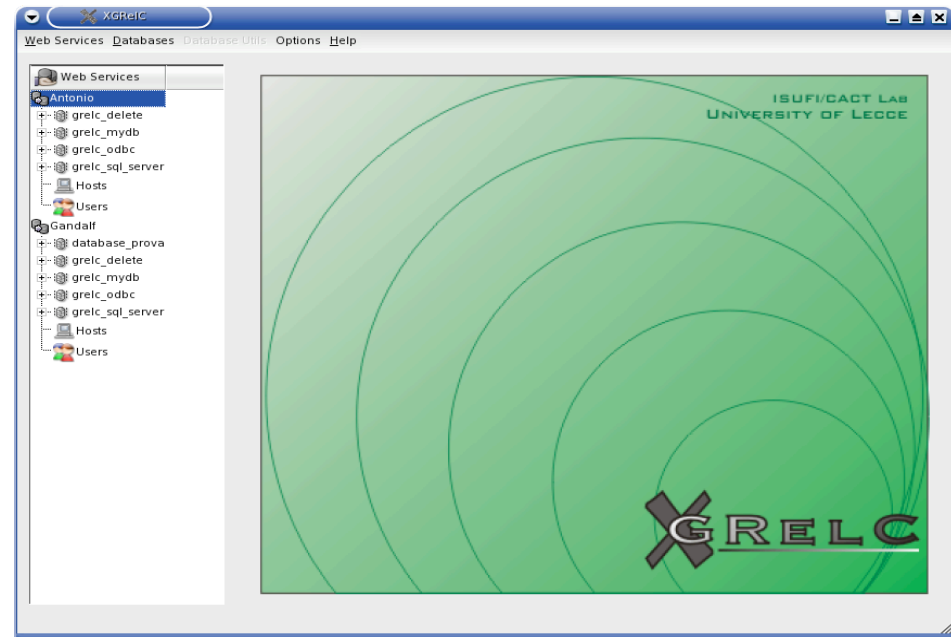
```
};
```



**Low level details concerning
with soap, gsi, conn. status
are concealed**

Functionalities:

- User management
- Web Service registration
- Host Management
- Logging
- DBMS configuration
- Database creation
- Import Database
- Database configuration
- Query submission
- Map deployment
- GUI no longer supported -> Migration to Web-based support
 - See **GReIC Portal** in the next slides

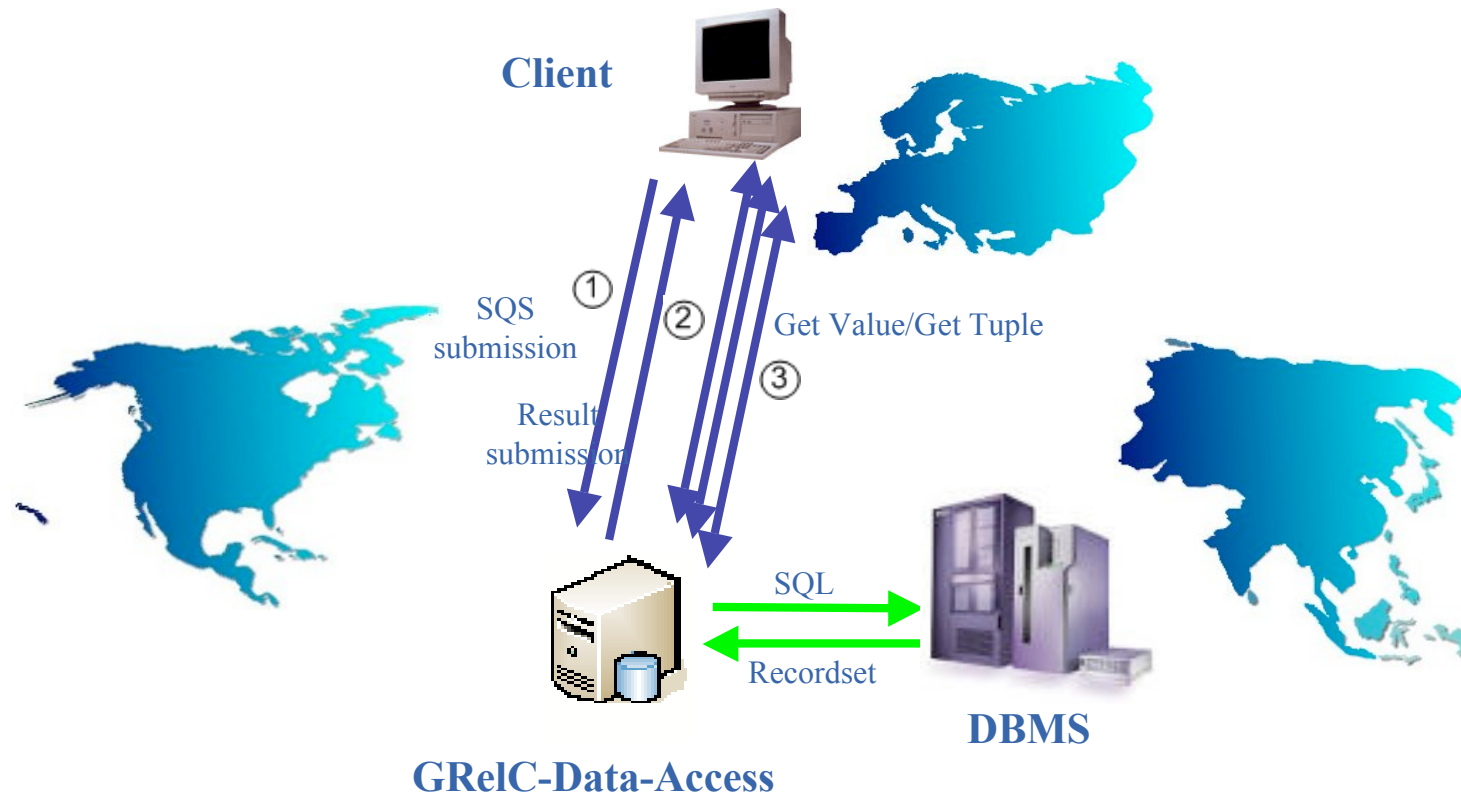


We developed the following kinds of queries:

- Single Query Online (*no longer available*)
- Single Query Memory (+ chunk management)
- Single Query File (+ chunk management)
- Single Query File + ZIP (+ chunk management)
- Single Query Prefetch (parallel chunk donwload/processing)
- Single Query Stream (resultset streaming)
- Web Single Query XHTML (+ chunk management / paging)
 - CSS v2.0, XHTML v1.0 Strict

Results displayed in the following formats:

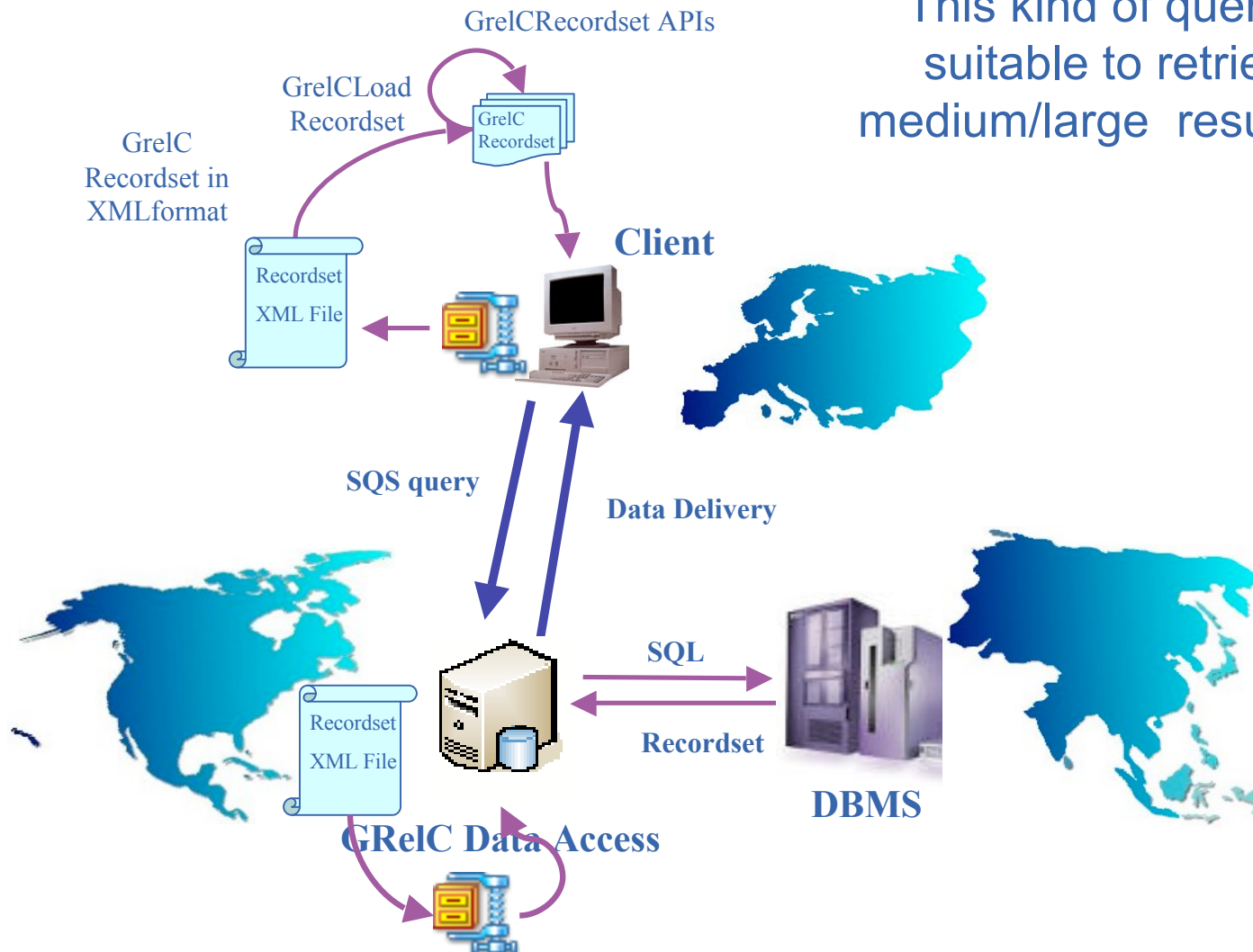
- Tabular
- XML
- HTML
- RAW



This kind of query is suitable for DML statements or to retrieve small resultsets.

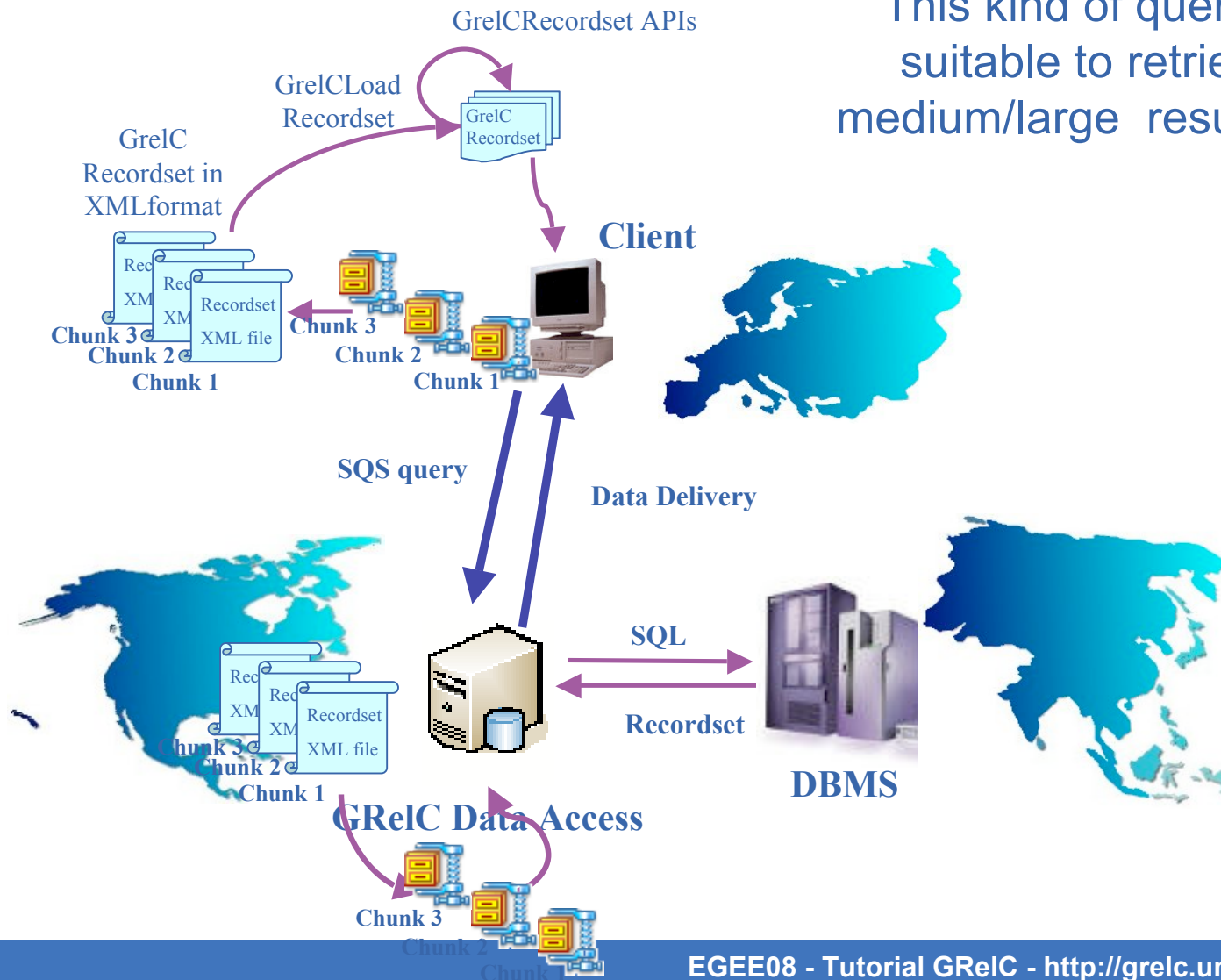
Single Query File Approach (Zip)

This kind of query is suitable to retrieve medium/large resultsets



Single Query File + chunk (Zip)

This kind of query is suitable to retrieve medium/large resultsets





GReIC Data Access: Clients

```

grelc@gandalf:~/PRODUCTION/access/bin> ./grelc_service_submission dime -Q "select * from person" -D phonebook -x
[Sat Dec 16 00:21:11 2006] GSI plugin for gSOAP v2.6a: Established security context with: /O=Grid/OU=unile.it/OU=
description: 1..12552 /source=
<RECORDS>
  <RECORD>
    <id>1</id>
    <name>John</name>
    <surname>Smith</surname>
    <address>1, Piccadilly Circus - London</address>
    <phonenumbers>+4412309876543</phonenumbers>
  </RECORD>
  <RECORD>
    <id>2</id>
    <name>Tony</name>
    <surname>Blair</surname>
    <address>10, Downing Street - London</address>
    <phonenumbers>+4412313667841</phonenumbers>
  </RECORD>
  <RECORD>
    <id>3</id>
    <name>Alessandro</name>
    <surname>Negro</surname>
    <address>Via Ungheria,16</address>
    <phonenumbers>3388781201</phonenumbers>
  </RECORD>
</RECORDS>
grelc@gandalf:~/PRODUCTION/access/bin>

```

```

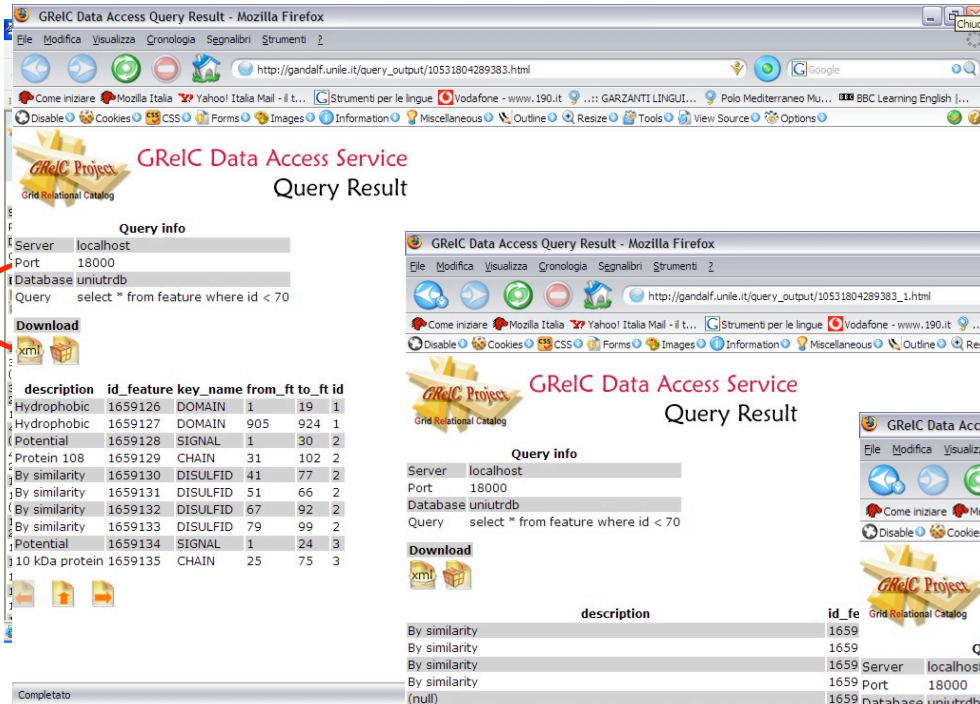
967 35 1 6147 5.99 2005-07-11 09:13:08 2006-02-15 22:12:37
968 35 1 6401 4.99 2005-07-11 22:44:34 2006-02-15 22:12:37
969 35 1 6565 4.99 2005-07-12 05:39:50 2006-02-15 22:12:37
970 35 1 6572 4.99 2005-07-12 05:56:38 2006-02-15 22:12:37
971 35 1 7140 4.99 2005-07-27 06:54:12 2006-02-15 22:12:37
972 35 1 8822 6.99 2005-07-29 22:20:21 2006-02-15 22:12:37
973 35 1 8971 5.99 2005-07-30 04:03:58 2006-02-15 22:12:37
974 35 2 9033 2.99 2005-07-30 06:07:42 2006-02-15 22:12:37
975 35 1 9579 6.99 2005-07-31 02:59:20 2006-02-15 22:12:37
976 35 1 11298 1.99 2005-08-02 15:32:32 2006-02-15 22:12:37
977 35 1 11452 7.99 2005-08-02 20:59:52 2006-02-15 22:12:37
978 35 1 11645 4.99 2005-08-17 04:50:56 2006-02-15 22:12:37
979 35 1 12055 4.99 2005-08-17 21:02:19 2006-02-15 22:12:37
980 35 1 13735 2.99 2005-08-20 10:31:01 2006-02-15 22:12:37
981 35 1 14110 0.99 2005-08-21 00:53:09 2006-02-15 22:12:37
982 35 2 14124 2.99 2005-08-21 01:31:51 2006-02-15 22:12:37
983 35 2 14735 4.99 2005-08-21 22:25:09 2006-02-15 22:12:37
984 36 1 349 0.99 2005-05-27 04:53:11 2006-02-15 22:12:37
985 36 1 716 0.99 2005-05-29 04:35:29 2006-02-15 22:12:37
986 36 2 2741 0.99 2005-06-19 16:05:41 2006-02-15 22:12:37
987 36 2 4135 0.99 2005-07-07 08:15:03 2006-02-15 22:12:37
988 36 2 4560 4.99 2005-07-08 04:58:48 2006-02-15 22:12:37
989 36 2 4762 4.99 2005-07-08 14:54:42 2006-02-15 22:12:37
990 36 1 5403 0.99 2005-07-09 20:07:09 2006-02-15 22:12:37
991 36 2 6030 0.99 2005-07-11 02:37:51 2006-02-15 22:12:37
992 36 1 7205 6.99 2005-07-27 09:06:13 2006-02-15 22:12:37
993 36 1 7647 0.99 2005-07-28 01:35:17 2006-02-15 22:12:37
994 36 2 7919 6.99 2005-07-28 11:59:45 2006-02-15 22:12:37
995 36 2 8099 0.99 2005-07-28 18:35:12 2006-02-15 22:12:37
996 36 1 8391 2.99 2005-07-29 05:52:50 2006-02-15 22:12:37
997 36 1 8952 4.99 2005-07-30 03:20:38 2006-02-15 22:12:37
998 36 1 9369 2.99 2005-07-30 18:52:19 2006-02-15 22:12:37
999 36 2 9805 0.99 2005-07-31 11:11:10 2006-02-15 22:12:37

```

```

grelc@gandalf:~/PRODUCTION/access/bin> [fiore@glite-tutor fiore]#

```



GReC Data Access Service Query Result

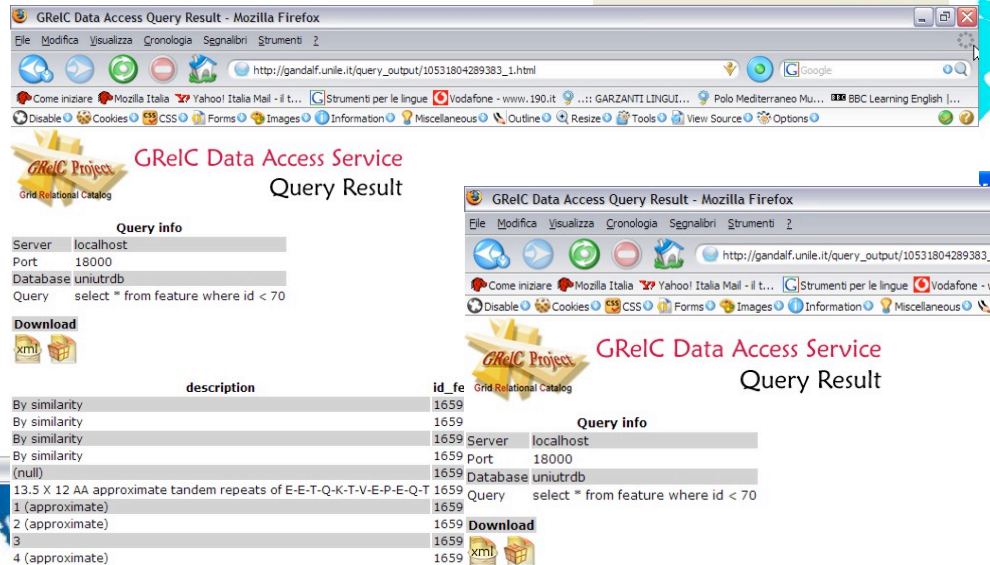
Query info

Server localhost
Port 18000
Database uniutrd
Query select * from feature where id < 70

