

ES metadata OGSA-DAI

NA4 GA Meeting, 09.01.06

D. Weissenbach, IPSL , France

H. Schwichtenberg, SCAI, Germany

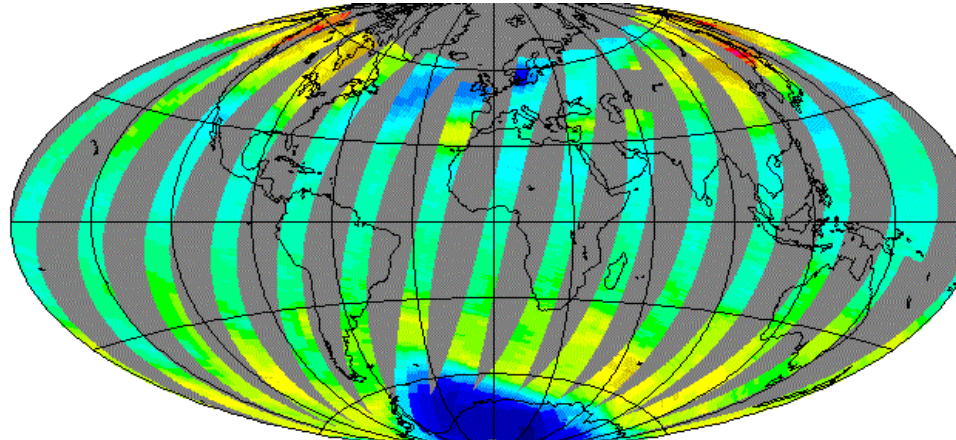
Overview

- ES Requirements
- OGSA-DAI
- ES meta data clients
- Future Plans

ES Requirements

- In ES, Metadata and data bases are very common as large data sets are handled by different teams. The RDBMS generally used are MySQL, PostgreSQL or Oracle
- Many databases already exist the aim is the implementation of an interface with EGEE or at least to access a copy of them.
- If new bases are created on EGEE they need to be accessible outside Grid.
- Some metadata and data are only accessible to authorized persons. Others available on web site have rules for publications(acknowledgement, co-author).
- Many queries concern matching in time and/or space
- Another specific point is the satellite data that are stored by orbit. The orbits are described by their footprint, expressed in geographical coordinates.

ES Requirements



The goal is to develop for a specific case a prototype that includes the needed tools:

Example: Validation of GOME/ERS experiment with Lidar data

Two different instruments : **Ground-based Lidar, spectrometer aboard the satellite, ERS.**

The satellite data stored by orbit or pixel; different algorithms

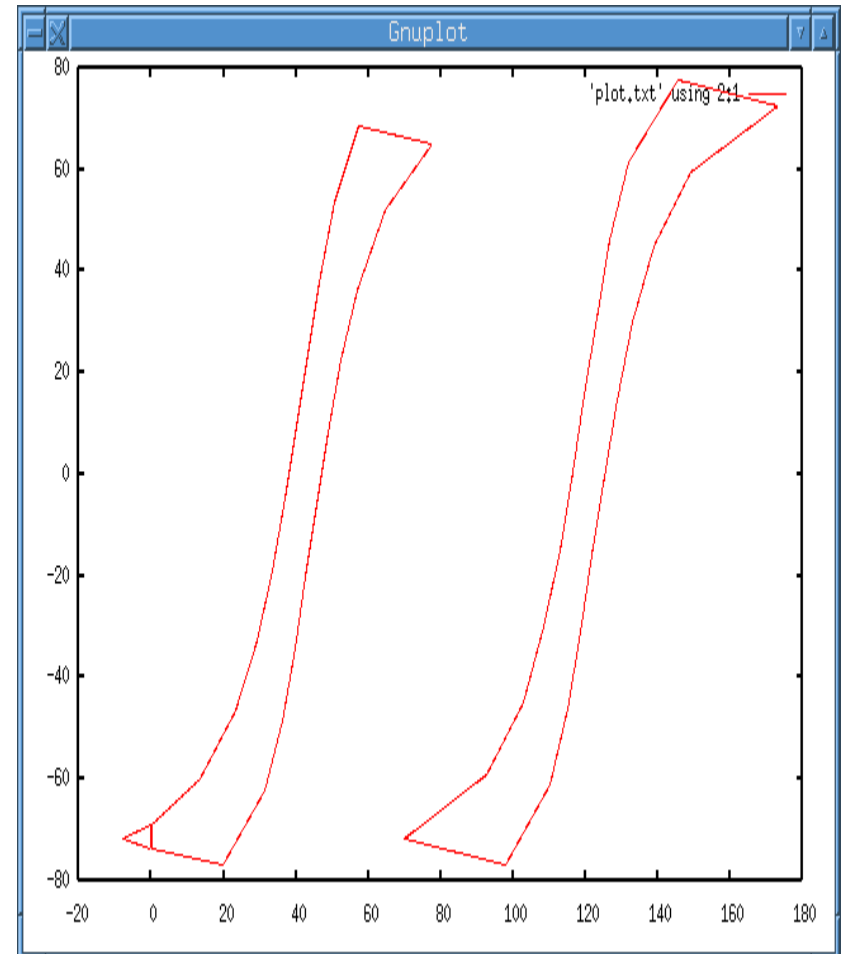
The Lidar data stored in monthly files with one profile/night

