

# Applications and the Grid

## The European DataGrid Project Team

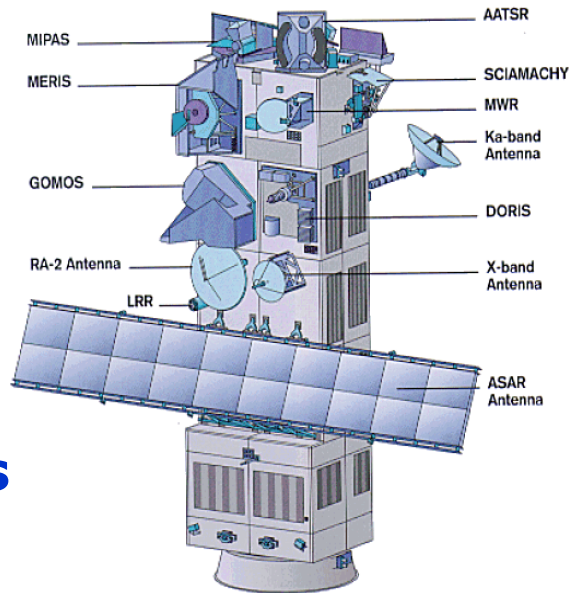
<http://www.eu-datagrid.org>



# EDG Application Areas

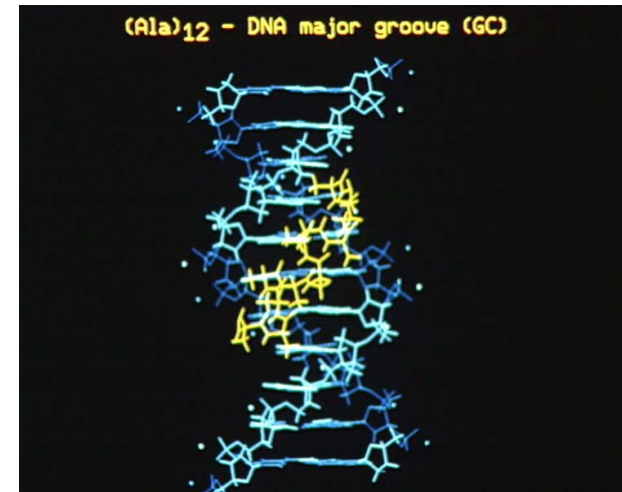


## Earth Observation Science Applications



## High Energy Physics

## Biomedical Applications



# High Energy Physics

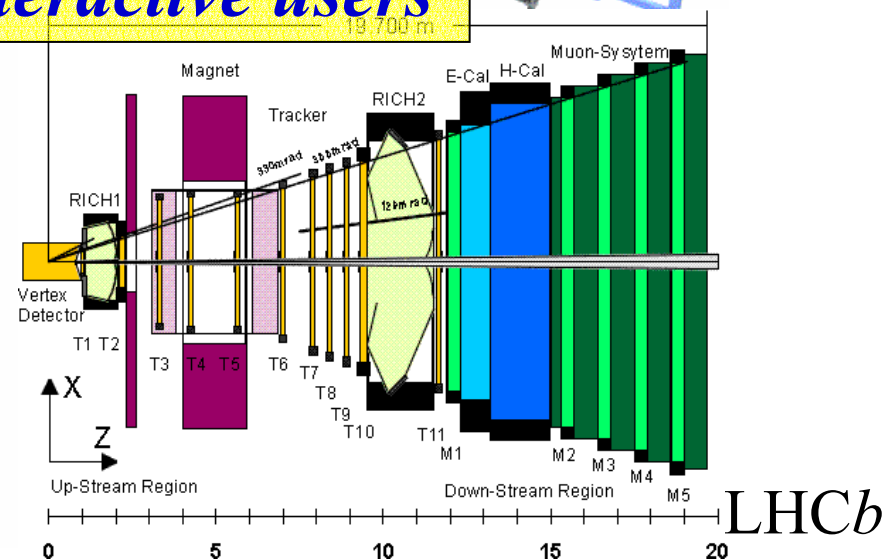
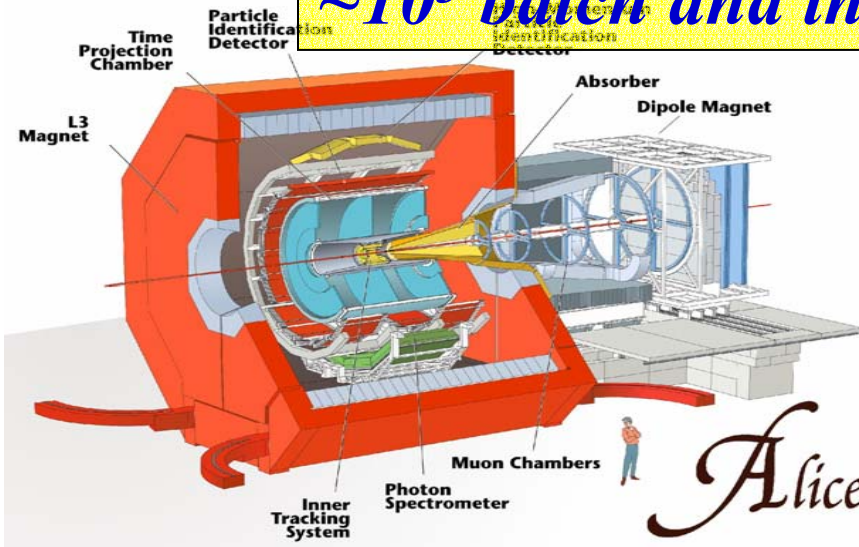
## 4 Experiments on LHC