Download

xml

description	id_feature	key_name	from_ft	to_ft	id
Hydrophobic	1659126	DOMAIN	1	19	1
Hydrophobic	1659127	DOMAIN	905	924	1
Potential	1659128	SIGNAL	1	30	2
Protein 108	1659129	CHAIN	31	102	2
By similarity	1659130	DISULFID	41	77	2
By similarity	1659131	DISULFID	51	66	2
By similarity	1659132	DISULFID	67	92	2
By similarity	1659133	DISULFID	79	99	2
Potential	1659134	SIGNAL	1	24	3
10 kDa protein	1659135	CHAIN	25	75	3



GReC Data Access Service Query Result

Query info

Server localhost
Port 18000
Database uniutrd
Query select * from feature where id < 70

Download

xml

description	id_fe
By similarity	1659
By similarity	1659
By similarity	1659
By similarity	1659
(null)	1659
13.5 X 12 AA approximate tandem repeats of E-E-T-Q-K-T-V-E-P-E-Q-T	1659
1 (approximate)	1659
2 (approximate)	1659
3	1659
4 (approximate)	1659



GReC Data Access Service Query Result

Query info

Server localhost
Port 18000
Database uniutrd
Query select * from feature where id < 70

Download

xml

description	id_feature	key_name	from_ft	to_ft	id
5	1659146	REPEAT	180	191	1
6 (approximate)	1659147	REPEAT	192	203	1
7	1659148	REPEAT	204	215	1
8	1659149	REPEAT	216	227	1
9	1659150	REPEAT	228	239	1
10	1659151	REPEAT	240	251	1
key_n 11	1659152	REPEAT	252	263	1
from 12	1659153	REPEAT	264	275	1
to_ft 13 (approximate)	1659154	REPEAT	276	287	1
id: 7 14 (incomplete)	1659155	REPEAT	288	293	1

GReC Data Access Query Result - Microsoft Internet Explorer

Indirizzo: http://gandalf.unile.it/query_output/325846930886_205.html



GReC Data Access Service
Grid Relational Catalog

Query Result

Query info

Server: gandalf.unile.it
Port: 18000
Database: uniutrdb
Query: select * from feature where id<500

Download

	description	id_feature	key_name	fr
1..13711	/source="EMBL::AY029768:4536..18246" /gene="P" /product="phosphoprotein"	3061520	3%27UTR	(ni
3160..3224	/source="EMBL::AY029768:7695..7759" /evidence="Experimental" /standard_name="Selenocysteine Insertion Sequence (SECIS)" /db_xref="UTRSITE:U0003"	3061521	SECIS	(ni
3160..3224	/source="EMBL::AY029768:7695..7759" /evidence="Pattern Similarity" /standard_name="Selenocysteine Insertion Sequence (SECIS) - type 2" /db_xref="UTRSITE:U0004"	3061522	SECIS-2	(ni
1..15318	/source="EMBL::AY029768:2929..18246" /gene="P" /product="non-structural C protein"	3061523	3%27UTR	(ni
4767..4831	/source="EMBL::AY029768:7695..7759" /evidence="Experimental" /standard_name="Selenocysteine Insertion Sequence (SECIS)" /db_xref="UTRSITE:U0003"	3061524	SECIS	(ni
4767..4831	/source="EMBL::AY029768:7695..7759" /evidence="Pattern Similarity" /standard_name="Selenocysteine Insertion Sequence (SECIS) - type 2" /db_xref="UTRSITE:U0004"	3061525	SECIS-2	(ni
1..12080	/source="EMBL::AY029768:6167..18246" /gene="M" /product="matrix protein"	3061526	3%27UTR	(ni
1529..1593	/source="EMBL::AY029768:7695..7759" /evidence="Experimental" /standard_name="Selenocysteine Insertion Sequence (SECIS)" /db_xref="UTRSITE:U0003"	3061527	SECIS	(ni
1529..1593	/source="EMBL::AY029768:7695..7759" /evidence="Pattern Similarity" /standard_name="Selenocysteine Insertion Sequence (SECIS) - type 2" /db_xref="UTRSITE:U0004"	3061528	SECIS-2	(ni
1..9952	/source="EMBL::AY029768:8295..18246" /gene="F" /product="fusion protein"	3061529	3%27UTR	(ni
1..13252	/source="EMBL::AY141760:2127..15378" /gene="U" /product="predicted protein U"	3062020	3%27UTR	(ni
1..12449	/source="EMBL::AY141760:2930..15378" /gene="V" /product="cysteine-rich protein V"	3062021	3%27UTR	(ni
1..10645	/source="EMBL::AY141760:4734..15378" /gene="M" /product="matrix protein M"	3062022	3%27UTR	(ni
1..14808	/source="EMBL::AY142960:4152..18959" /gene="VP35" /product="second nucleoprotein VP35"	3062026	3%27UTR	(ni
1..14808	/source="EMBL::AY142960:4152..18959" /gene="VP40" /product="second nucleoprotein VP40"	3062027	3%27UTR	(ni

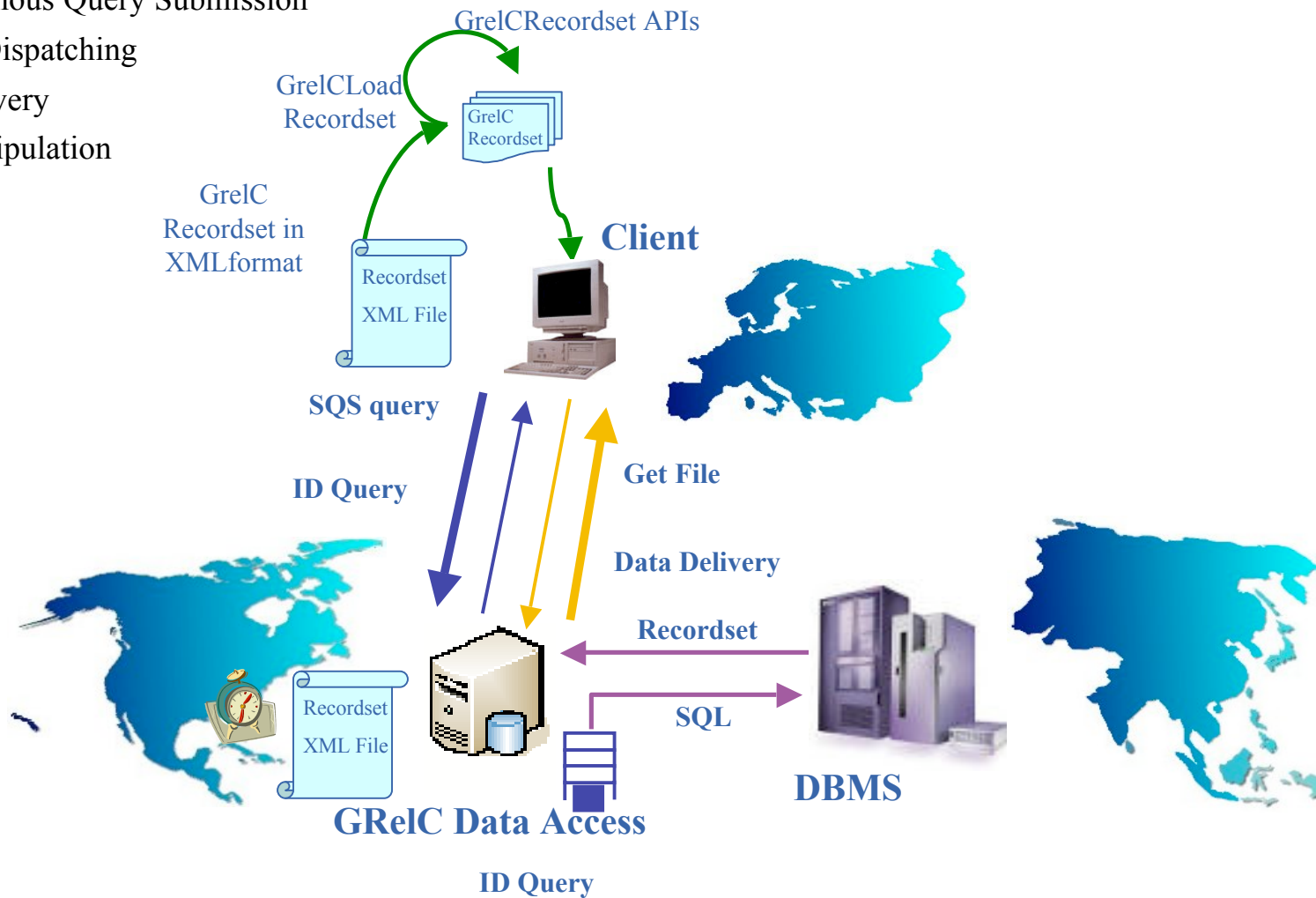
Operazione completata



Asynchronous queries

- Batch mode
- Users can define a lifetime for results availability on the GRelC DAS
- decoupling client/server (e.g. WN gLite)
- New clients (submission, status, abort...)
- Additional thread to manage requests
- Preliminary internal tests were ok
- Added within the release v2.2.0 to v2.3.0 (not included in 2.4.x)
 - Managed in a different way in the GRelC DAIS 3.0 release

- 1 – Asynchronous Query Submission
- 2 – Request Dispatching
- 3 – Data Delivery
- 4 – Data Manipulation



- **Insert Query in the Catalog**

```
grelc_service_insert_async_query -s <server_IP> -p <port> -d <db_name> -q <query>
```

- **Check Status**

```
grelc_service_check_status_async_query -s <server_IP> -p <port> -i <id_query>
```

- **Abort Query**

```
grelc_service_abort_async_query -s <server_IP> -p <port> -i <id_query>
```

- **Purge Query**

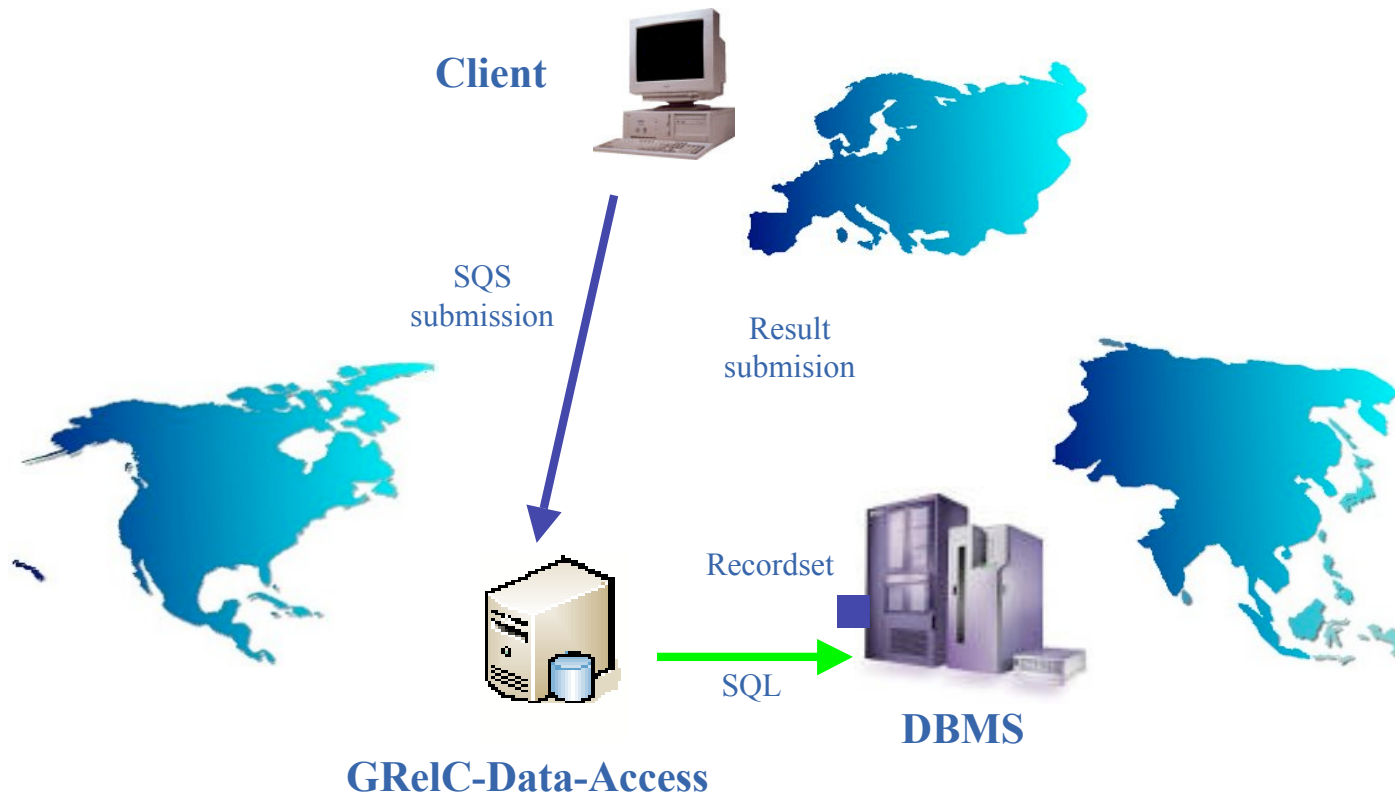
```
grelc_service_purge_async_query -s <server_IP> -p <port> -i <id_query>
```

- **Get File**

```
grelc_service_purge_async_query -s <server_IP> -p <port> -i <id_query> -f <destination_file_name>
```

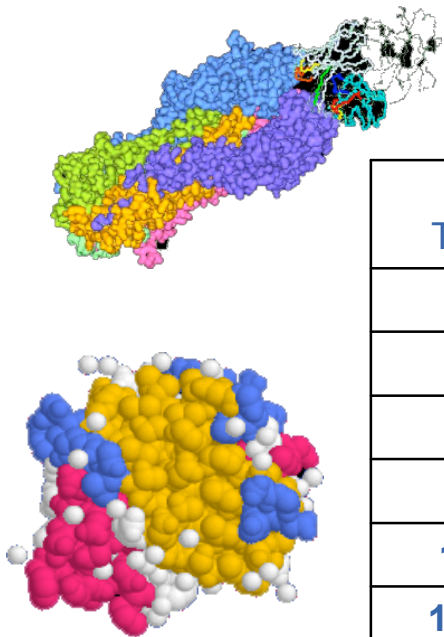
- **Get List**

```
grelc_service_purge_async_query -s <server_IP> -p <port> -d <dn> -S <status>
```



This kind of query is suitable to retrieve **VERY LARGE resultsets**

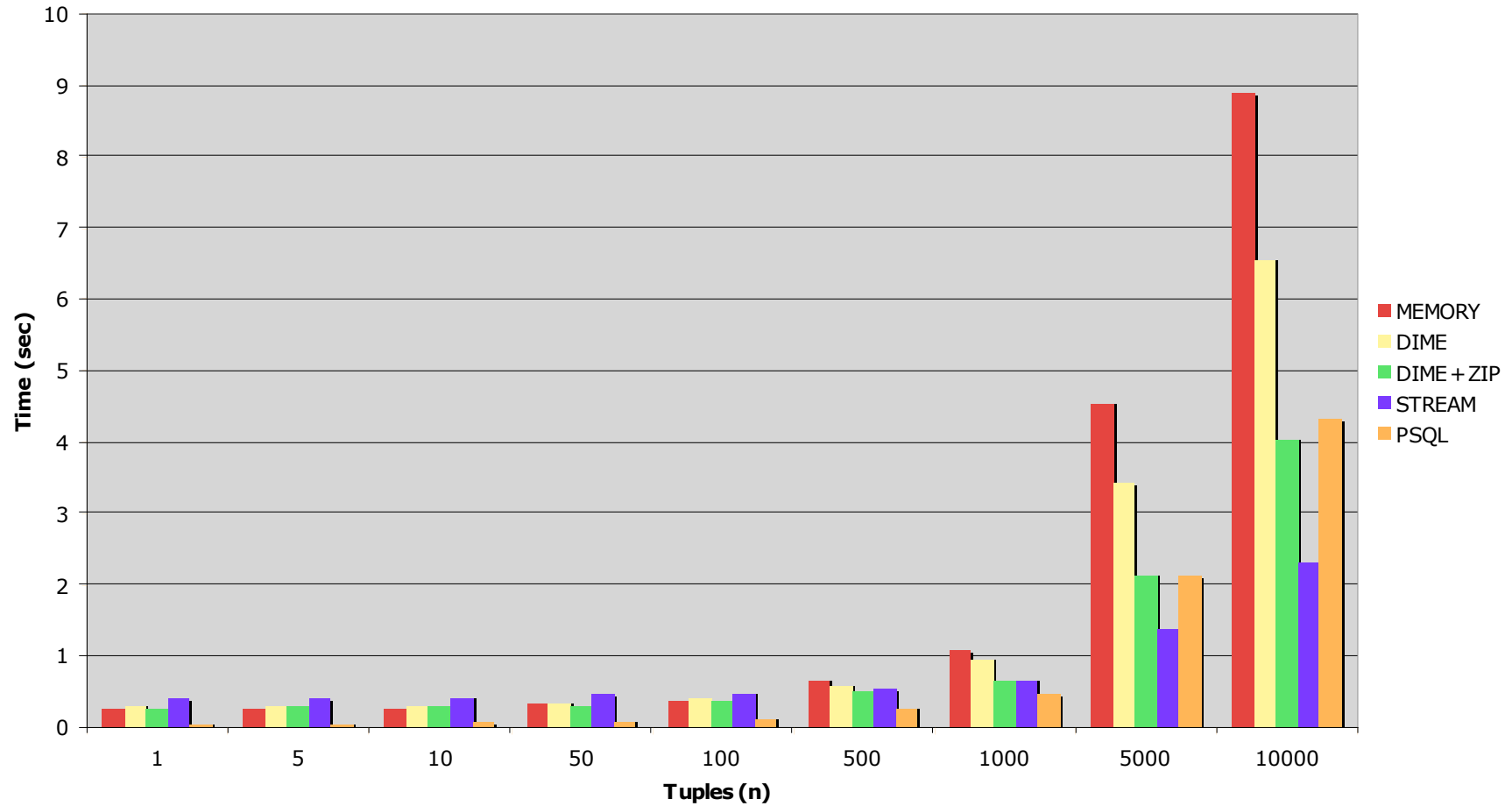
- **SQs Comparison**
- **Test DB: bioinformatics relational database**
- **Sequential tests**
- **SELECT statements**



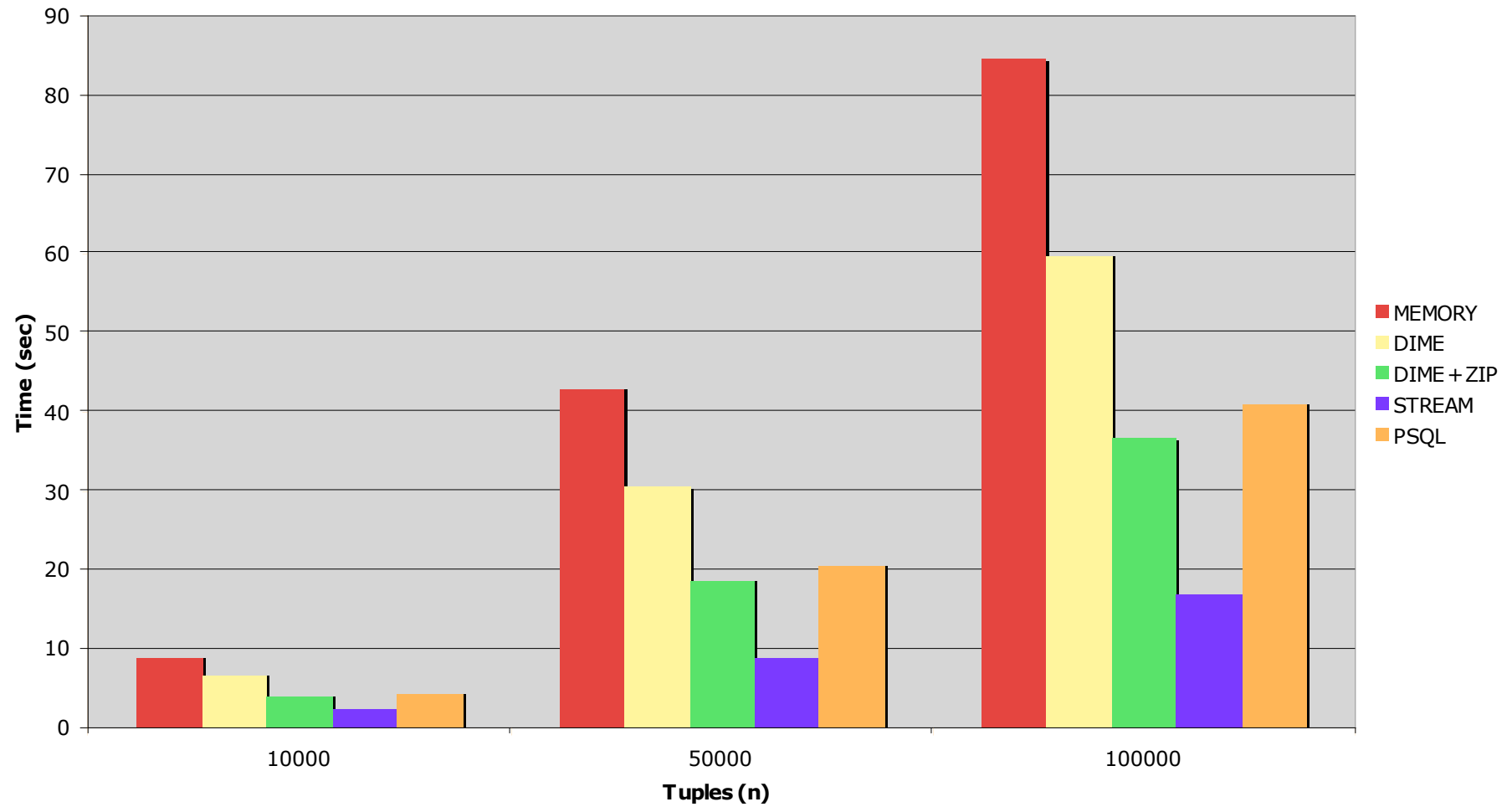
Tuples N	Resultset Size
1	2KB
10	9KB
100	81KB
1000	929KB
10000	9.9MB
100000	96.8MB