ES Requirements

ES evaluates common Metadatamodels (Standards) Today:

- Opera/NNO meta data scheme:

Column	Type
dataset	character varying(50)
level	character varying(5)
version	character varying(4)
orbit	integer
file_name	character varying(50)
start_date	timestamp without time zone
stop_date	timestamp without time zone
lat	numeric(8,2)
lon	numeric(8,2)
proc_center	character varying(50)
proc_date	timestamp without time zone
file_input	character varying(50)
proc_description	character varying(50)
footprint	geometry (Multipolygon)



ES Requirements

- Lidar meta data scheme (data stored in monthly files)

Column	Type	
		startdate,stopdate for /grid/esr/lidar/2000/oho30001.gol
station	character varying(50)	2000-01-03 19:25:00.0,2000-01-03 23:50:00.0
stationname	character varying(50)	2000-01-04 19:22:00.0,2000-01-05 00:33:00.0
sensor	character varying(50)	2000-01-05 17:22:00.0,2000-01-05 21:31:00.0
cpd	character varying(50)	2000-01-06 17:52:00.0,2000-01-06 22:01:00.0
producer	character varying(50)	2000-01-07 19:35:00.0,2000-01-07 23:44:00.0
fileinput	character varying(50)	2000-01-10 17:53:00.0,2000-01-10 22:11:00.0
file_name	character varying(50)	2000-01-11 17:34:00.0,2000-01-11 21:13:00.0
parameter	character varying(50)	2000-01-12 17:38:00.0,2000-01-12 22:09:00.0
footprintlat	numeric(8,2)	2000-01-13 20:50:00.0,2000-01-14 01:15:00.0
footprintlon	numeric(8,2)	2000-01-15 17:49:00.0,2000-01-16 06:12:00.0
startdate	timestamp without time zone	2000-01-16 17:38:00.0,2000-01-17 01:51:00.0
stopdate	timestamp without time zone	2000-01-17 17:37:00.0,2000-01-17 22:35:00.0
procdate	timestamp without time zone	2000-01-18 17:27:00.0,2000-01-18 21:36:00.0
proccenter	character varying(50)	2000-01-19 17:33:00.0,2000-01-19 21:58:00.0
		2000-01-20 17:28:00.0,2000-01-20 21:33:00.0
		2000-01-21 17:41:00.0,2000-01-21 22:06:00.0
		2000-01-23 17:38:00.0,2000-01-23 22:16:00.0
		2000-01-24 17:38:00.0,2000-01-24 23:09:00.0
		2000-01-25 17:43:00.0,2000-01-25 22:47:00.0
		2000-01-26 18:06:00.0,2000-01-26 23:24:00.0



ES Requirements

- Conclusion:
 - consistent, fast and reliable access
 - secure and restricted access to Metadata or data
 - industrial standards are preferable
(industry is part of ES)
 - the RDBMS needs to support spatial data types
(OpenGIS conform)

OGSA-DAI

Open Grid Services Architecture – Data Access and Integration

- What is OGSA-DAI

OGSA-DAI is a middleware product to access data resources, such as relational databases, in Grids see: www.ogsadai.org.uk