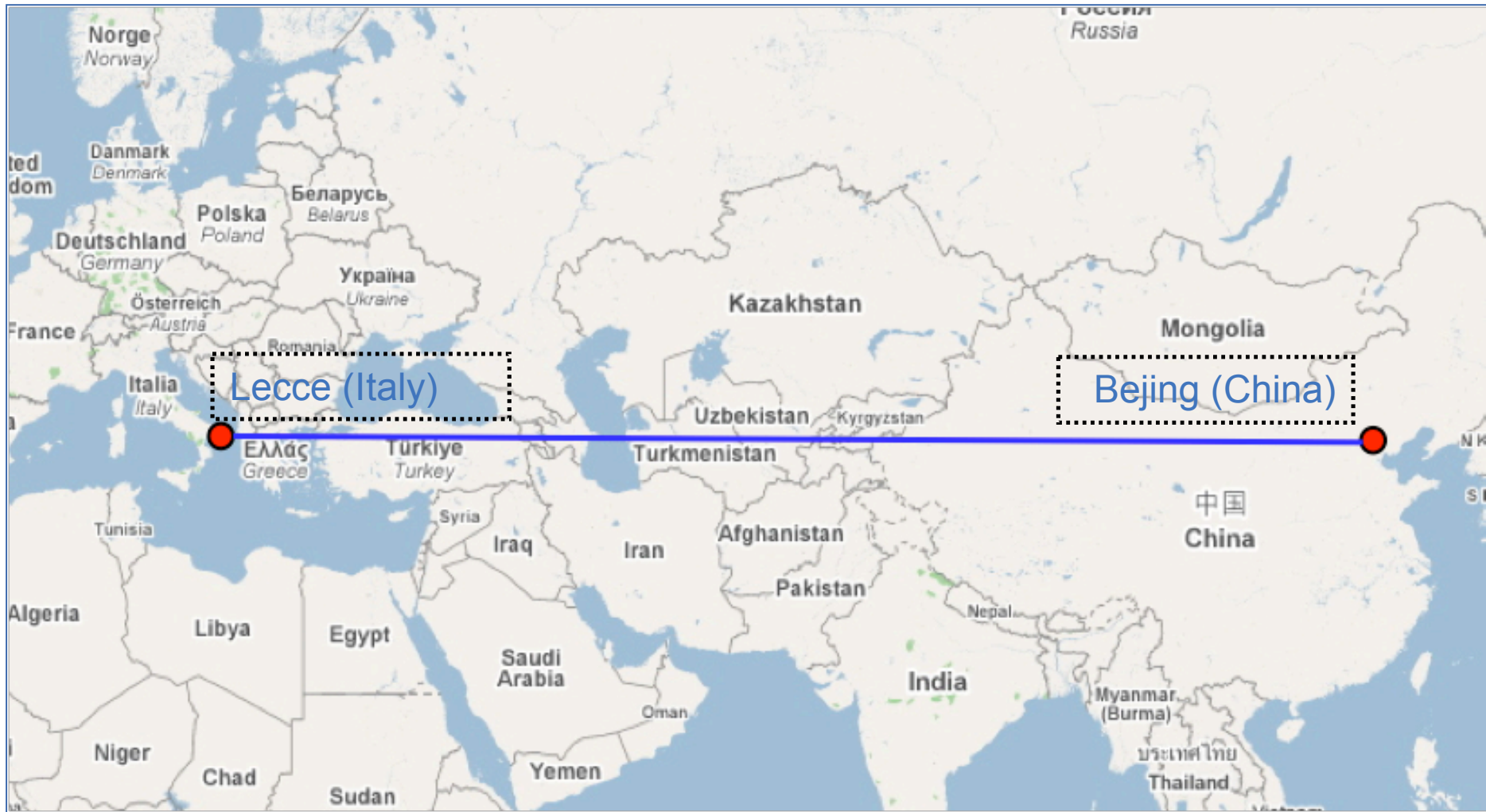
SQ Comparison - WAN (Bo-Le) - Small



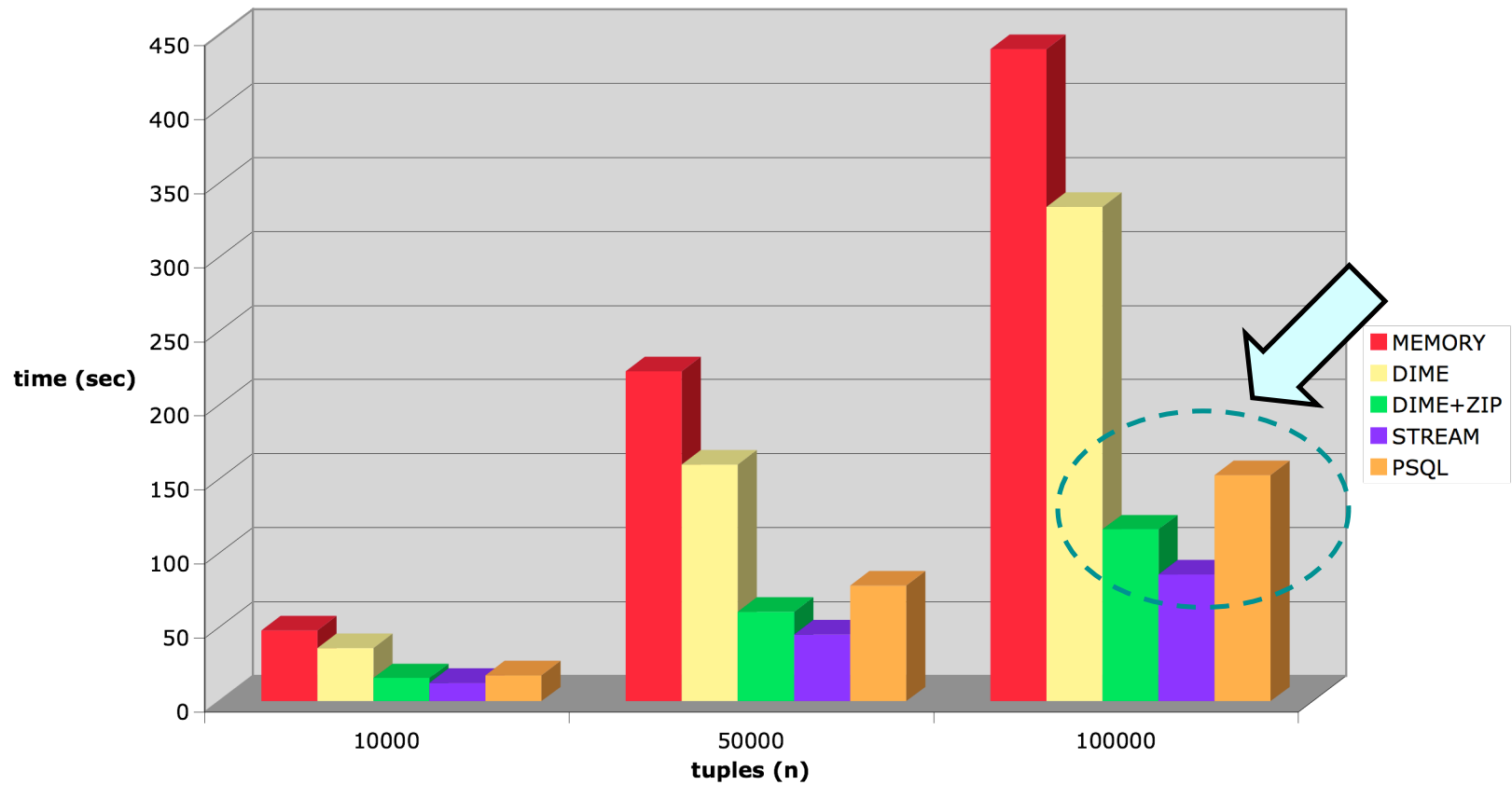
SQ Comparison - WAN (Bo-LE) - Large

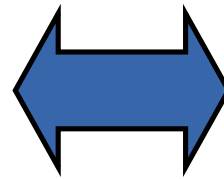


International Testbeds



SQ comparison (Beijing - Lecce) - large



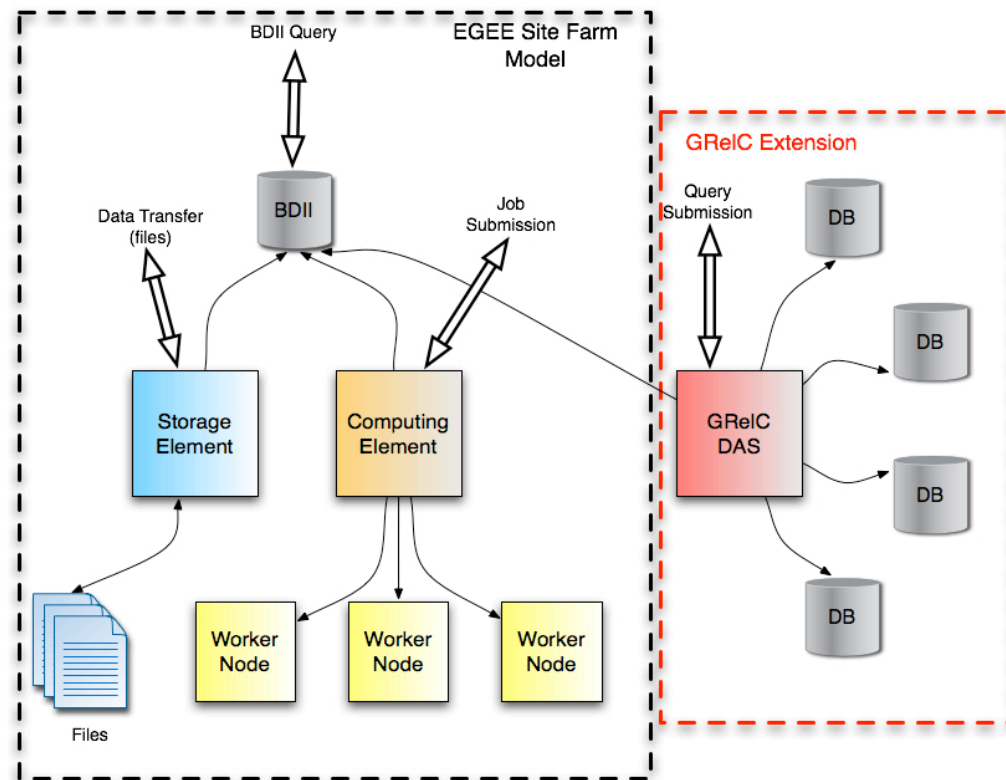
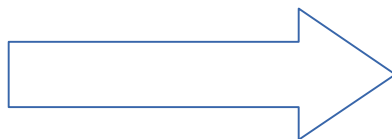


- Porting of GRelC on **gLite** was straightforward
- Porting on gLite is ok both for client and server side
- The middleware works fine both on **LCG-2-7-0** and current **gLite 3.x** middleware
- GRelC DAS runs also on several platforms:
 - **Linux**
 - **MAC OS X**
 - **FreeBSD**
- Both **IA64** and **IA32** platforms are supported (we currently installed on **SPACI-LECCE-IA64 (EGEE SA1 partner)** the GRelC DAS)

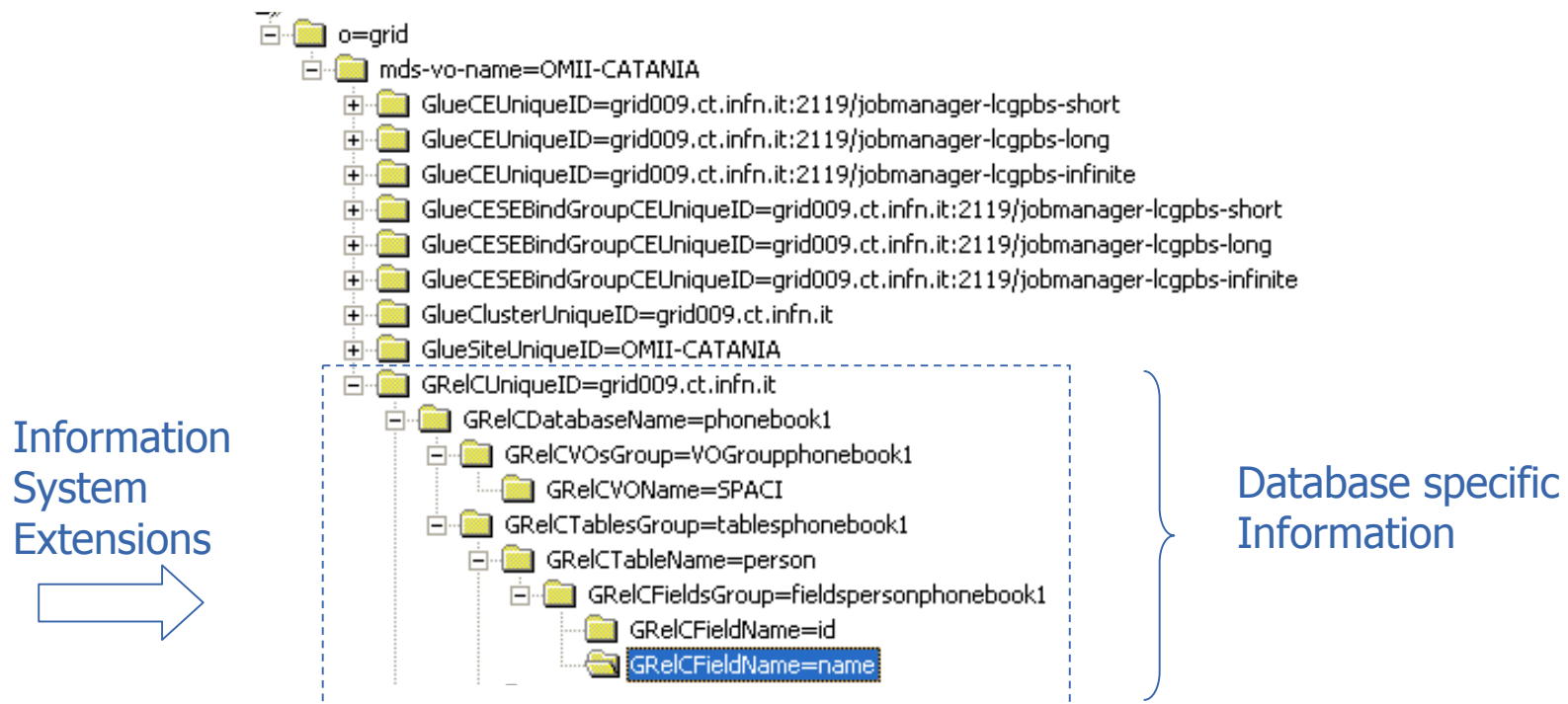
GReIC on gLite: A New Service

- Straightforward integration within the EGEE farm model
- GReIC 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

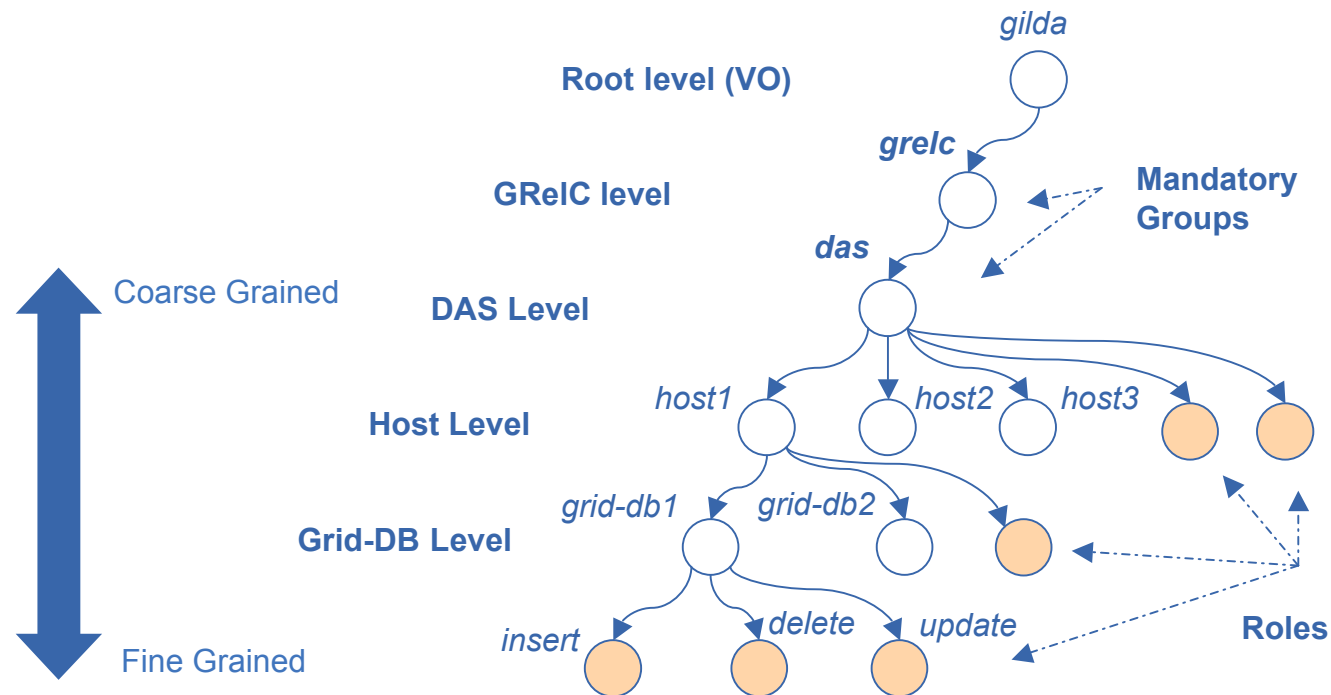


- Proprietary GLUE schema extensions providing information about VOs and Databases, etc.
- Local admin can set up the Information Provider Level parameter
Min: 0 to publish just basic info (only the contact string)
Max: 7 for all info (contact string, VOs, DBs, tables, fields, etc.)



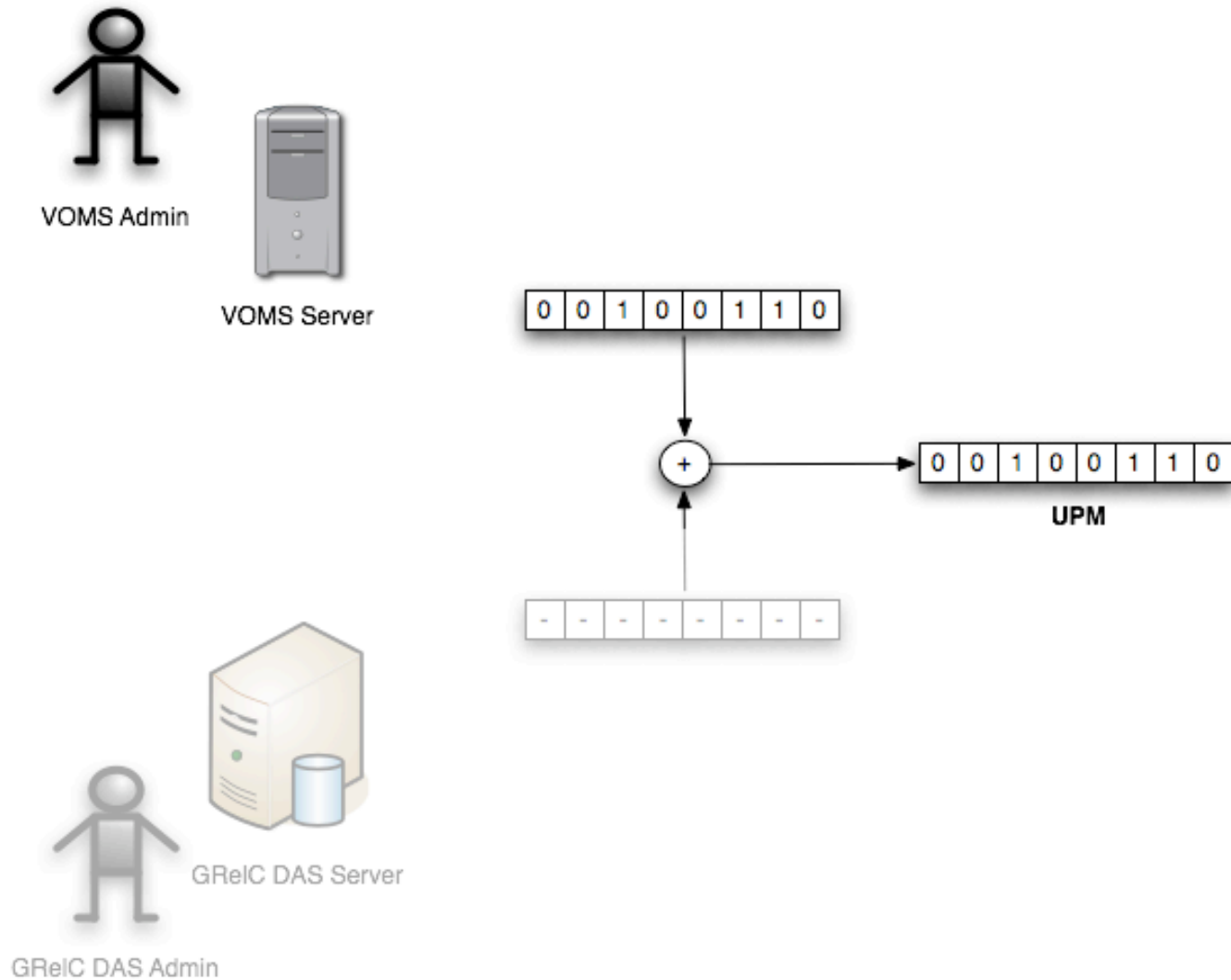
- **GRelCUniqueID = <GRelC Connection String>**
 - GRelCDatabaseName = <database name>
 - GRelCVosGroup = VoGroup<database name>
 - GRelCVoName = <vo name>
 - GRelCTablesGroup = tables<database name>
 - GRelCTableName = <table name>
 - GRelCFieldsGroup = fields<table name><database name>
 - GRelCFieldName = <field name>

- We provide global authorization by means of VOMS proxy extensions
- High level of scalability concerning DAPs related to VOs
- Double level authorization framework: both local and global policies management can be provided (**combined mode**)

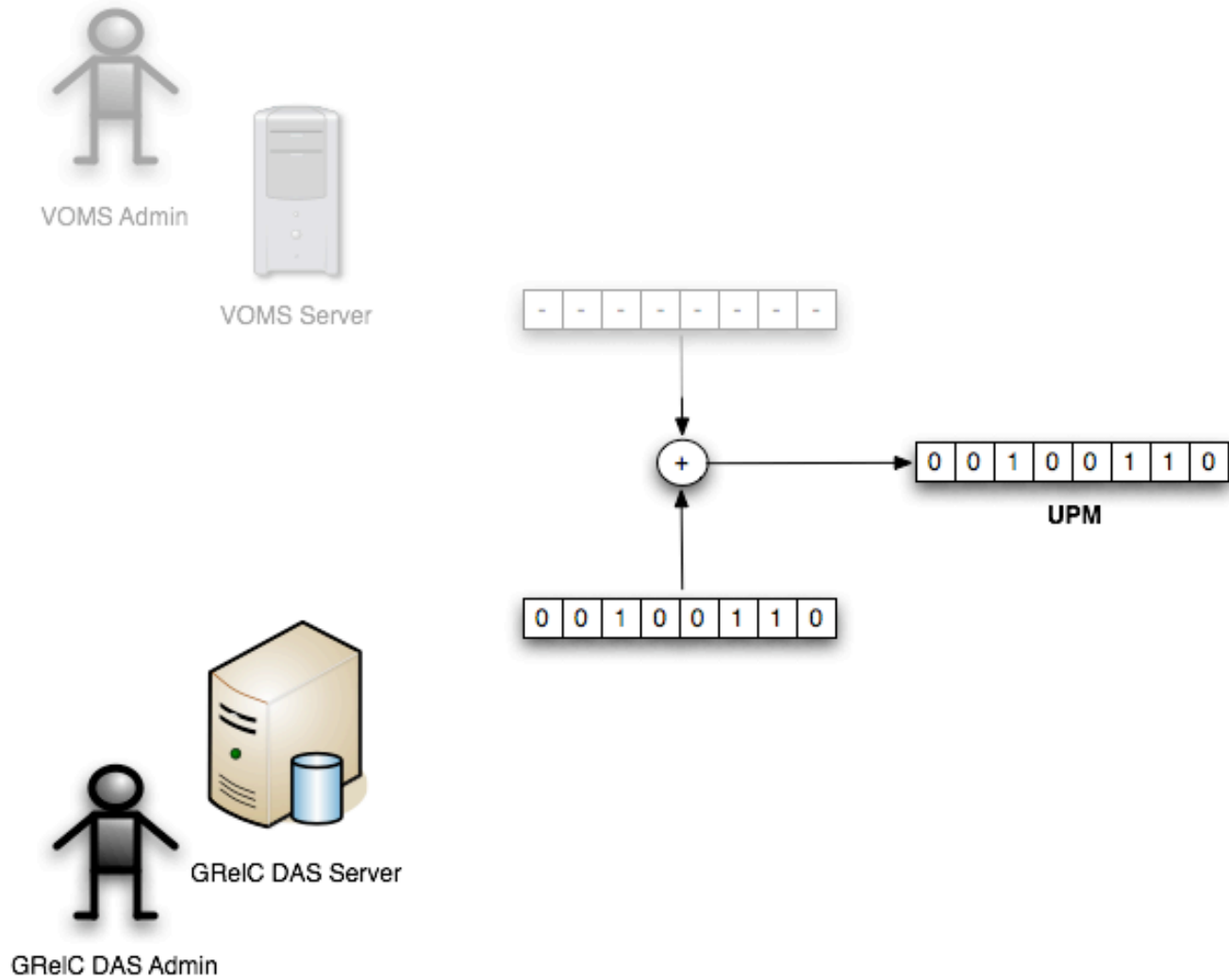


- **Global authorization** (through VOMS extensions)
- **Local authorization** (by means of the local GRelC DAS authorization framework)
- The two masks obtained from global and local authorization are combined to infer the final **User Privileges Mask (UPM)**
- **3 scenarios**
 - **global** mode, coarse grained approach
 - **local** mode, fine grained approach
 - **combined** mode

- User credentials must be obtained through **voms-proxy-init**
- The UPM is inferred from the available **VOMS extensions**
- No additional authorization setting is required on the GRelC DAS
- Easy and fast setup procedure
- **It scales well!!!!**
- Flexible mechanism
- Role setting relies on **VO admin**
- Feasible for a real production grid environments
 - EGEE
 - Currently adopted in GILDA for training activities

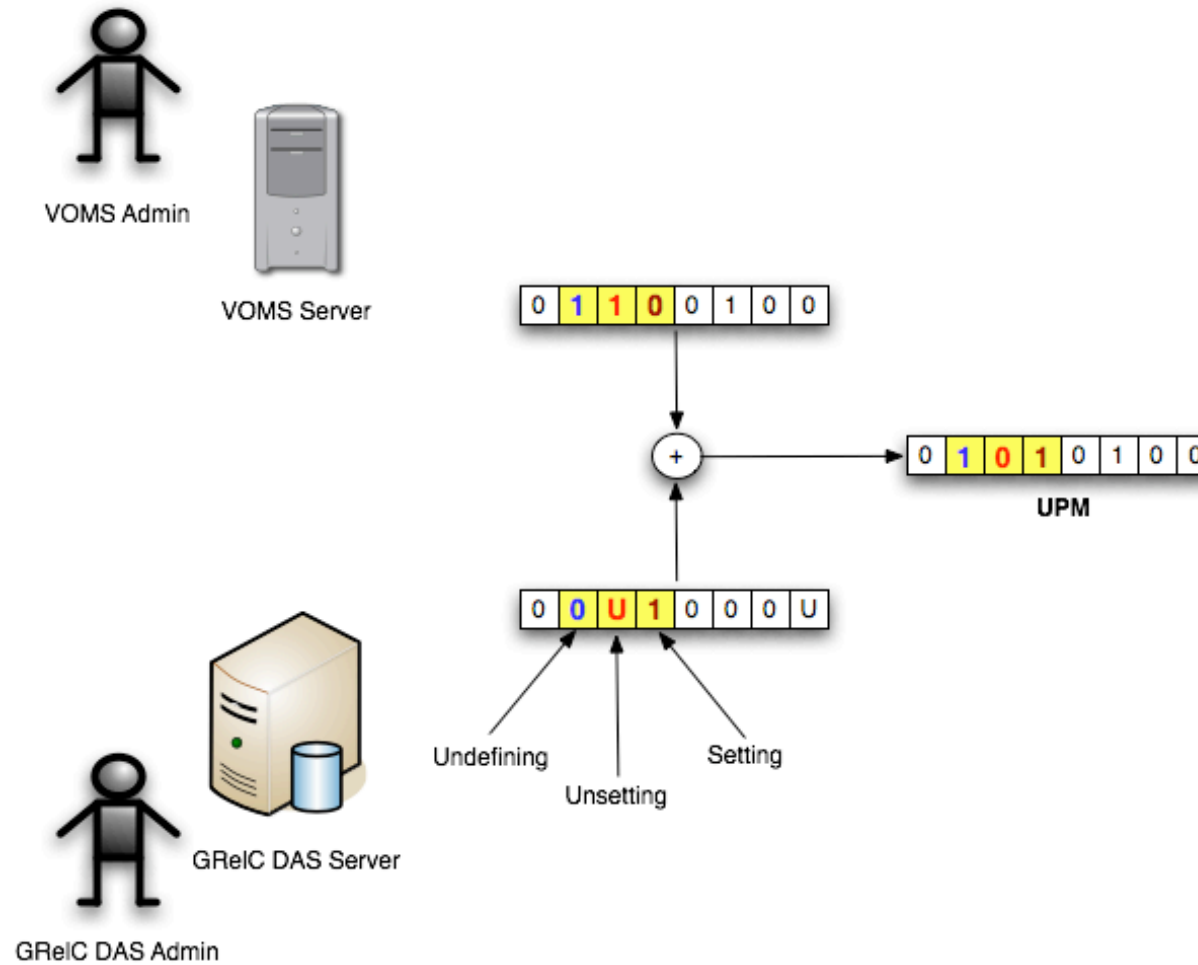


- User credentials must be obtained through *grid-proxy-init*
- The UPM is drawn out of the **GReIC DAS metadata catalogue**
- **No VOMS extensions** are added to the user proxy
- The setup procedure must be carried out on each GReIC DAS
- **Scalability is worse!**
- Good for a simple deployment scenario
- Privileges setting relies on **GReIC DAS admin**
- VO-related settings/management are not needed on the GReIC DAS side

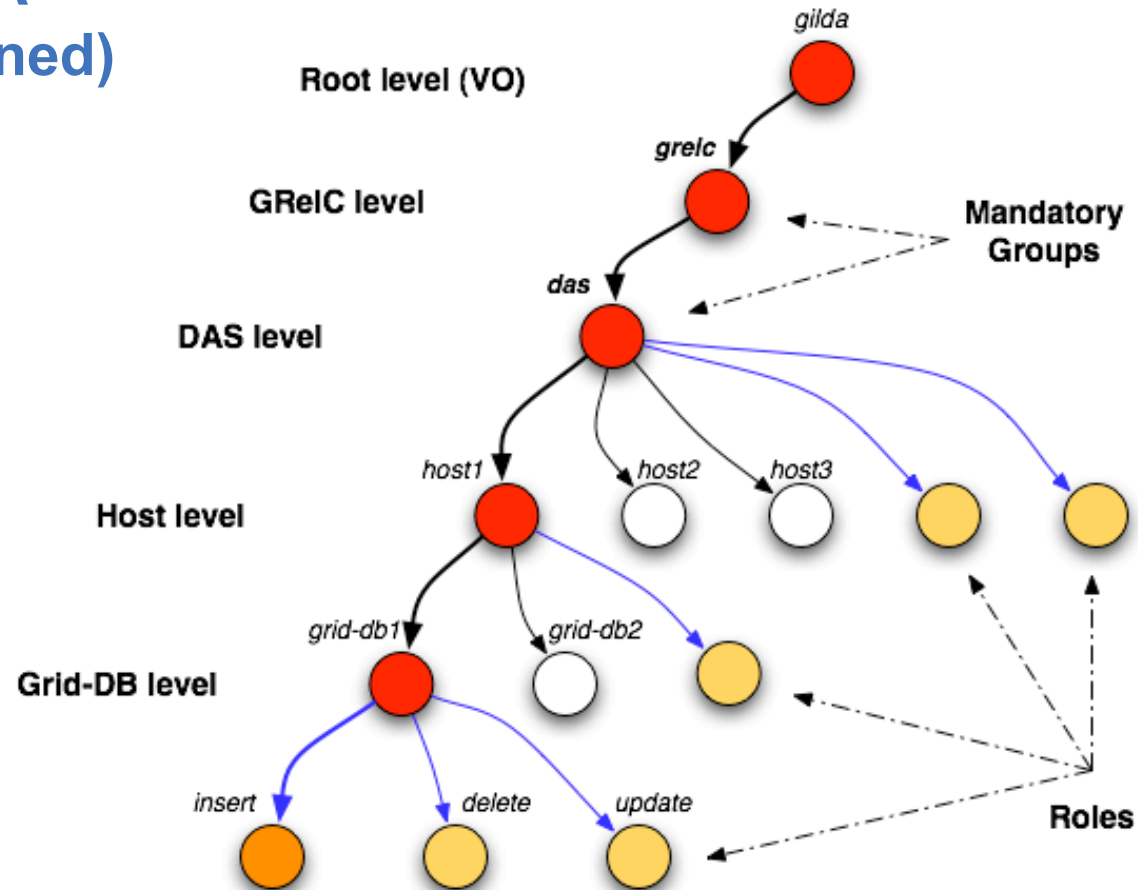


- User credentials must be obtained through ***voms-proxy-init***
- The UPM is inferred joining information on access policies coming from VOMS extensions and the GReIC DAS metadata catalogue
- Privileges/Roles setting rely on **GReIC DAS admin** and **VO-Admin**
- VOMS level (**two states**: grant or revoke)
- GReIC DAS level (**three states**: setting, undefining, unsetting)
- Local autonomy is preserved
- Represents the best way to combine local and global needs
- Currently adopted in GILDA for training activities

Combined Mode - An Example

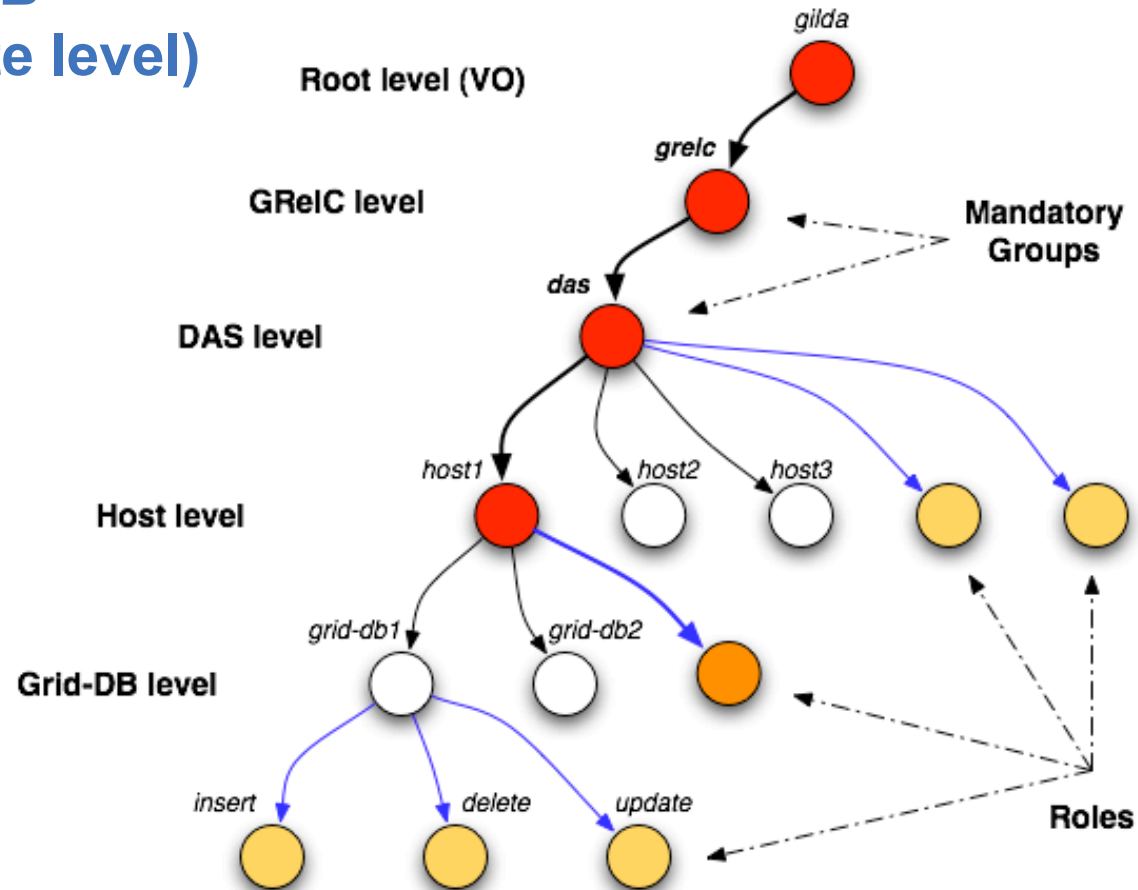


Case A (fine grained)



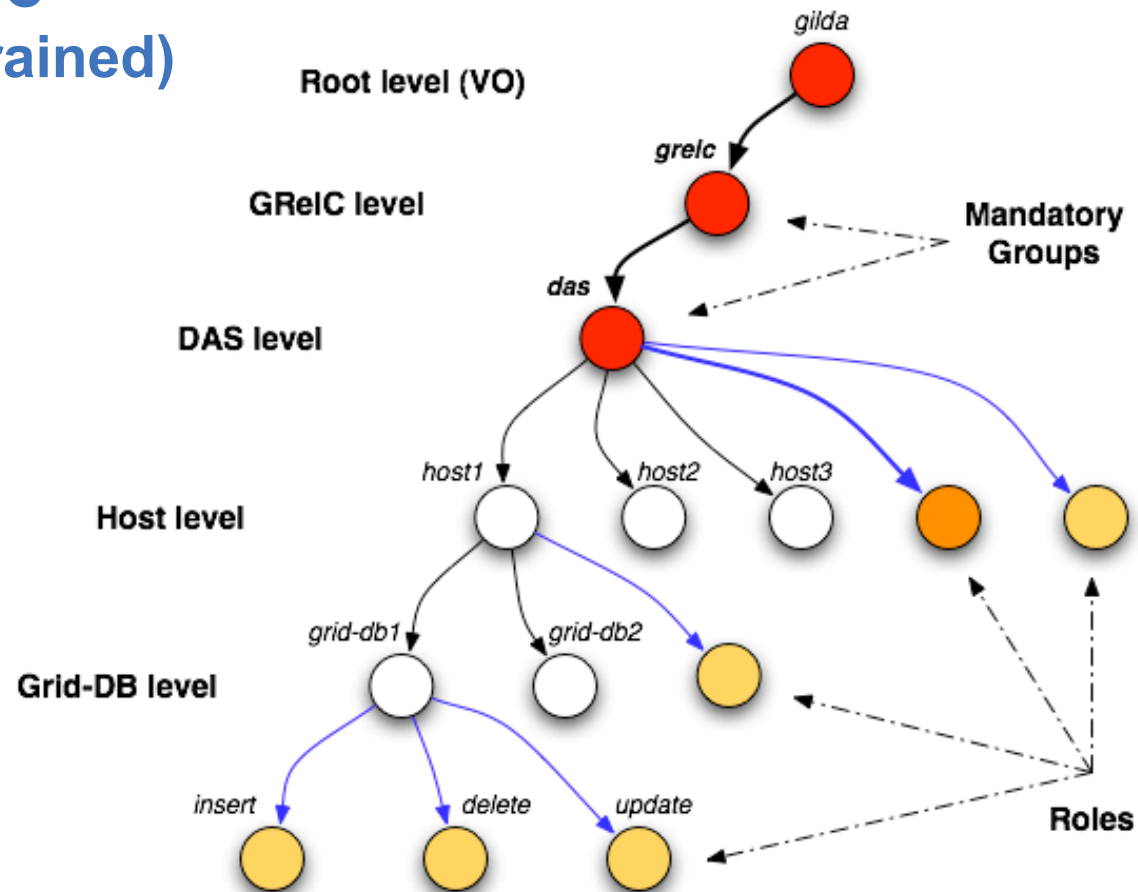
`/gilda/grelc/das/host1/grid-db1/Role=grelc-db-insert`

Case B (intermediate level)



/gilda/grelc/das/host1/Role=grelc-db-insert

Case C (coarse grained)



`/gilda/grelc/das/Role=grelic-db-insert`

- Porting on **SLC4.x** is ok
- Support for VOMS continue to be ok!
- Support for BDII ok!
- A release for SLC4.x is already available on the GRelC website
- Current tests are connected both with **IA32** and **IA64** (Itanium2 processors) platforms
- This activity is part of the **SPACI-LECCE-IA64 SA1** activity in the EGEE-III Project



GReLC DAS & INFNGRID Release

The GReLC DAS is now part of the INFNGRID release (ig_GRELC)
 The ig_GRELC profile is available for gLite (full yaim compliant)

Server

ig_GRELC (GReLC DAS)

Clients

Installed by default on
 UI and WN of the INFNGRID

Documentation

Full support in terms of
 guides, tutorials, etc.

WebSite

www.grelc.unile.it

Data Grid Portal

<https://grelc.unile.it:8443/GReLCPortal/>

If you find errors in this document please open a ticket to the "Release & Documentation" group using the INFNGRID t

Released profiles

Here below an updated lists (alphabetically sorted) of currently deployed profiles with related metapackage and nod

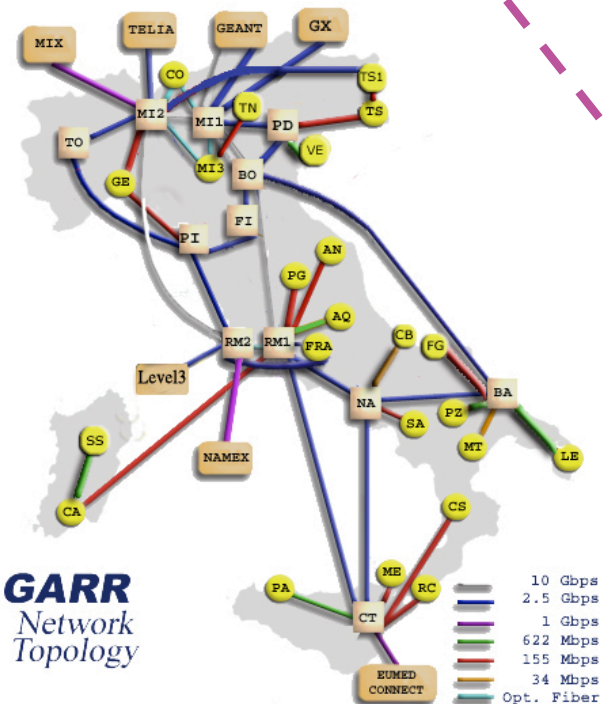
Profiles	INSTALLATION Metapackages	CONFIGURATION Nodetypes	Note
BDII site	ig_BDII	ig_BDII_site	both SL3, SL4
BDII top	ig_BDII	ig_BDII_top	both SL3, SL4
FTA	ig_FTA2	ig_FTA2	-
FTM	-	-	available only for SL4
FTS	ig_FTS2	ig_FTS2	-
gLite CE	-	-	no more supported
GReLC	ig_GRELC	ig_GRELC	
HLR	ig_HLR	ig_HLR	-
LB	ig_LB	ig_LB	-
lcg-CE	ig_CE ig_CE_LSF ig_CE_torque	ig_CE ig_CE_LSF ig_CE_torque	both SL3, SL4
LFC	ig_LFC_mysql ig_LFC_oracle	ig_LFC_mysql ig_LFC_oracle	both SL3, SL4
MON	ig_MON	ig_MON	-
PP	-	-	new installation no

The Italian e-Infrastructure

+ Several Universities
+ Compute Centers



GARR Network



The IT Grid



Slide courtesy of M. Mazzucato

EGEE User & Application Portal (NA4)

Home ► RESPECT

Main Menu

- Home
- About Grids
- Grid Applications
- Grid Infrastructure
- Grid Software
- RESPECT
- Participate
- Support
- News & Info
- Collaborative Tools
- VO Mgrs. Group
- Problems
- Contacts
- Links
- Search

RESPECT

Thursday, 29 March 2007

The **RESPECT program** (Recommended External Software for EGEE CommunitiEs) aims to publicize grid software and services that work well in concert with the EGEE gLite software and thereby:

- Expand the functionality of the grid infrastructure for users,
- Reduce duplicated development when porting applications, and
- Speed the porting of new application to the grid.