De facto standard for databases on the grid

contributes to standardisation efforts of GGF DAIS

reference implementation of DAIS spec, part of OMII Bundle,

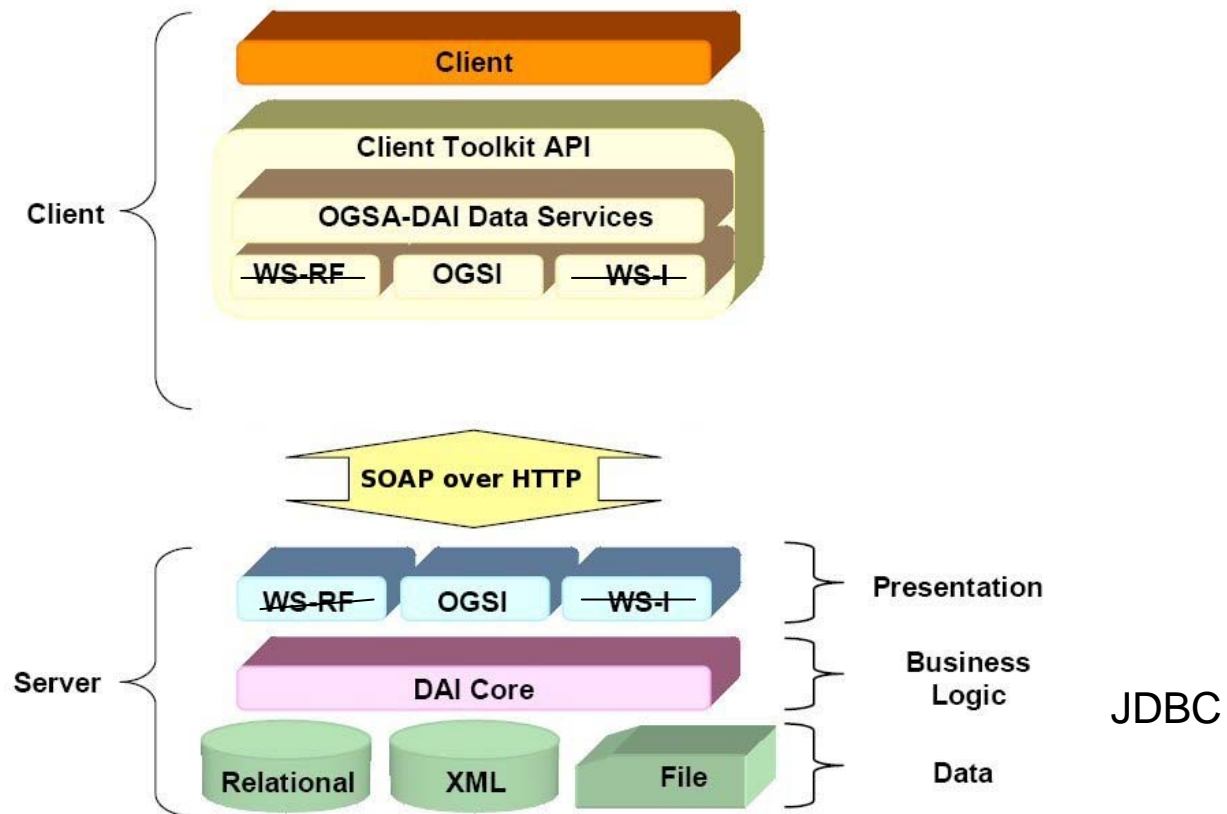
IBM and Oracle are partners

consists of 3 main services: **DAISGR** (registry) for discovery,

GDSF (factory) present data resource, **GDS** to access data resource

supports MYSQL; DB2, Oracle, PostgreSQL, SQLserver, XML (Xindice,..) and files

OGSA-DAI



OGSA-DAI

Installed Environment at SCAI

- SL 4.1
- Web-Service Container:
Tomcat 4.1.31
- OGSA-DAI:
OGSA-DAI OGSI 6.0 with GLOBUS 3.2.1
- RDBMS:
 - MySQL 4.1.10 (testing)
 - MySQL spatial extensions only support convex polygons
 - PostgreSQL 7.4.8 + PostGIS (production)
 - PostGIS adds support for geographic objects to Postgres: <http://postgis.refrations.net/>