The **procedure to propose new packages** for this program is available from EDMS.

GridWay Metascheduler ([Website](#), [Documentation](#), [Download](#), [Support](#))
The GridWay Metascheduler performs job execution management and resource brokering, allowing unattended, reliable, and efficient execution of jobs, job arrays, and workflows on heterogeneous and dynamic Grids.

P-GRADE Portal ([Website](#), [Documentation](#), [Download](#), [Support](#) [pgportal at lpd.sztaki.hu])
The Parallel Grid Run-time and Application Development Environment Portal (P-GRADE Portal) is a workflow-oriented graphical environment that covers every stage of Grid application lifecycles. It supports the development, execution and monitoring of workflows and workflow based parametric studies built up from various types of executable components.

Ganga ([Website](#), [Documentation](#), [Download](#), [Support](#))
Ganga is an easy-to-use frontend for job definition and management, implemented in Python. Ganga allows trivial switching between testing on a local batch system and large-scale processing on Grid resources. Ganga is readily extended and customised to meet the needs of different user communities.

DIANE ([Website](#), [Documentation](#), [Download](#), [Support](#))
DIANE is a lightweight distributed framework for parallel scientific applications in master-worker model. It assumes that a job may be split into a number of independent tasks which is a typical case in many scientific applications. The DIANE framework takes care of all synchronization, communication and workflow management details on behalf of the application.

i2glogin ([Website](#), [Documentation](#), [Download](#), [Support](#) [i2glogin at gup.jku.at])
i2glogin is a tool that enables interactive communication between a grid job and the user. The tool opens a fast, secure channel between the job and the user interface allowing, for example, debugging of the remote environment, application steering, or remote visualization.

GReIC ([Website](#), [Documentation](#), [Download](#), [Support](#) [grelc-user at sara.unile.it])
The Grid Relational Catalog Project (GReIC) aims at providing a set of advanced data grid services to transparently, efficiently and securely manage grid-databases. At the moment the GReIC Data Access Service (GReIC DAS) allows users accessing and interacting both with heterogeneous RDBMS (PostgreSQL, MySQL, SQLite, etc), as well as XML DB engines (XIndex, eXist, etc.) providing an uniform and grid enabled access interface to wide spread data sources.





User tutorials: GILDA t-Infrastructure

GReC DAS User Tutorial
on GILDA Grid CT Wiki Website

Info about:

- Log in to the grid
- Query Submission



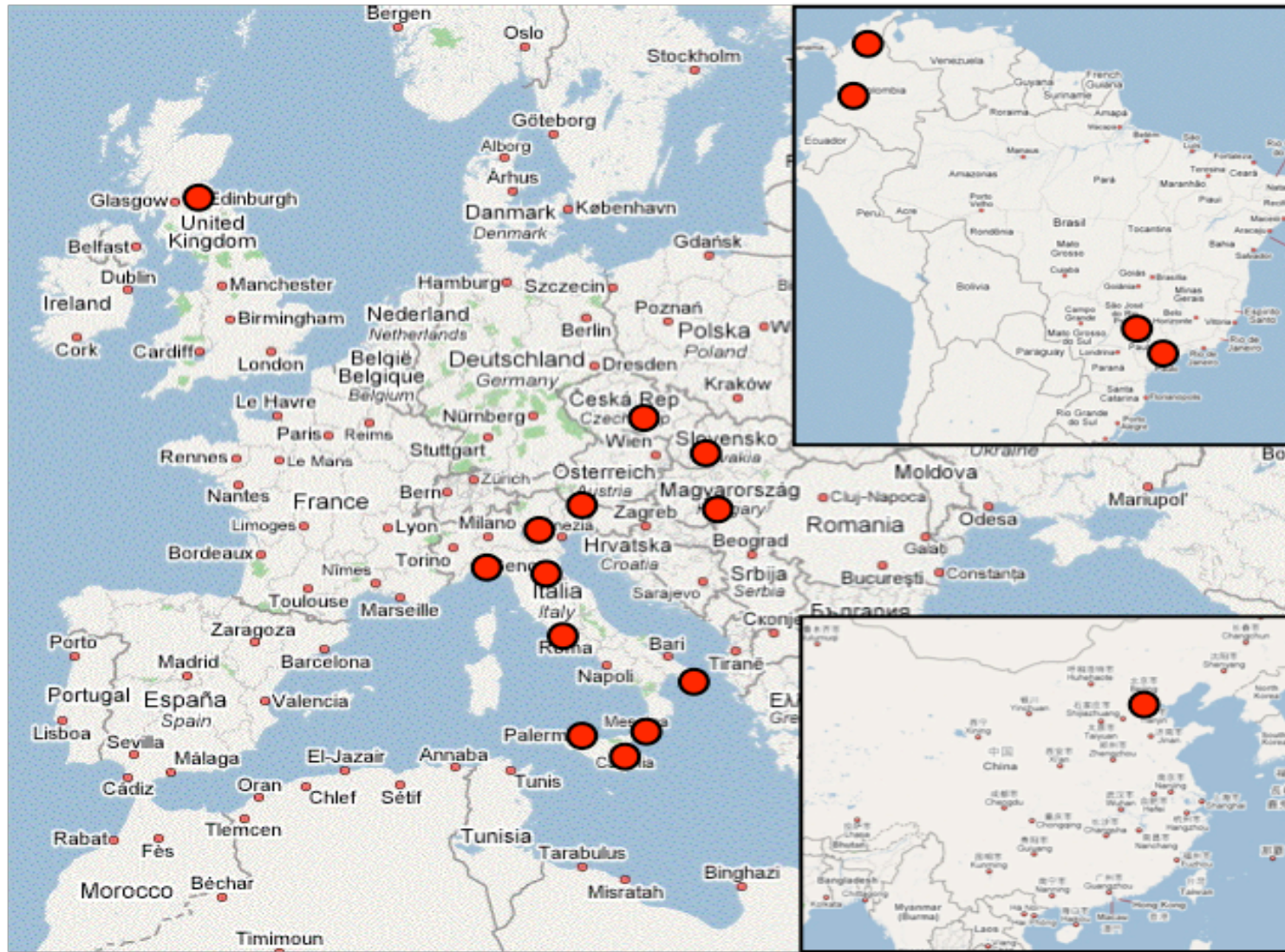
The screenshot shows a Mozilla Firefox browser window displaying the GReC Data Access Service 2.2.0 page. The address bar shows the URL: <https://grid.ct.infn.it/twiki/bin/view/GILDA/GReCDataAccessService>. The page content includes the GReC Project logo, a navigation menu with options like 'Log In or Register', 'GILDA Web', 'Create New Topic', 'Index', 'Search', 'Changes', 'Notifications', 'Statistics', and 'Preferences'. Below the menu is a 'Webs' section with links to EUMEDGRID, GILDA, ICEAGE, Main, PI2S2, Sandbox, TRIGRID, and TWiki. The main content area features the title 'GReC Data Access Service 2.2.0' and lists the supervisor (Prof. Giovanni Aloisio), project P.I. (Sandro Fiore, Ph.D.), team members (Massimo Cafaro, Alessandro Negro, Salvatore Vadacca), and website (<http://grelc.unile.it>). An 'Introduction' section explains that many data grid applications manage and process huge datasets distributed across multiple grid nodes. A 'GReC Project' section describes the project's goal of providing efficient and transparent data grid services.

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

Special thanks to the GILDA
Staff for their support

GReC DAS Tutorial link:

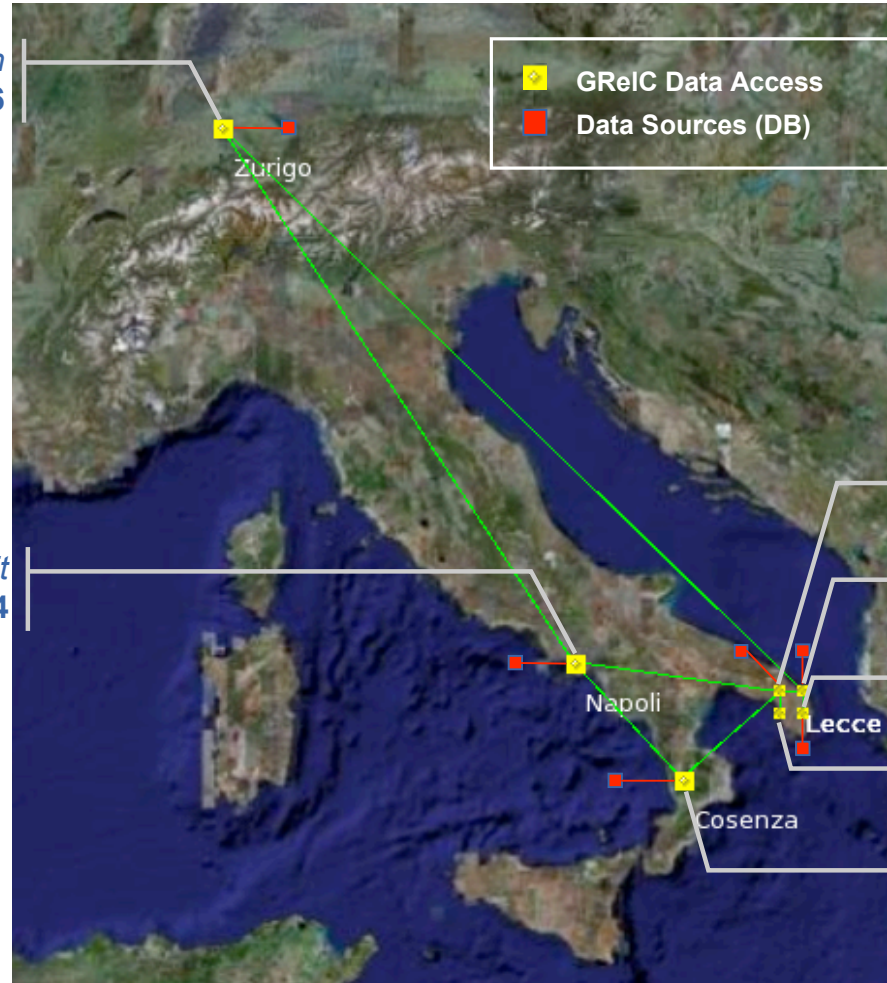
<https://grid.ct.infn.it/twiki/bin/view/GILDA/GReCProject>



SEPAC Grid Deployment

sepac00.projects.cscs.ch
Linux x86

spacina.na.infn.it
Linux IA64





GReIC WebSite

Tuesday, 3rd April 2007 10:54 AM +0200

GReIC
AN EASY WAY TO MANAGE GRID DATABASES

Home Components Downloads Documentation News Events Publications Staff Deployment

Headlines

- 29-03-2007 GReIC DAS clients version 2.1.2 are now available
- 29-03-2007 GReIC DAS Clients belong to the GILDA Virtual Services
- 27-03-2007 GReIC DAS version 2.1.1 is now available for gLite 3.0
- 21-03-2007 GReIC DAS version 2.1.0 is now available for gLite 3.0
- 26-02-2007 GILDA hosts the training version of GReIC DAS 2.1.0 (server and clients)
- 05-02-2007 GILDA hosts the training version of GReIC DAS clients
- 03-02-2007 GReIC DAS clients version 2.0.1 are now available for gLite 3.0 User Interfaces
- 01-02-2007 GReIC website is now online!

Events

- 11-05-2007 2nd EGEE User Forum - Manchester, UK
- 12-04-2007 CYCLOPS First Training Workshop - Bologna, Italy

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. It is developed by the GReIC Team (University of Salento, Lecce & SPACI Consortium).

The GReIC middleware is currently used within the SPACI and SEPAC Production Grids 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 two services are already available:

1. GReIC Data Access Service (GReIC DAS) which allows users accessing and interacting with heterogeneous DBMS (PostgreSQL, MySQL, Oracle, DB2, SQLite, etc), providing an uniform access interface to data sources
2. GReIC Data Gather which aims at integrating transparently and securely distributed and geographically spread heterogeneous grid data sources through Data Gather Service nodes connected in a P2P fashion.

[Top]

Components

GReIC Standard Database Access Interface (GReIC SDAI)

To address the transparency and uniformity requirements, we conceived the **Standard Database Access Interface**, a C library leveraging a plug-in based approach. The library virtualizes the data access operations.

Main sections:

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

GReIC Website

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

Mailing List

mail: greic-user@sara.unile.it

Functionalities:

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



Wed Sep 26 13:53:24 CEST 2007

GReIC
AN EASY WAY TO MANAGE GRID DATABASES

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

Query

Synchronous Query

GReIC DASs

Servers List

gandalf
grelc02
grelc01

Operations - gandalf

System Users

Hosts

Instances

Grid databases

Grid Database Users

Virtual Organizations

Association

Session

Logout

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

Features:

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



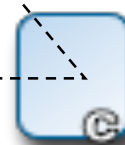
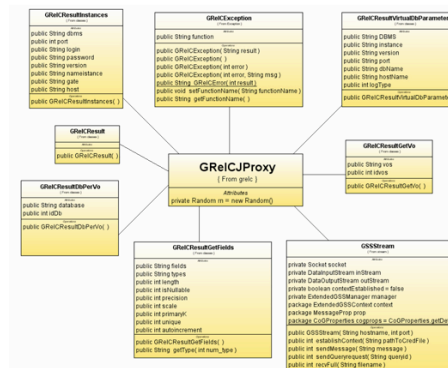
GRELC Portal: Some Snapshots



GReIC
AN EASY WAY TO MANAGE GRID DATABASES

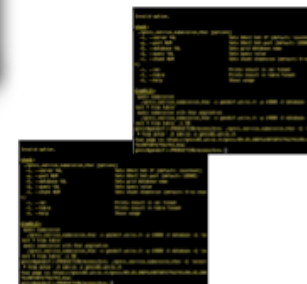
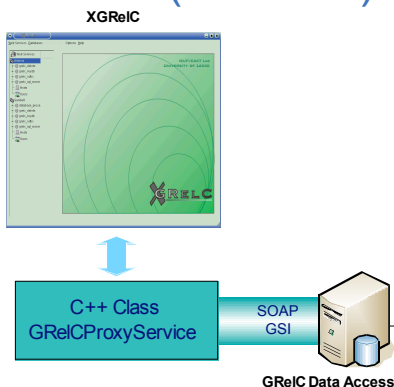
NAME	COMPONENTS	DOWNLOADS	DOCUMENTATION	NEWS	EVENTS	PUBLICATIONS	HELP	DEPLOYMENT	FEEDBACK
Query									
Synchronous Query Result									
Server: gprof									
Database: hella									
Query: select * from actor									
SELECT ID	ACTOR_NAME	ACTOR_LASTNAME	ACTOR_UPDATE						
1	PULLMAN	DAVID	2006-02-13 04:34:33						
2	NOX	DAVID	2006-02-13 04:34:33						
3	ED	CHRIS	2006-02-13 04:34:33						
4	JANSEN	DAVID	2006-02-13 04:34:33						
5	JOHNSON	LOU	2006-02-13 04:34:33						
6	BETTE	NICHOLSON	2006-02-13 04:34:33						
7	GRACE	MARY	2006-02-13 04:34:33						
8	BEYONCE	JAY	2006-02-13 04:34:33						
9	JOE	BRANK	2006-02-13 04:34:33						
10	CHRISTINA	GARY	2006-02-13 04:34:33						
11	ZERO	CASEY	2006-02-13 04:34:33						
12	KARL	BERRY	2006-02-13 04:34:33						
13	UMA	WOOD	2006-02-13 04:34:33						
14	VIGAN	BRIANNA	2006-02-13 04:34:33						
15	CLAIR	ALVIN	2006-02-13 04:34:33						
16	FRID	GOSTNER	2006-02-13 04:34:33						
17	HELEN	OSGOTT	2006-02-13 04:34:33						
18	DAI	TONK	2006-02-13 04:34:33						
19	BON	FARRETT	2006-02-13 04:34:33						
20	LUCILLE	TRACY	2006-02-13 04:34:33						

Grid Enabled Web Apps (GReIC Portal)

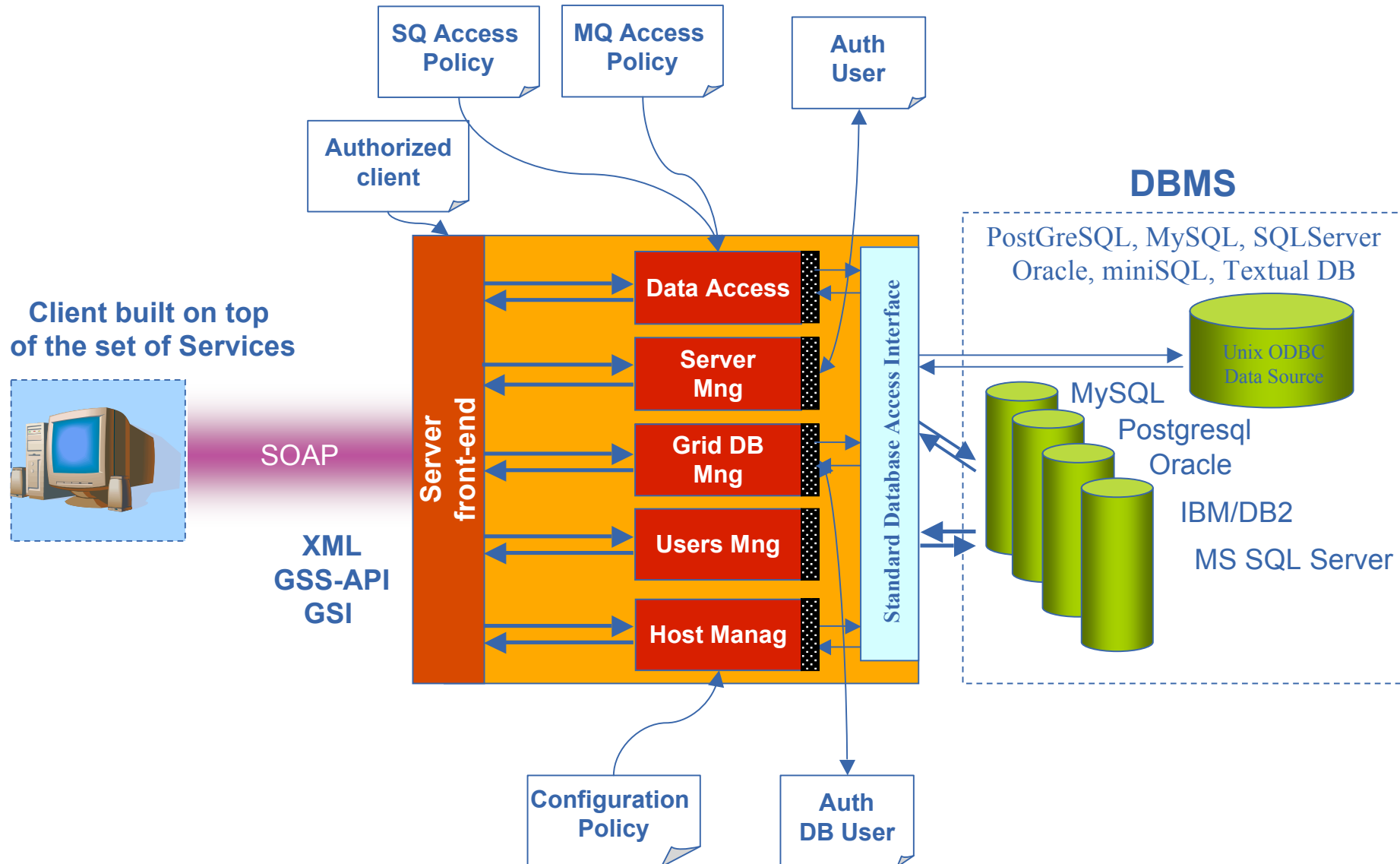


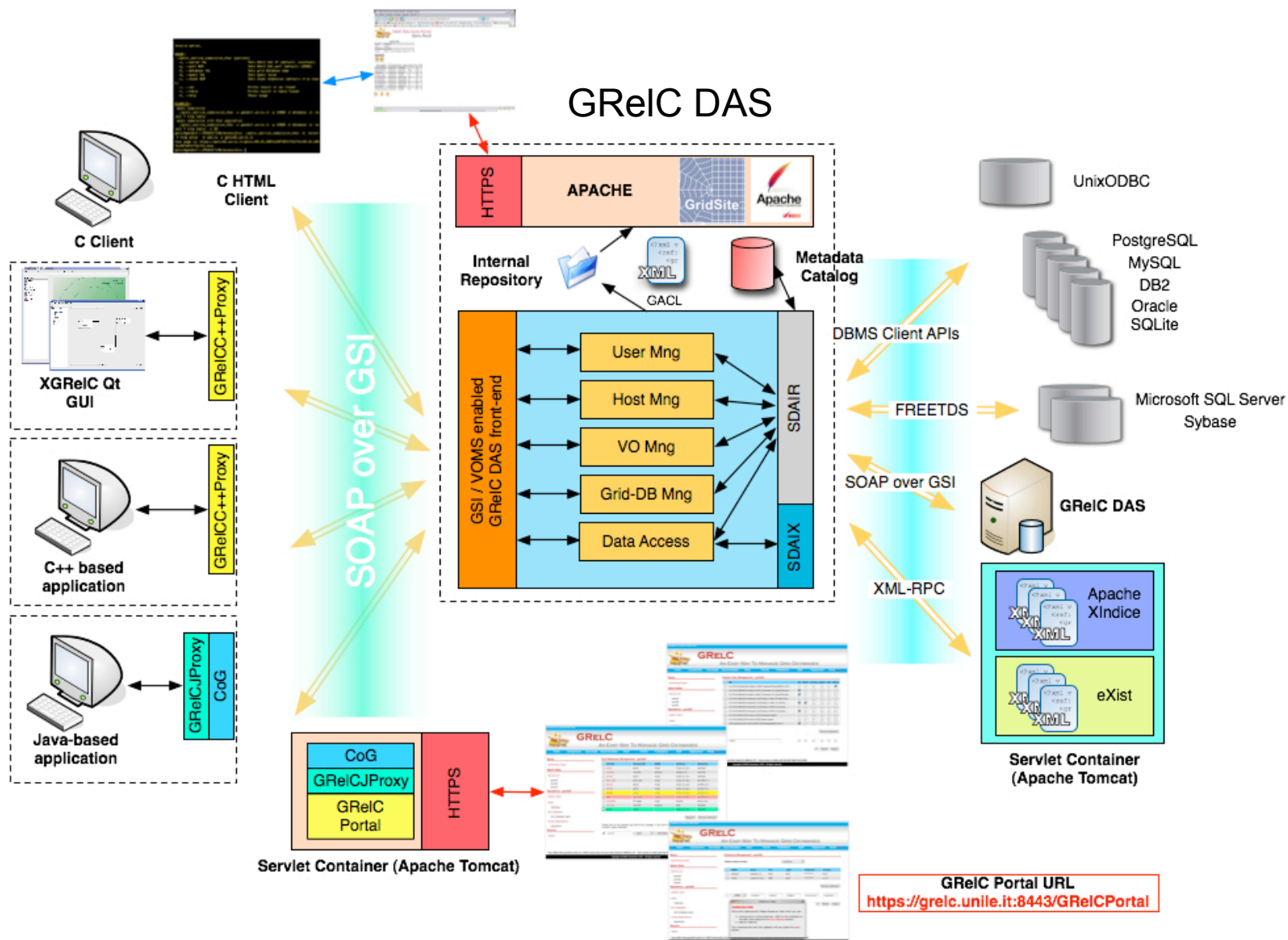
Grid Enabled Stand alone GUI (XGReIC)

High Performance Grid Services & Clients (CLI)



From the first architectural design in 2001...



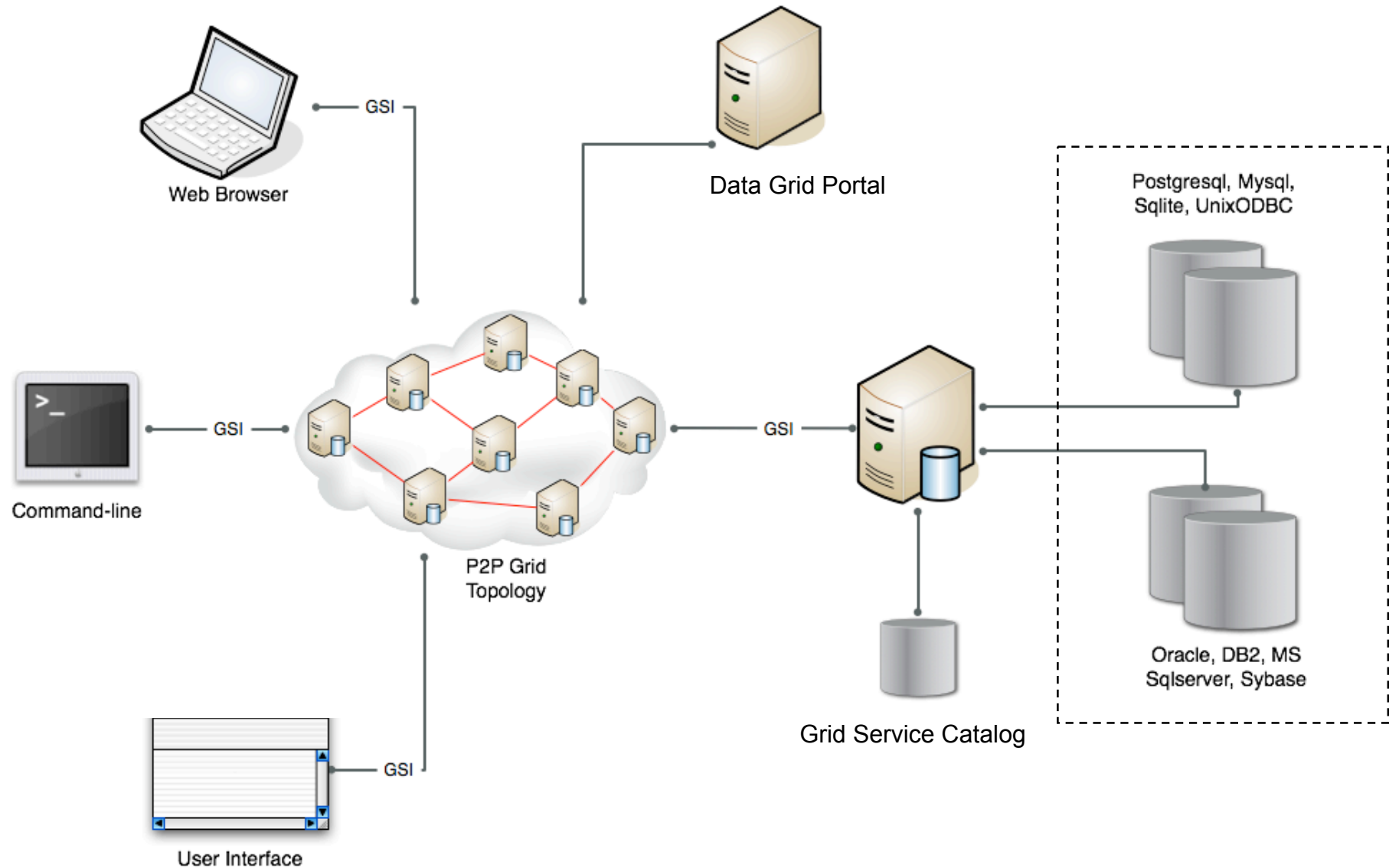




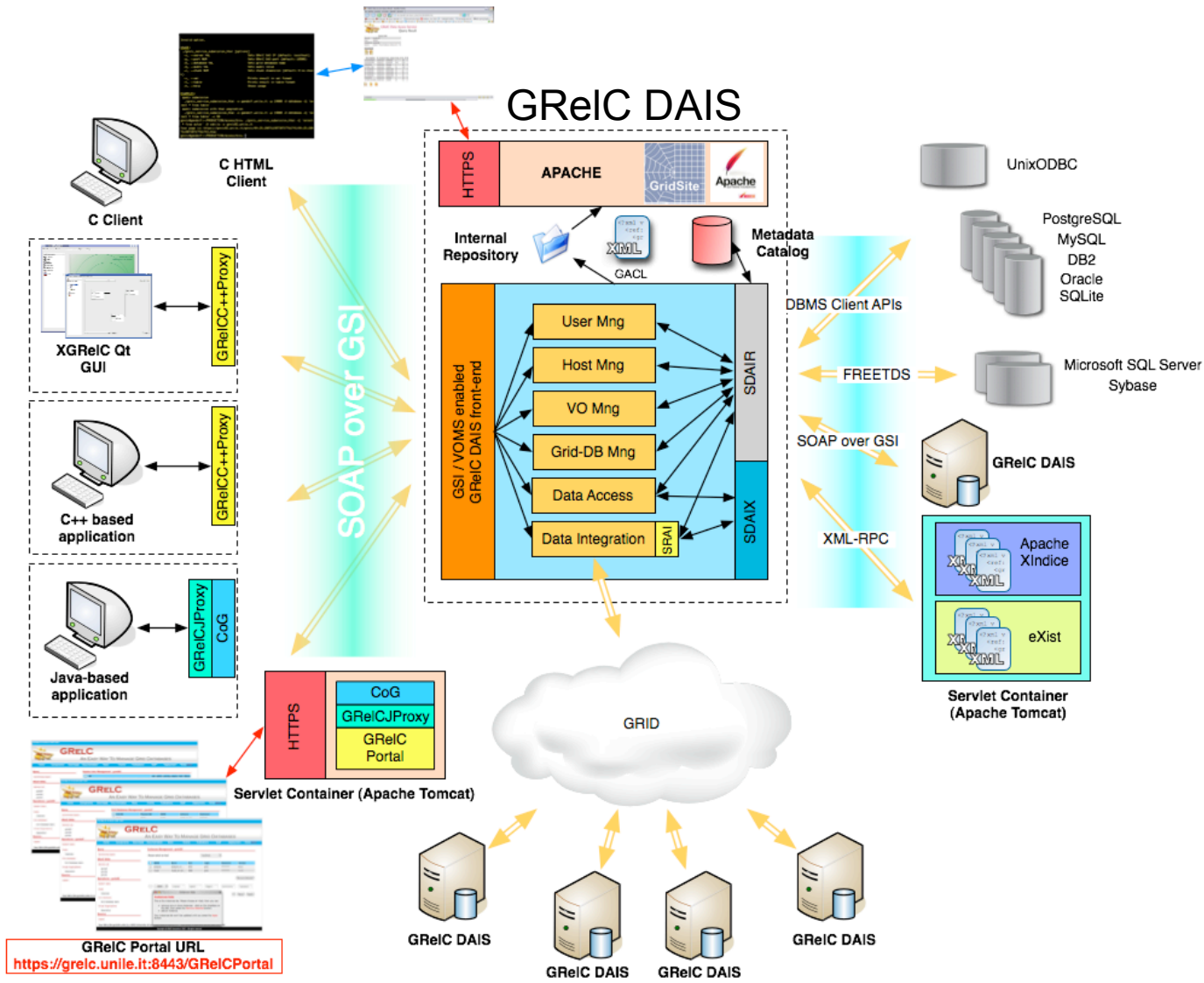
Last release: GReIC DAIS 3.0

- **GReIC DAIS: Data Access and Integration** service
 - Based on **GReIC DAS experience**
 - GReIC DAS extension
- New modules for **data integration**
- Data Grid Middleware to access to **distributed databases**
- New architectural (design) and infrastructural (implementation) issues
- No centralization: P2P approach
- Scalable approach
- Support for different data models (relational, hierarchical, etc.)
- Compatibility with different production grid middleware (gLite, Globus, Unicore, etc.)
- Fixed schema constraint

Architecture in the large



Architecture in the small



- **Transparency:**
 - Access: transparent access to heterogeneous DBMS
 - RDBMS (Oracle, MySQL, Postgresql, IBM/DB2, etc.)
 - XML (Xindice, eXist, libxml2 based documents, etc.)
 - Location: physical database location completely hidden
 - Naming: mapping between physical parameters and virtual/grid DB name
- **Efficiency:**
 - Multithreaded grid service
 - Prethreaded support to enhance performance
 - Advanced delivery mechanisms supporting:
 - Compression
 - GridFTP
 - Streaming
 - etc.

Requirements (II)

- **Scalability:**

- Management of virtual organizations
- Role-based management based on existing services:
 - VOMS (gLite - EGEE)
 - etc.
- Decentralized solutions for data-service/data-service interaction in grid
 - P2P data grid services
 - ...

- **Interoperability:**

- Standardized data request interface
 - Web Services Technologies
- Wide set of standards for underlying technologies
 - OASIS
 - W3C
 - OGF
- Working Groups
 - OGF, EGEE, etc.

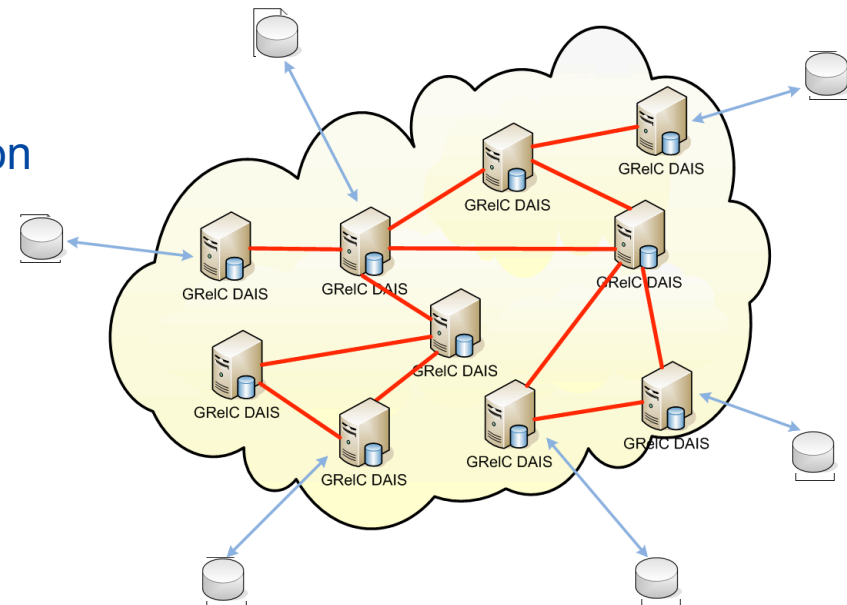


- **Security:**

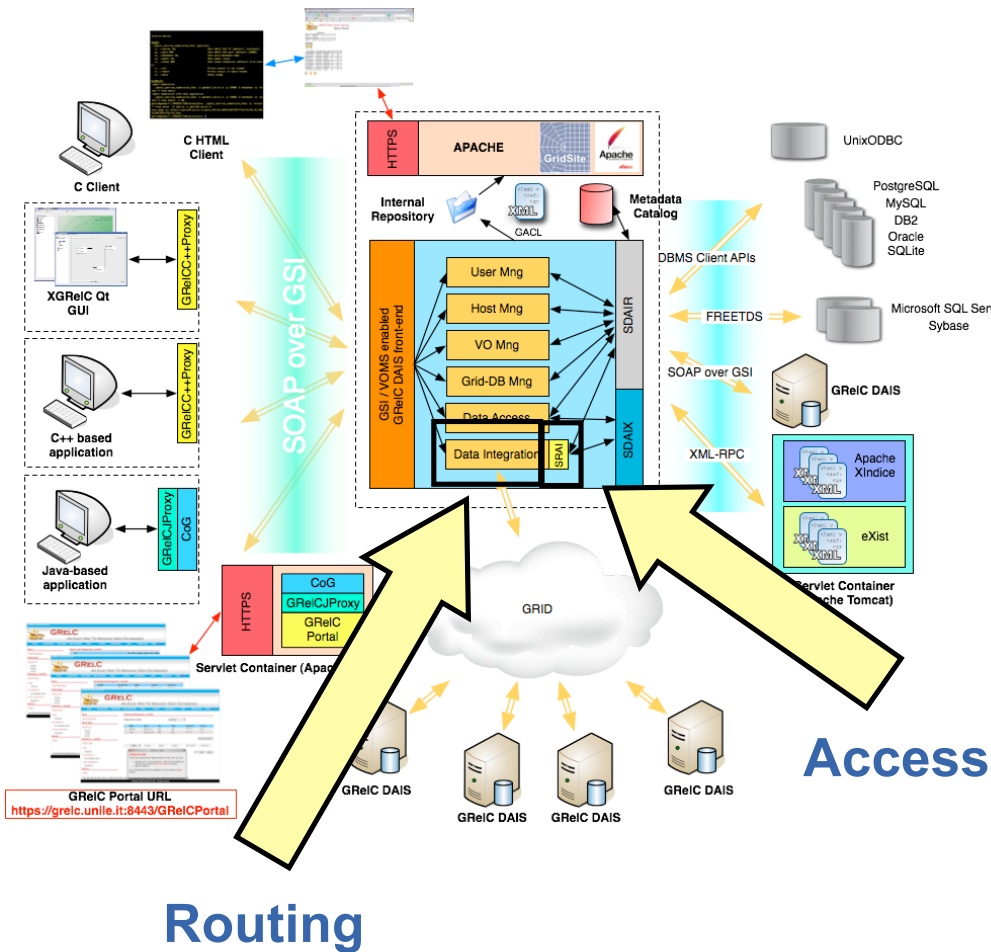
- Full security support by means of Grid Security Infrastructure (GSI) :
 - Mutual Authentication
 - Authorization based on local ACL and global VOMS support
 - Data Integrity
 - Data Confidentiality
 - ...

- **Loosely Coupled System:**

- Decoupling routing/access issues
- Higher (second) level of Data virtualization
- Based on P2P grid services & protocols
- Data integrated will concern:
 - Relational databases
 - XML Databases
 - Flat files
 - *Etc.*



Internals on GReIC DAIS



Routing

- Data integration module
- P2P (proprietary) Protocol
- Cycles discovery
- No duplication of query

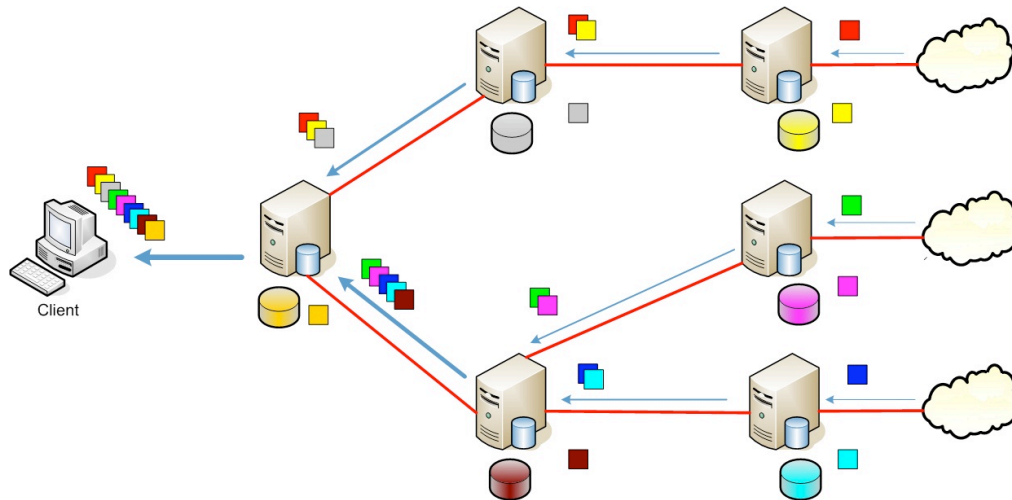
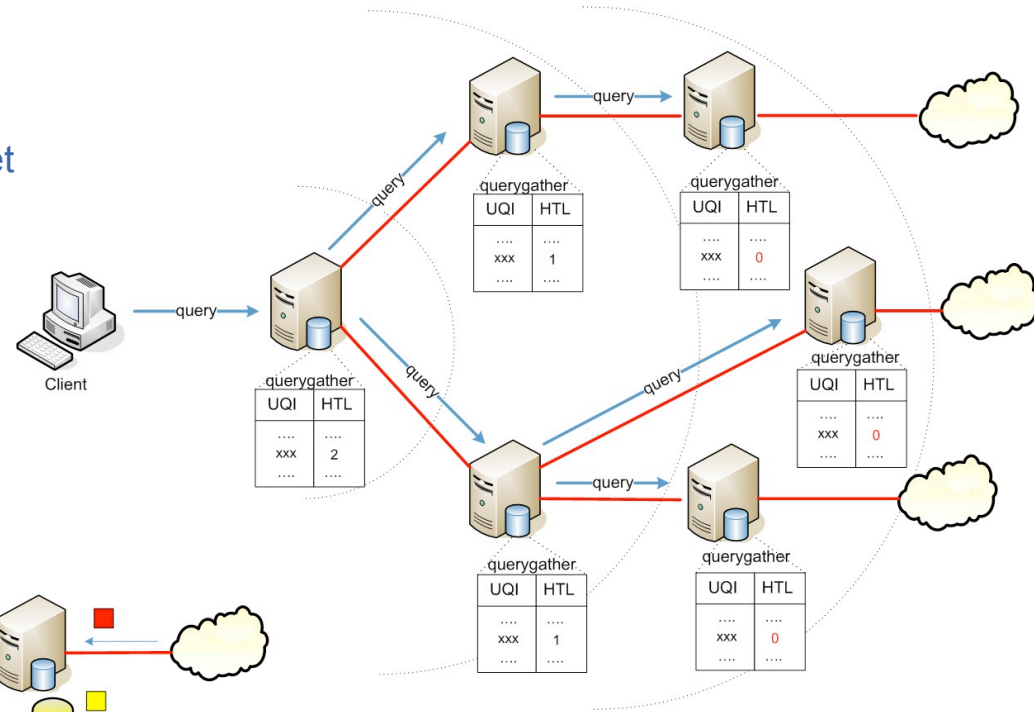
Access

— SRAI

- It decouples routing/access
- It interacts directly with:
 - *SDAIR (relational data sources)*
 - *SDAIX (XML data sources)*
- No direct interactions with end-systems.
- Completely independent of new back-end systems!

Key issues

1. UQI to identify the queries within the P2Pnet
2. HTL to provide a bound in terms of HOP
3. TTL to provide a bound in terms of time
4. VO centric support (scalable and flexible)
5. Full GSI Support (delegation of credentials)
6. Synchronous and asynchronous support



Four phases

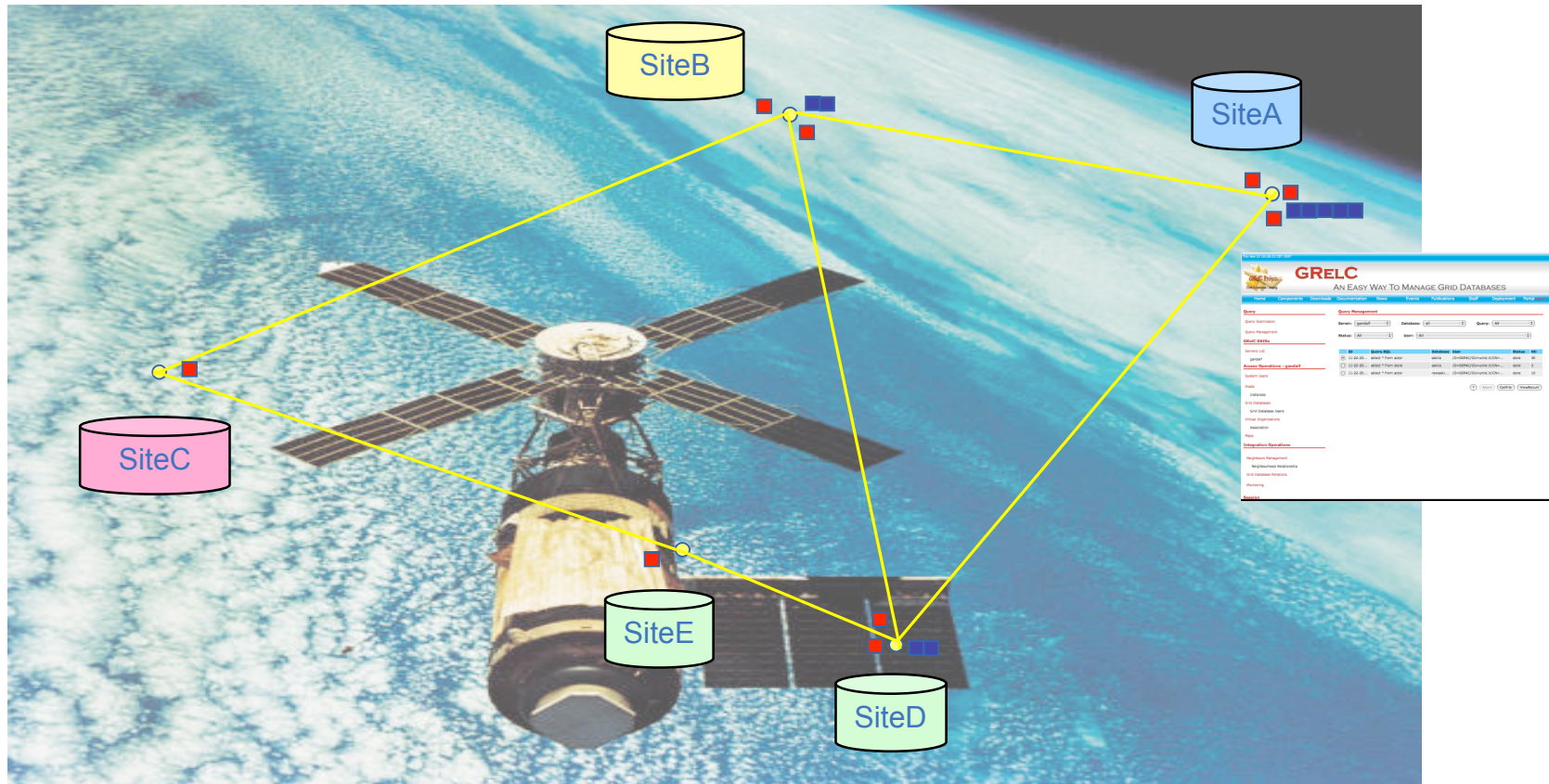
1. Query Forward
2. Querying local resources
3. Merging of partial results
4. Data retrieving

A Data Grid Integration Scenario for Metadata Mng

1. Query Forward

2. Query Submission

3. Collecting resultset





GReIC Portal New Web Pages DAIS

- P2P Mng

The screenshot displays the GReIC Portal interface for managing grid databases. The main header includes the GReIC logo and the tagline "AN EASY WAY TO MANAGE GRID DATABASES". A navigation menu contains links for Home, Components, Downloads, Documentation, News, Events, Publications, Staff, Deployment, and Portal (BETA). The left sidebar lists various management categories: Query, GReIC DAISs, Access Operations - gandalf, System Users, Hosts, Grid Databases, Virtual Organizations, Integration Operations, and Session.