OGSA-DAI

- User Authentication

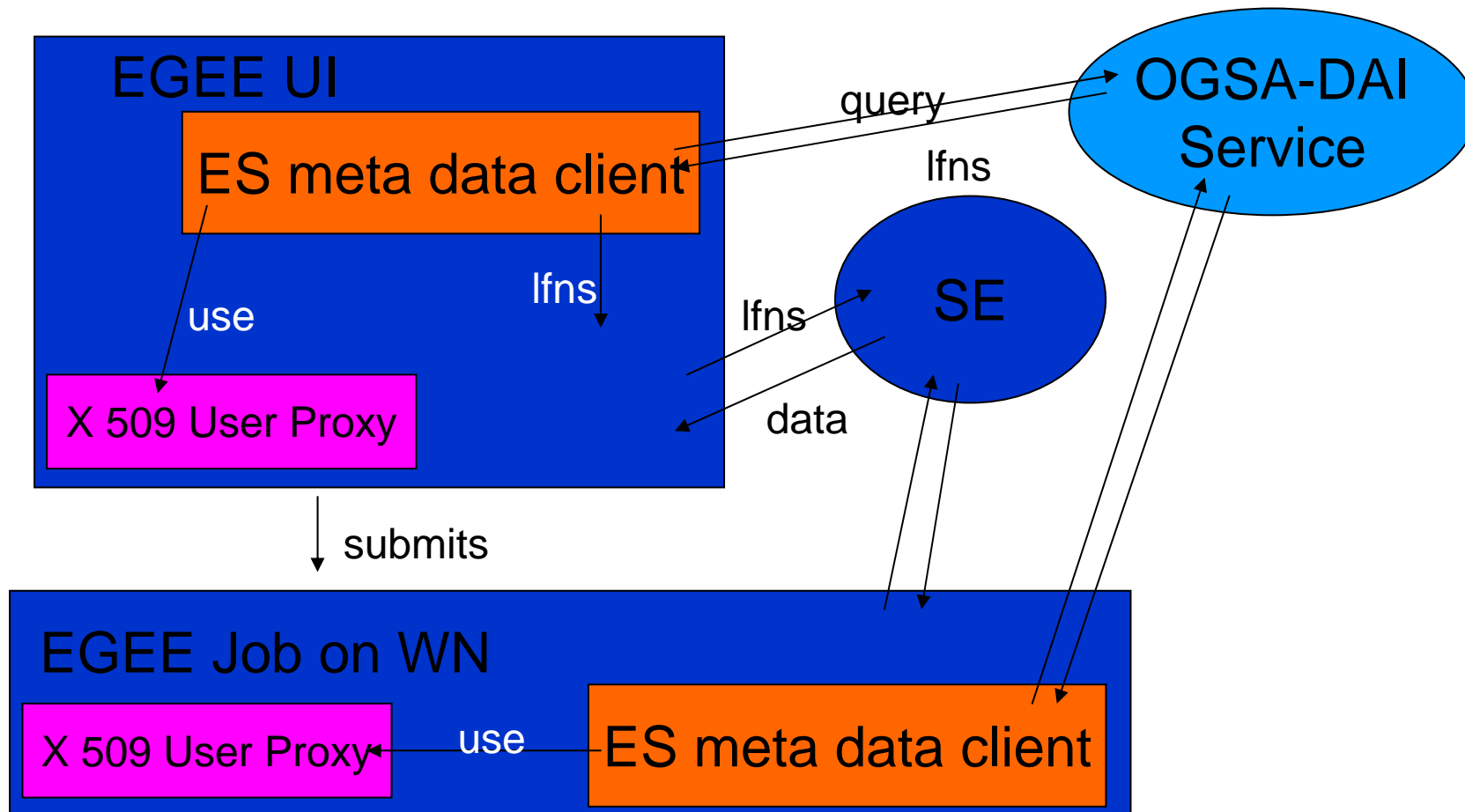
- with grid proxy certificates
- mapping to db roles for every user, no VO / VOMS support

```
<User dn="/O=GRID-FR/C=FR/O=CNRS/OU=IPSL/CN=David Weissenbach" userid="lidar_writer" password="****" />
```