The main content area is titled "MANAGE GRID DATABASES" and shows the current GReIC DAS as "gandalf". It features a dropdown menu for selecting a database, currently set to "sakila". Below this is a table of database relations:

DB virtual name	DB virtual name
<input type="checkbox"/> sakila	newsakila

Below the table, there is a "+ sakila" dropdown, a "- SELECT DATABASE -" dropdown, and buttons for "Remove Selected", "?", "Reset", and "Apply".

On the right side, there is a table for managing database instances:

Server port	DB virtual name
00	mydb


Below this table, there is a "Remove Selected" button, a dropdown menu for "<db virtual name>", and buttons for "?", "Reset", and "Apply".



GRELC Portal New Web Pages DAIS

- Database Log
- System Log

Mon Jan 28 19:01:16 CET 2008



GRELC

AN EASY WAY TO MANAGE GRID DATABASES

Home Components Downloads Documentation

Query

Query Submission

Query Management

GRELC DAISs

Servers List

grelc02

giancarlo

Databases List

Access Operations - giancarlo

System Users

Hosts

Instances

Grid Databases

Grid Database Users

Grid Database Log

Virtual Organizations

Association

GRELC DAIS Log

Mon Jan 28 19:01:39 CET 2008



GRELC

AN EASY WAY TO MANAGE GRID DATABASES

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

Query

Query Submission

Query Management

GRELC DAISs

Servers List

grelc02

giancarlo

Databases List

Access Operations - giancarlo

System Users

Hosts

Instances

Grid Databases

Grid Database Users

Grid Database Log

Virtual Organizations

Association

GRELC DAIS Log

Grid Database Log - giancarlo

Grid Database:

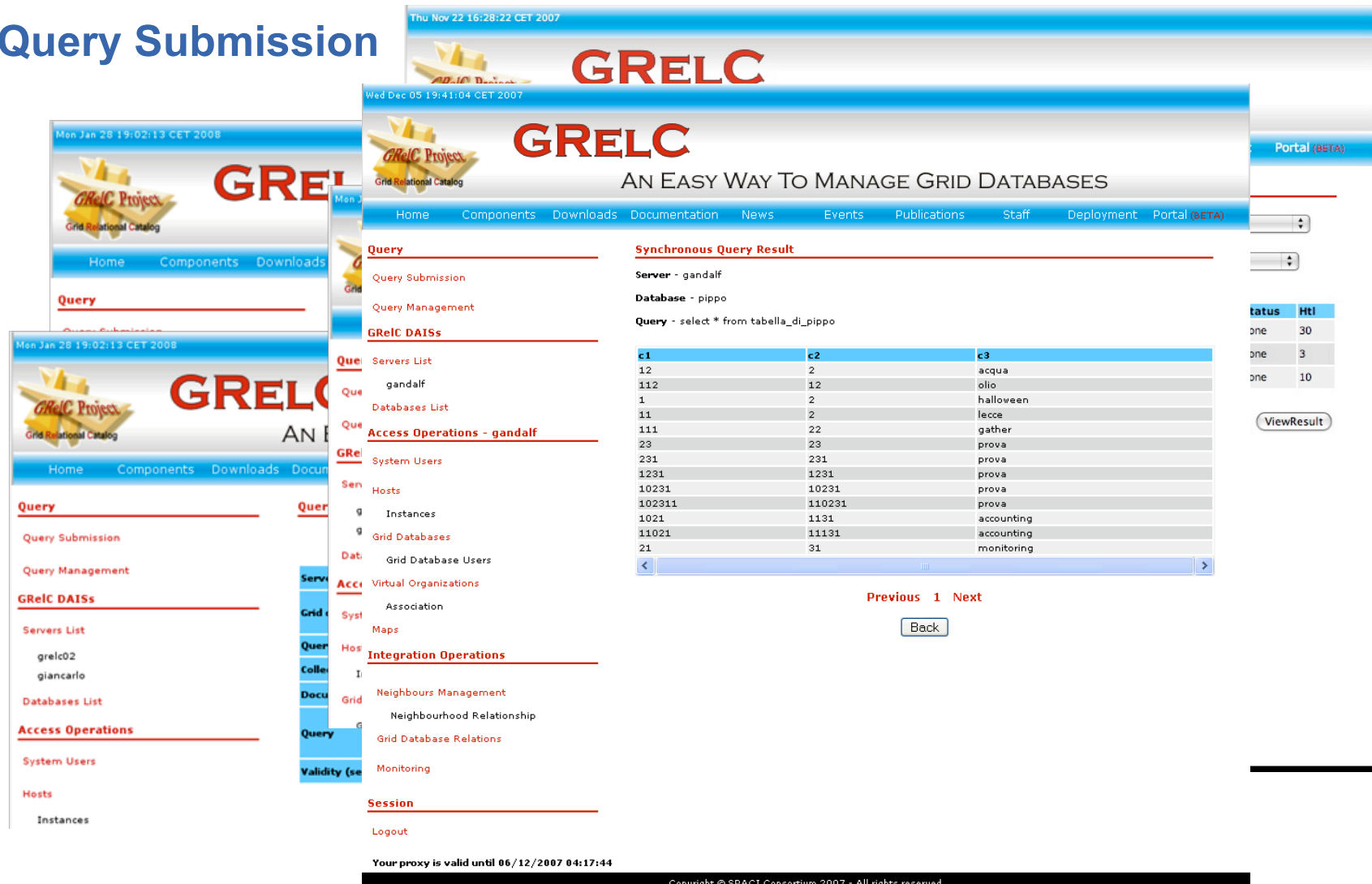
Log Type:

Number of lines:

[View Log](#)

Date	Operation	Type	Message
16/01/2008 17:13:50	ACCESS	SYSTEM	(grelc_data_access_xml_query_submission_one_shot) Submitting one-shot query (/) - user (/C=IT/O=INFN/OU=Host/L=HPCC University of Lecce/CN=grelc04.unile.it)
28/01/2008 18:47:57	ACCESS	DEBUG	(grelc_data_access_grelc_service_get_log_database) Executing get database log - user (/C=IT/O=INFN/OU=Host/L=HPCC University of Lecce/CN=grelc04.unile.it)
28/01/2008 18:50:26	ACCESS	SYSTEM	(grelc_data_access_xml_query_submission_one_shot) Submitting one-shot query (/Product[Color=Yellow]) - user (/C=IT/O=INFN/OU=Host/L=HPCC University of Lecce/CN=grelc04.unile.it)
28/01/2008 18:50:46	ACCESS	SYSTEM	(grelc_data_access_xml_query_submission_one_shot) Submitting one-shot query (/Product[Color=Yellow]) - user (/C=IT/O=INFN/OU=Host/L=HPCC University of Lecce/CN=grelc04.unile.it)
28/01/2008 18:50:46	ACCESS	ERROR	(grelc_data_access_xml_query_submission_one_shot) Query submission failed
28/01/2008 18:50:56	ACCESS	SYSTEM	(grelc_data_access_xml_query_submission_one_shot) Submitting one-shot query (/Product[Color=Yellow]) - user (/C=IT/O=INFN/OU=Host/L=HPCC University of Lecce/CN=grelc04.unile.it)
28/01/2008 18:52:11	ACCESS	SYSTEM	(grelc_data_access_xml_query_submission_one_shot) Submitting one-shot query (/Product[Color=Yellow]) - user (/C=IT/O=INFN/OU=Host/L=HPCC University of Lecce/CN=grelc04.unile.it)
28/01/2008 19:01:01	ACCESS	DEBUG	(load_system_privileges) Connection from /C=IT/O=INFN/OU=Host/L=HPCC University of Lecce/CN=grelc04.unile.it
28/01/2008 19:01:16	ACCESS	DEBUG	Accepted connection from 127.0.0.1
28/01/2008 19:01:16	ACCESS	DEBUG	(load_system_privileges) Connection from /C=IT/O=INFN/OU=Host/L=HPCC University of Lecce/CN=grelc04.unile.it
28/01/2008 19:01:16	ACCESS	DEBUG	(grelc_data_access_grelc_service_get_log) Executing get system log - user (/C=IT/O=INFN/OU=Host/L=HPCC University of Lecce/CN=grelc04.unile.it)

- Query Submission



The screenshot displays the GReLC Portal interface. The main header includes the GReLC logo and the tagline "AN EASY WAY TO MANAGE GRID DATABASES". A navigation menu contains links for Home, Components, Downloads, Documentation, News, Events, Publications, Staff, Deployment, and Portal (BETA). The central content area is titled "Synchronous Query Result" and shows the following details:

- Server:** gandalf
- Database:** pippo
- Query:** select * from tabella_di_pippo

The query results are presented in a table with three columns: c1, c2, and c3. The table contains 21 rows of data, including values like 12, 112, 1, 11, 111, 23, 231, 1231, 10231, 102311, 1021, 11021, and 21. Navigation controls for the table include "Previous", "1", and "Next" buttons, along with a "Back" button.

On the right side of the results, there is a small table with the following data:

tatus	Htl
one	30
one	3
one	10

A "ViewResult" button is located below this table. The footer of the page indicates "Your proxy is valid until 06/12/2007 04:17:44" and "Copyright © SPACI Consortium 2007 - All rights reserved".



GRELC Portal New Web Pages DAIS

- Monitoring

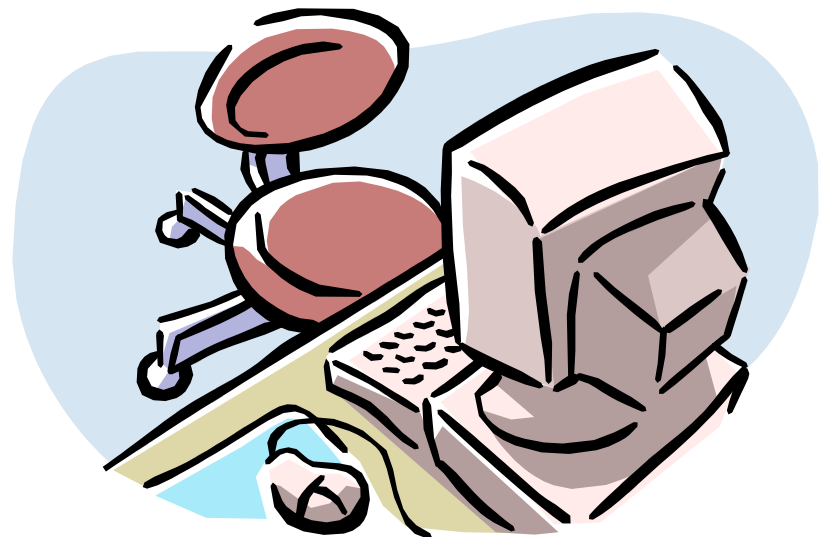
The screenshot displays the GRELC web portal interface. At the top, it shows the date 'Wed Feb 06 13:29:23 CET 2008' and the GRELC logo with the tagline 'AN EASY WAY TO MANAGE GRID DATABASES'. A navigation menu includes 'Home', 'Components', 'Downloads', 'Documentation', 'News', 'Events', 'Publications', 'Staff', 'Deployment', and 'Portal (BETA)'. The main content area is titled 'Monitoring - Agent node: gandalf' and features a map of Europe with a blue pin over the location of SPACI - Lecce. A pop-up window for 'SPACI - Lecce' provides the administrator's email (salvatore.vadacca@unile.it) and the number of queries (51). Below this, a pie chart shows the status of queries: Done (the largest slice), Queued, Other, and Failed. A bar chart below the pie chart shows 'Databases per VO (Total no.: 9)', with 'spaci' having the highest count at 6, followed by 'gilda' at 3, 'ingv' at 2, and 'libri' at 1. The left sidebar contains a 'Query' section with links for 'Query Submission' and 'Query Management', a 'GRELC DAISs' section with a 'Servers List' (including 'gandalf', 'greld02-gather', 'greld03', and 'grid009'), and an 'Access Operations - gandalf' section with various sub-sections like 'System Users', 'Hosts', 'Instances', 'Grid Databases', 'Virtual Organizations', 'Maps', 'Integration Operations', and 'Session'. A 'Logout' link is also present at the bottom of the sidebar.

Some important numbers about the GReIC project:

- 1 Patent
- About 20 International works
- More than 100.000 **code lines**
- 103 **services**
- **Wide documentation**
- **8 Major releases**
- 5 online tutorials on **GILDA**

Technologies:

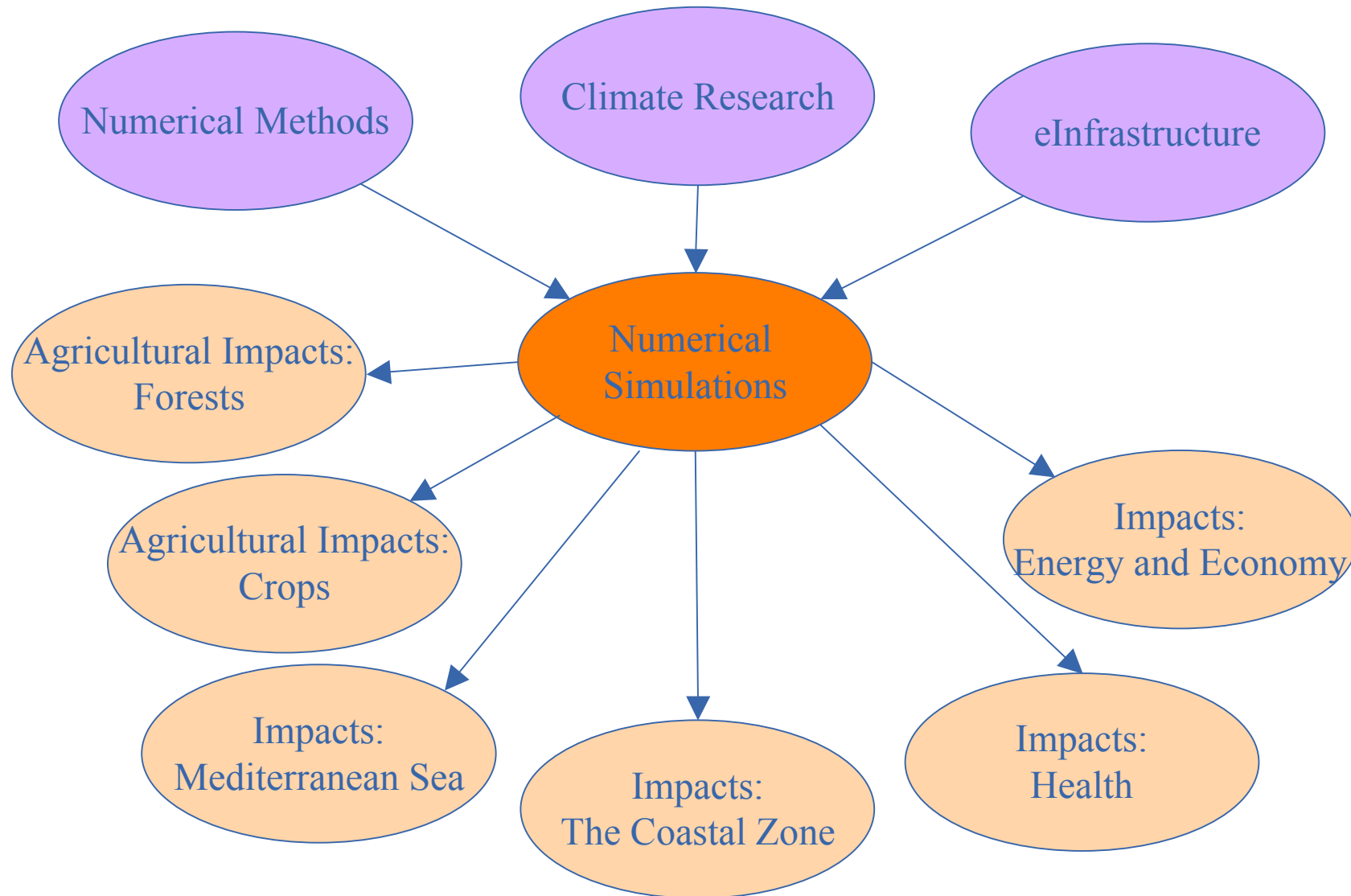
- GSI
- gSOAP
- GSI-plugin
- SDAI Library



The Euro-Mediterranean Centre for Climate Change (CMCC) is a facility of scientific research that aims to deepen our knowledge in the field of climate change, its causes and consequences, by developing numerical simulations of global models of the Earth System as well as regional models.

- Research Divisions (**SCO**, ANS, CIP, ISC, IAFENT, FDD)
- Partners (INGV, UNILE, CIRA, etc.)
- Associated Centres (SPACI, CRMPA, etc.)





CMCC: An Integrated and Ubiquitous Grid Environment for Climate Changes

Grid Paradigm:

Data & Computational grid

CMCC Environment:

acts as an incubator for the proposed technologies

Interdisciplinary:

Climate and Computer Scientists

Key Points:

Transparency and Interoperability

Expertise and Know-how

Grid & Computer Scientists (Unile)
SPACI support

Middleware:

Globus, Unicore, gLite, etc.

Metadata Mng:

Grid Metadata Handling System (GMHS)

● Associate Centers ● Partners



CMCC Data Grid Middleware Design

- Assumptions
 - Few Services
 - WS based for interoperability
 - Choosing the best for each service



- State of the art analysis
 - Middleware
 - *Globus (USA)*
 - *gLite (Europe)*
 - ...
 - Focus on Data Grid Services
 - Metadata Services
 - Data Access & Integration
 - Distributed Storage Mng
 - Data Grid Portals





Data Management @ CMCC - Phase1

Alias METADATA

CMCC Metadata Services

- Grid based solutions
- Centralized Solution
 - Initially deployed
 - Based on GReIC DAS
 - *Centralized Metadata Mng*
- Distributed Solution
 - Work in progress
 - GReIC DAIS
 - *P2P Solution*
 - *Metadata distribution*

CMCC Data Distribution Center

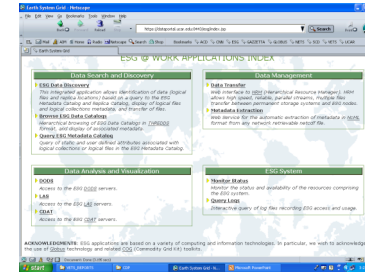
- CMCC Data Grid Portal
- Data oriented functionalities
- Search and discovery of metadata
 - Dataset browsing
 - Editing functionalities

CMCC Metadata Agreement

- Standard Analysis
 - ISO19115 / ISO19139
 - Dublin Core Metadata
 - Other standards and schema currently used
 - Link with FP7/METAFOR
- Schema definition
 - Design and schema implementation
 - **CMCC Working Group**
 - *Interdisciplinary Group*
 - *Climate and Computer scientists*
 - Schema describes
 - *Models*
 - *Algorithms*
 - *Datasets*
 - ...

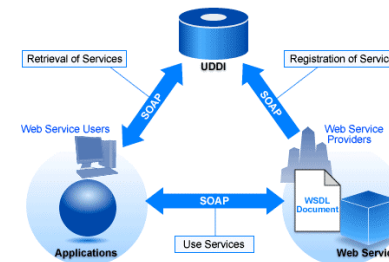
Data Grid Portals, Federated Gateways

Several entrance point to the P2P network
metadata catalogue, browsing, search, discovery, etc.



Standardized data request interface

Hides the complexity of specific data access mechanisms and addresses interoperability (SOA approach)



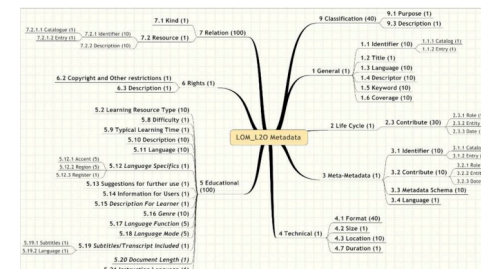
Automatic ingestion of metadata

Metadata is ingested after an automatic process of extraction from data (plugin based architecture, to extract metadata from different file formats)



Standardized metadata

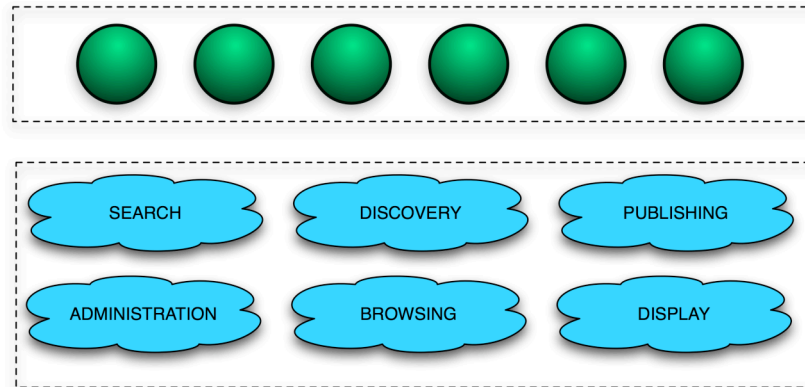
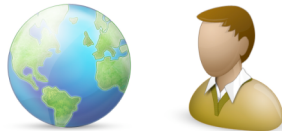
ISO 19115, XML, etc.



- A Metadata Grid Service Infrastructure
 - GReIC Project based solution
 - Moving from GReIC DAS to GReIC DAIS
 - Data Access and Integration capabilities
 - Scalable approach to distributed database management
 - P2P and Grid Protocols/Services
 - CMCC customization
 - GReIC DAIS
 - Deployment is ongoing
 - Testing phase, experimental results in the next two months
 - Data Grid Portal for metadata access
 - Two step search & discovery process based on different data models
 - SOA based approach with full security support through GSI

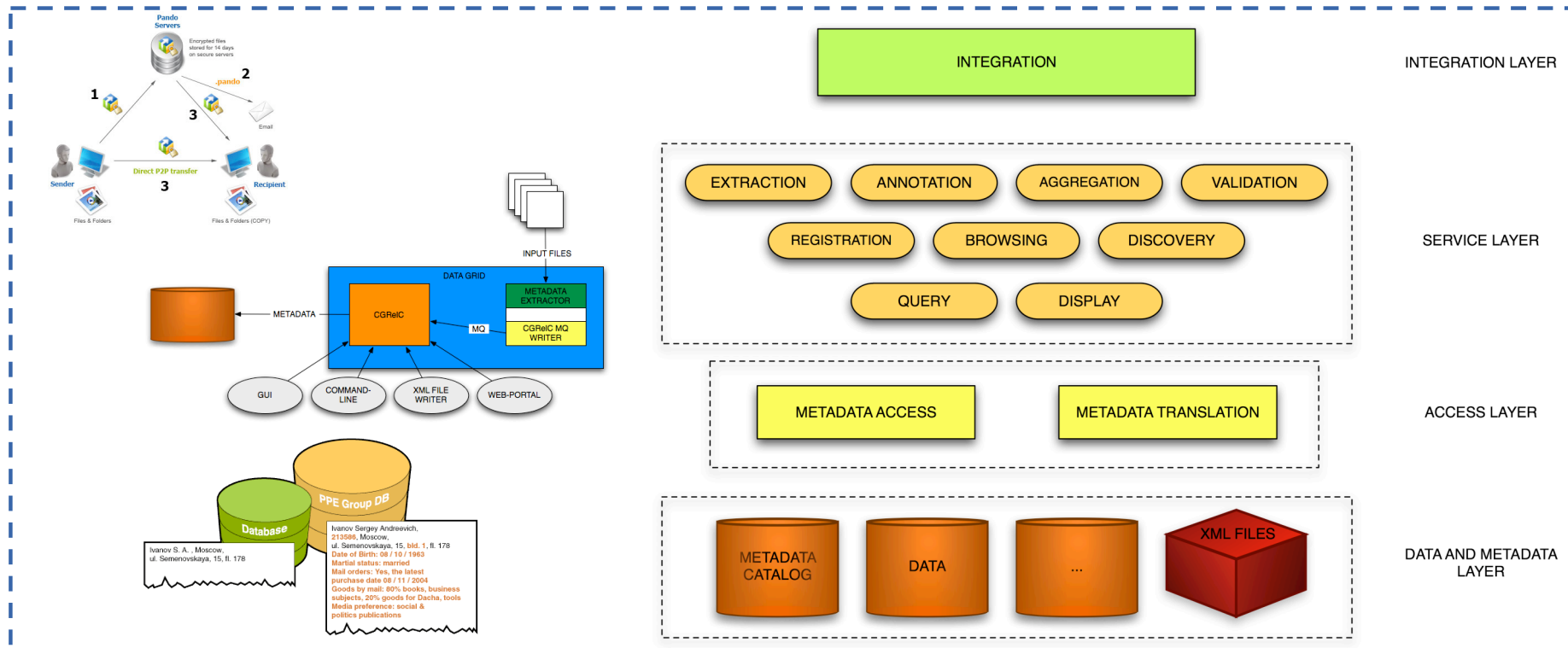


Metadata Management: Stack



APPLICATION LAYER

CLIENT API LAYER



CMCC: a fully distributed data grid environment

Grid Layer

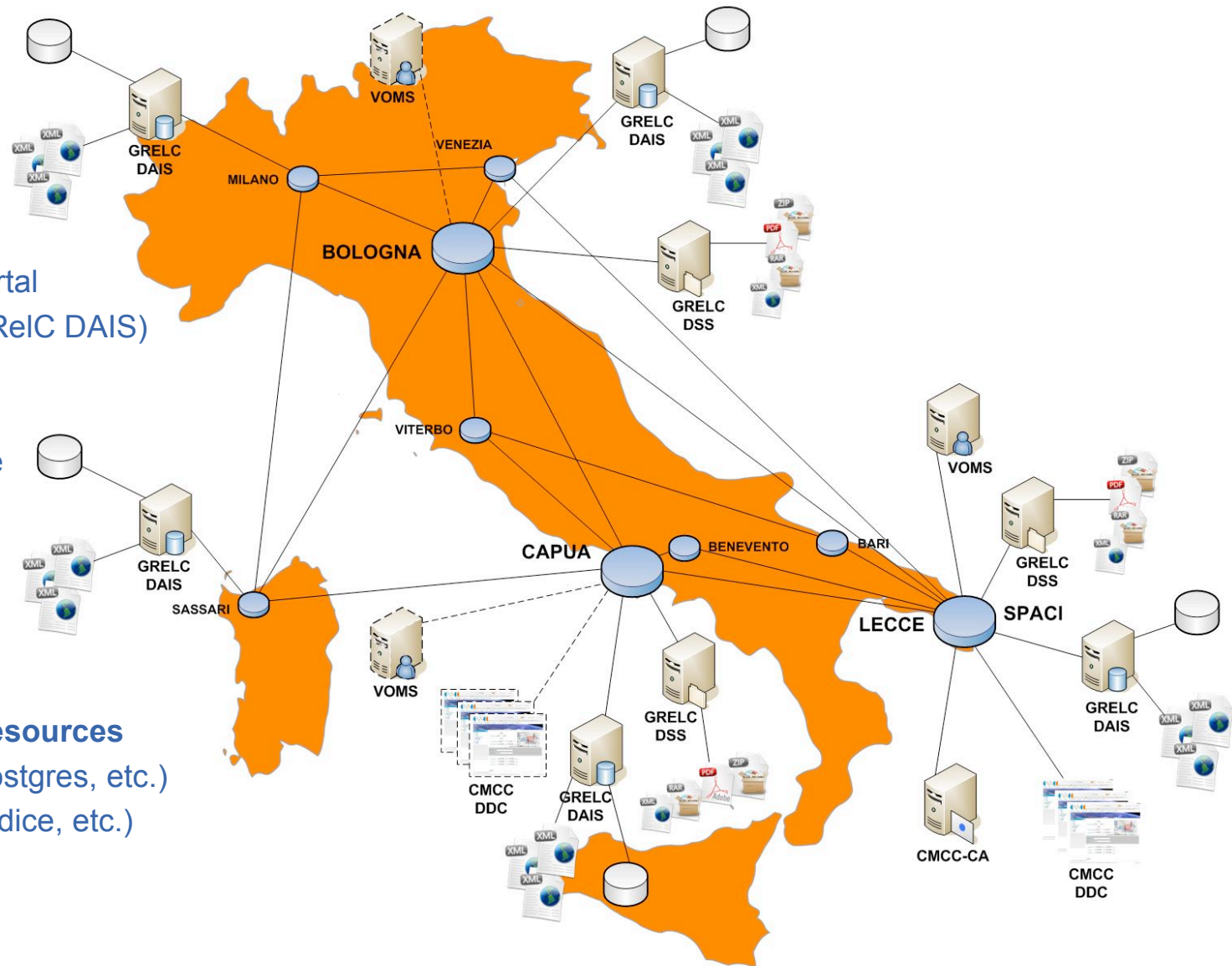
- VOMS Server
- CMCC Data Grid Portal
- Metadata Server (GReC DAIS)
- Grid Storage
 - GridFTP
 - GReC Storage
- CMCC CA
- P2P layer

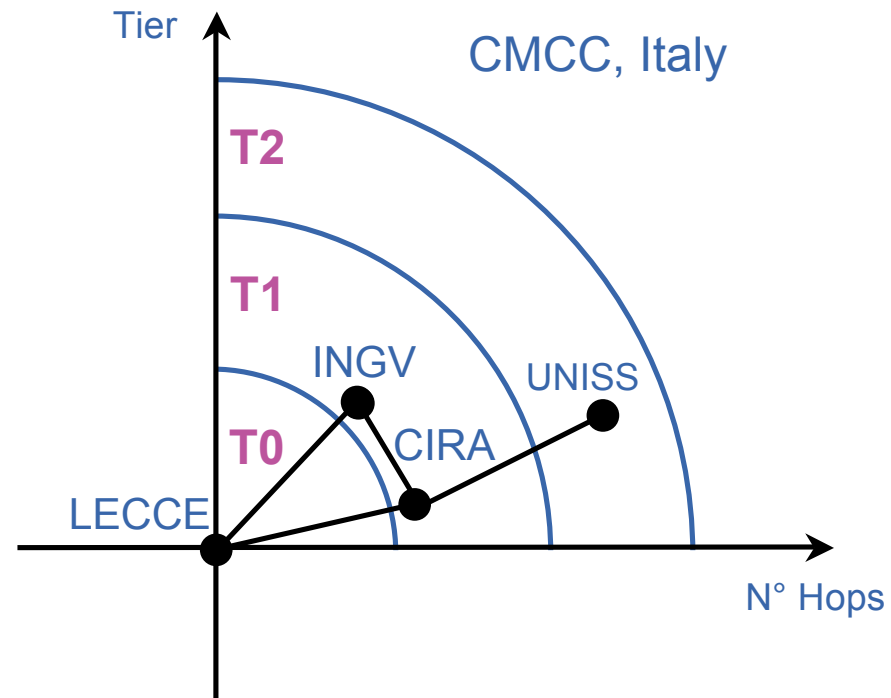
Data and Metadata resources

- RDBMS (MySQL, Postgres, etc.)
- XML DBs (eXist, XIndice, etc.)

Search & Discovery

- Two step process





- Deployment at Lecce (T0), Bologna, Capua (T1), Sassari (T2)
- Testing and certification phase on these 4 nodes
- Crossbar link T0 - T1
- T0, T1 superpeers for T2 sites

Key issues:

- GReIC DAIS 3.0
- CMCC GMHS
- EGEE08 (Poster)
- RESPECT Program
- CMCC Deployment



The screenshot shows the iSGTW website interface. The main content area features an article titled "Feature - GReIC @ Euro-Mediterranean Centre for Climate Change". The article discusses the challenges of managing large volumes of climate data and introduces the CMCC Grid Metadata Handling System (GMHS). It mentions the use of GReIC DAIS and the SPACI Consortium. A sidebar on the right contains search and navigation options, including a search bar, a calendar for July 2008, and announcements.

Feature - GReIC @ Euro-Mediterranean Centre for Climate Change

Drowning in data

Talk to any [climate scientist](#), and they'll tell you that one of their biggest problems is dealing with the sheer volume of data they acquire when dealing with something as complex as worldwide climate change. They don't lack for sources of data; if anything, they have too much to choose from—everything from tree ring data to ice cores to collections of air samples from hot-air balloons.

With this in mind, managing the metadata can be a bottleneck. But distributed, peer-to-peer and grid-enabled solutions can help to provide a secure, centralized, transparent and scalable solution for managing petabytes of datasets spread among several sites, say researchers Sandro Fiore and Giovanni Aloisio at the [EuroMediterranean Center for Climate Change \(CMCC, Italy\)](#). They have developed what they call the CMCC Grid Metadata Handling System (GMHS), which provides both access to and integration of climate metadata stored in different and widespread data sources, while allowing scalability, transparency and efficiency.

The middleware for their system is based upon [GReIC DAIS](#), adopted as a data grid enabling technology, developed by the [SPACI Consortium](#) and the University of Salento under the [Grid Relational Catalog \(GReIC\) Project](#). The software, compatible both with gLite and Globus, is part of the [IGI](#) release, and was recently included in the [EGEE Respect Program](#) and presented at the last two EGEE events, at the EGEE07 conference in Budapest and the 3rd EGEE User Forum in Clermont-Ferrand. It was mentioned as a runner-up during the related demo sessions—in both demos, Fiore and the GReIC Team showed practical scenarios involving the distributed climate metadata management at CMCC.

Supporting science

"The CMCC Data Grid framework," says Aloisio, head of Scientific Computation and Operation

Main Sites:

- **CMCC**

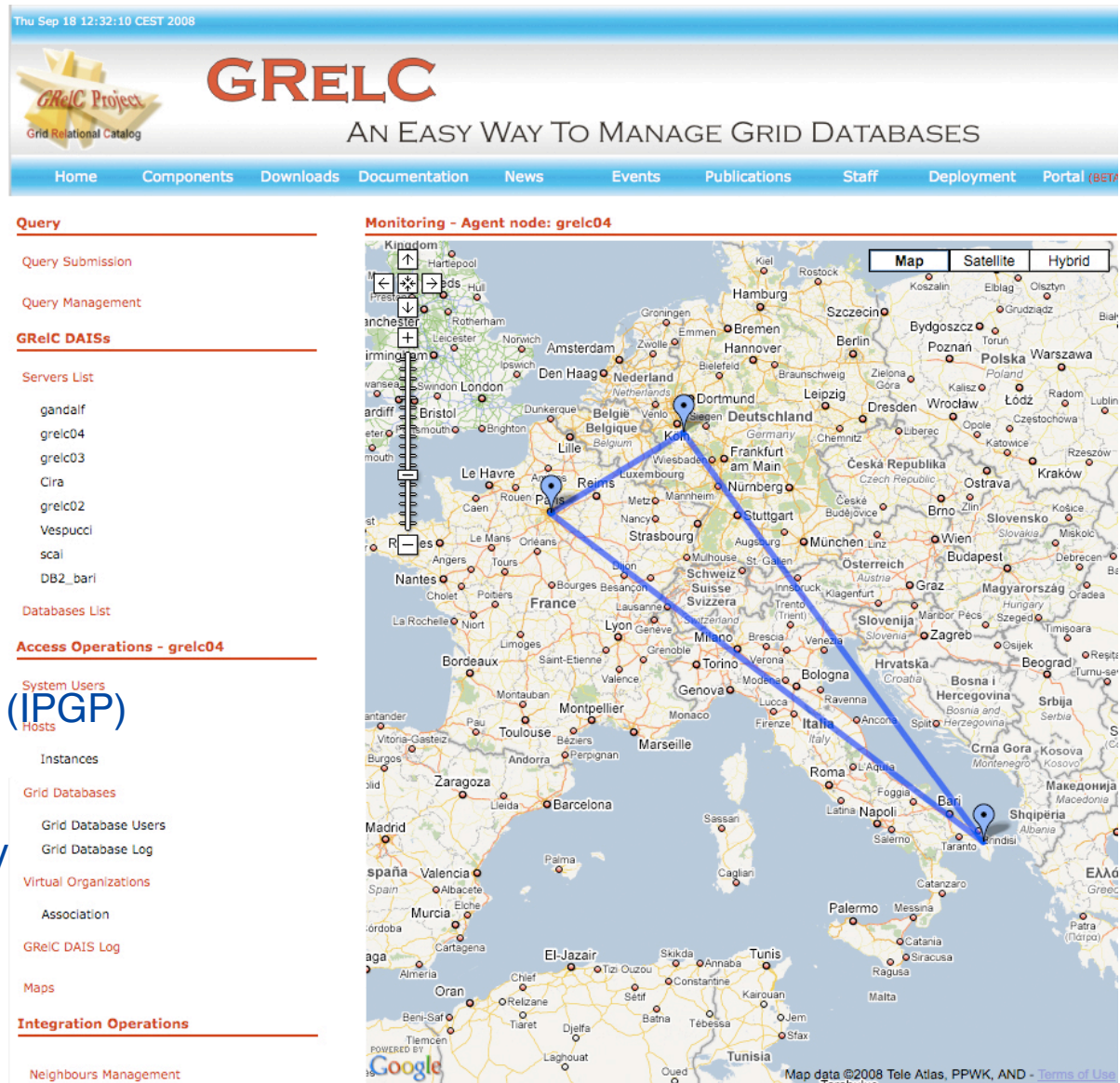
- Site: Lecce, Italy
- Resp: G. Aloisio
- Fiore, Negro, Vadacca

- **IPSL-IPGP**

- Site: Paris, France
- Resp: M. Petitdidier
- Denvil (IPSL), Weissenbach (IPGP)

- **Fraunhofer SCAI**

- Site: St. Augustin, Germany
- Resp: H. Schwichtenberg
- Gemuend



Thu Sep 18 12:32:10 CEST 2008

GRELC
AN EASY WAY TO MANAGE GRID DATABASES

Home Components Downloads Documentation News Events Publications Staff Deployment Portal

Query

- Query Submission
- Query Management

GRELC DAIS

Servers List

- gandalf
- grelc04
- grelc03
- Cira
- grelc02
- Vespucci
- scal
- DB2_barl

Databases List

Access Operations - grelc04

- System Users
- Hosts
- Instances
- Grid Databases
 - Grid Database Users
 - Grid Database Log
- Virtual Organizations
 - Association
 - GRELC DAIS Log
- Maps
- Integration Operations
 - Neighbours Management

Monitoring - Agent node: grelc04

Map Satellite Hybrid

The map shows a blue line connecting several cities across Europe: Paris, France; Dortmund, Germany; and Bari, Italy. Other cities visible include London, Amsterdam, Berlin, Frankfurt, Munich, Rome, and Athens.

Map data ©2008 Tele Atlas, PPWK, AND - Terms of Use



European Testbed

Thu Sep 18 12:34:19 CEST 2008
Thu Sep 18 12:29:22 CEST 2008

GRELC

AN EASY WAY TO MANAGE GRID DATABASES

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

Query

[Query Submission](#)

[Query Management](#)

GRELC DAISs

[Servers List](#)

[Databases List](#)

Access Operations - grelc04

[System Users](#)

[Hosts](#)

[Grid Databases](#)

[Virtual Organizations](#)

[Maps](#)

Integration Operations

[Neighbours Management](#)

[Monitoring](#)

Session

[Logout](#)

Gather Query Result

Server - grelc04

idmetadata	mname
1	ENSEMBL
2	ENSEMBL
1	ENSEMBL
2	ENSEMBL
3	ENSEMBL
1	ENSEMBL
2	ENSEMBL
3	ENSEMBL

Synchronous Query Result

```

- <result count="6">
+ <CONTACT contact_type="metadata" src:col="/db/cmcc" src:key="data1.xml">
  </CONTACT>
+ <CONTACT contact_type="metadata reviewer" src:col="/db/cmcc"
  src:key="data1.xml"></CONTACT>
+ <CONTACT contact_type="investigator" src:col="/db/cmcc" src:key="data1.xml">
  </CONTACT>
+ <CONTACT contact_type="metadata" src:col="/db/cmcc" src:key="data2.xml">
  </CONTACT>
+ <CONTACT contact_type="metadata reviewer" src:col="/db/cmcc"
  src:key="data2.xml"></CONTACT>
- <CONTACT contact_type="investigator" src:col="/db/cmcc" src:key="data2.xml">
  - <PERSON>
    <PERSON_ID>2000513</PERSON_ID>
    <FIRST_NAME>Jean-Louis</FIRST_NAME>
    <SECOND_NAME>n/a</SECOND_NAME>
    <LAST_NAME>Dufresne</LAST_NAME>
    <TITLE>Dr.</TITLE>
    <TELEPHONE>33 1 44 27 50 14</TELEPHONE>
    <FAX>33 1 44 27 62 72</FAX>
    <URI>n/a</URI>
    <EMAIL>jean-louis.dufresne@lmd.jussieu.fr</EMAIL>
  </PERSON>
  - <INSTITUTE>
    <INSTITUTE_ID>2000143</INSTITUTE_ID>
    <INSTITUTE_NAME>Institut Pierre Simon Laplace</INSTITUTE_NAME>
    <INSTITUTE_ACRONYM>IPSL</INSTITUTE_ACRONYM>
    <DEPARTMENT_NAME>IPSL Global Climate Modeling
    Group</DEPARTMENT_NAME>
    <DEPARTMENT_ACRONYM>IGCMG</DEPARTMENT_ACRONYM>
    <COUNTRY>France</COUNTRY>
    <STATE_OR_PROVINCE>n/a</STATE_OR_PROVINCE>
    <PLACE>Paris</PLACE>
    <STREET>4, Place de Jussieu</STREET>
    <STREET_POSTAL_CODE>75252</STREET_POSTAL_CODE>
    <POBOX>n/a</POBOX>
    <POBOX_POSTAL_CODE>cedex 05</POBOX_POSTAL_CODE>
    <URI>http://ipsl.jussieu.fr</URI>
    <ADDITIONAL_INFO>not filled</ADDITIONAL_INFO>
  </INSTITUTE>
  </CONTACT>
          
```

Your proxy is valid until 19/09/2008 00:00:40

Copyright © SPACI Consortium 2007 - All rights reserved

- GReIC DAS/DAIS provides support in Grid for a wide range of DBMSs.
- It is currently tested on several grid environments (**SEPAC, GILDA, INFNGRID**)
- A **wide SDK** (C, Java) is available for developers
- **CLI**, and **Portal Interfaces** to ease Grid-DB mng
- **gLite compliant** (porting on gLite 3.x/4.x and integration with **VOMS** framework, **BDII**, etc.)
- Support for several platforms (**IA32** and **IA64**)
- Part of the Italian (IGI) Release
- Currently included the **EGEE Respect** Program
- Adopted at the **Euro-Mediterranean Centre for Climate Change**



For any information



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

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

Team Members:

Ph. D. Massimo Cafaro

MSc Alessandro Negro

MSc Salvatore Vadacca

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

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