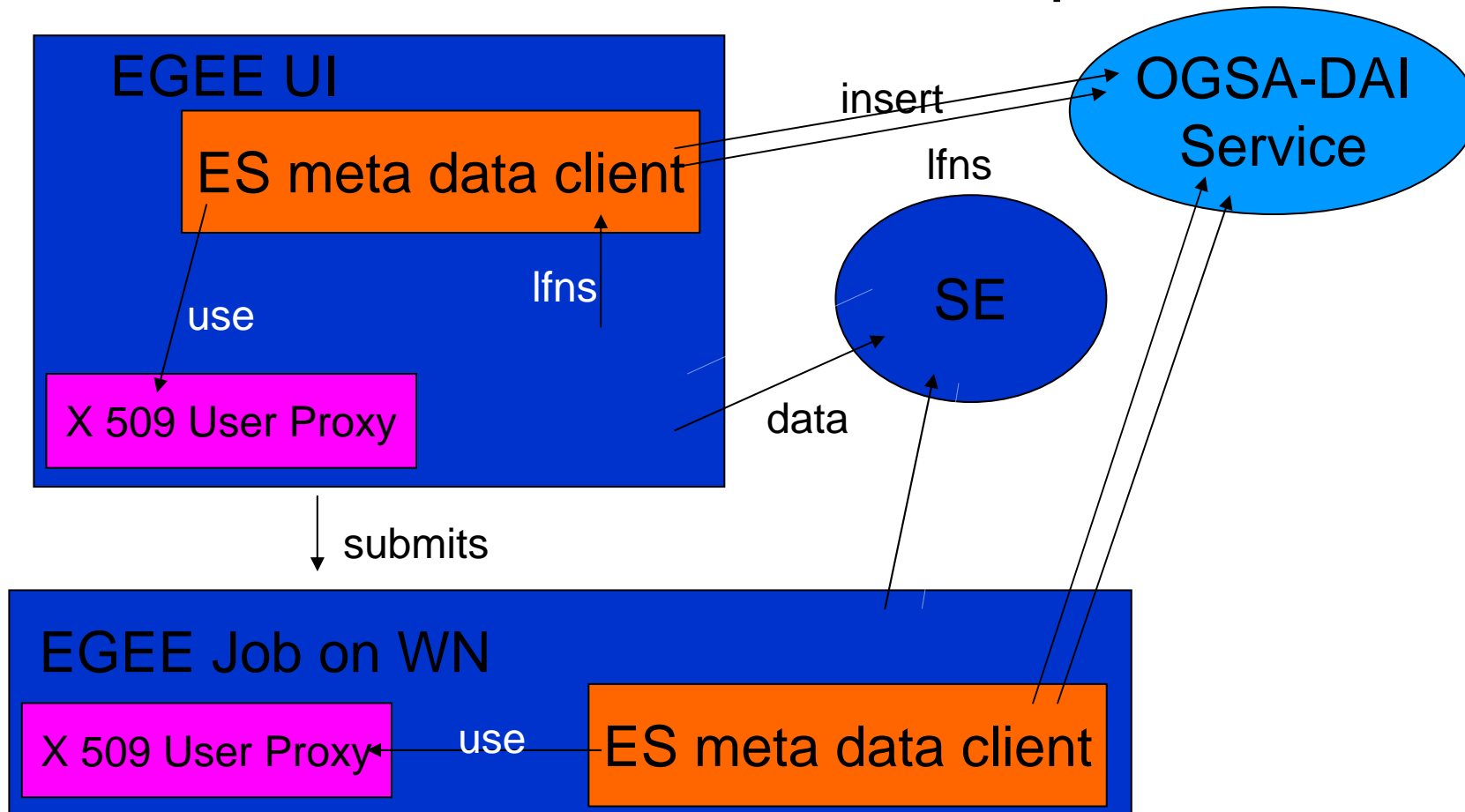
- Communication

- OGSA-DAI OGSI use message level security (mls)
(newest Version has mls and tls)

ES meta data clients : query



ES meta data clients : update



ES meta data clients

Database Aspects (Spatial Ext.)

- GIS representations as Text (WKT) or Binary (WKB)
 - eg `PolygonFromText ('POLYGON((1 1, 1 -1, -1 -1, -1 1, 1 1))')`
- Many types and mathematical operators available
 - functions (Area...), comparators (Contains, Intersects...)
- but Text is not interpreted by SQL engine
 - `SELECT a.x, a.y, b.geom FROM a, b WHERE Contains(PointFromText ('POINT(x y)') , geom)`
 x and y are Text, not numbers... Such a query is impossible.
- So we need to use embedded SQL to enable parametrized queries.

ES meta data clients

DB Aspects II (OGSA-DAI API)

- OGSA-DAI : a java API (good embeded SQL -JDBC- API)
- Very very easy to use with SCAI additionnal classes
 1. connect to service :


```
service = new OGSISecureDataService(URL)
```
 1. create the SQL query : `complex_query = new String()`
 2. send the query , eventually get the result or catch the Exception :


```
result=service.submitSQLQuery(new SQLQuery(str))
```
- One session can send as many queries as needed

No joint SELECT possible leads to multiple select operations with intermediate java String/Number conversions/combinations

Future Plans

- OGSA-DAI WSRF with GLOBUS 4
- Postgres 8
- Manage other (remote) ES applications (meta)data bases with the same OGSA-DAI server.
- Integrating with gLite java API

DEMO by David