ATLAS

CMS



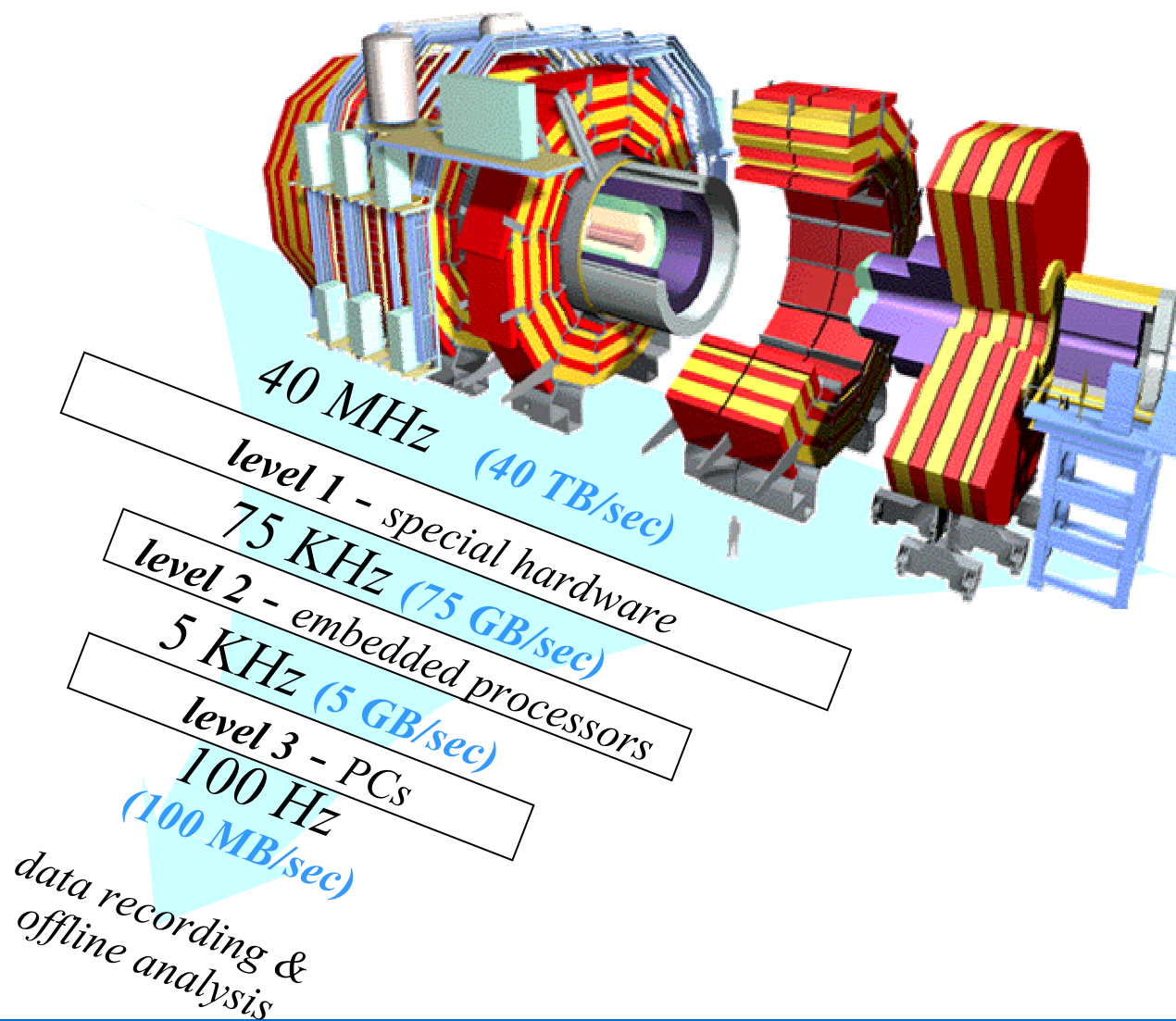
*~6-8 PetaBytes / year*  
*~10<sup>8</sup> events/year*  
*~10<sup>3</sup> batch and interactive users*





## online system

*multi-level trigger  
filter out background  
reduce data volume*



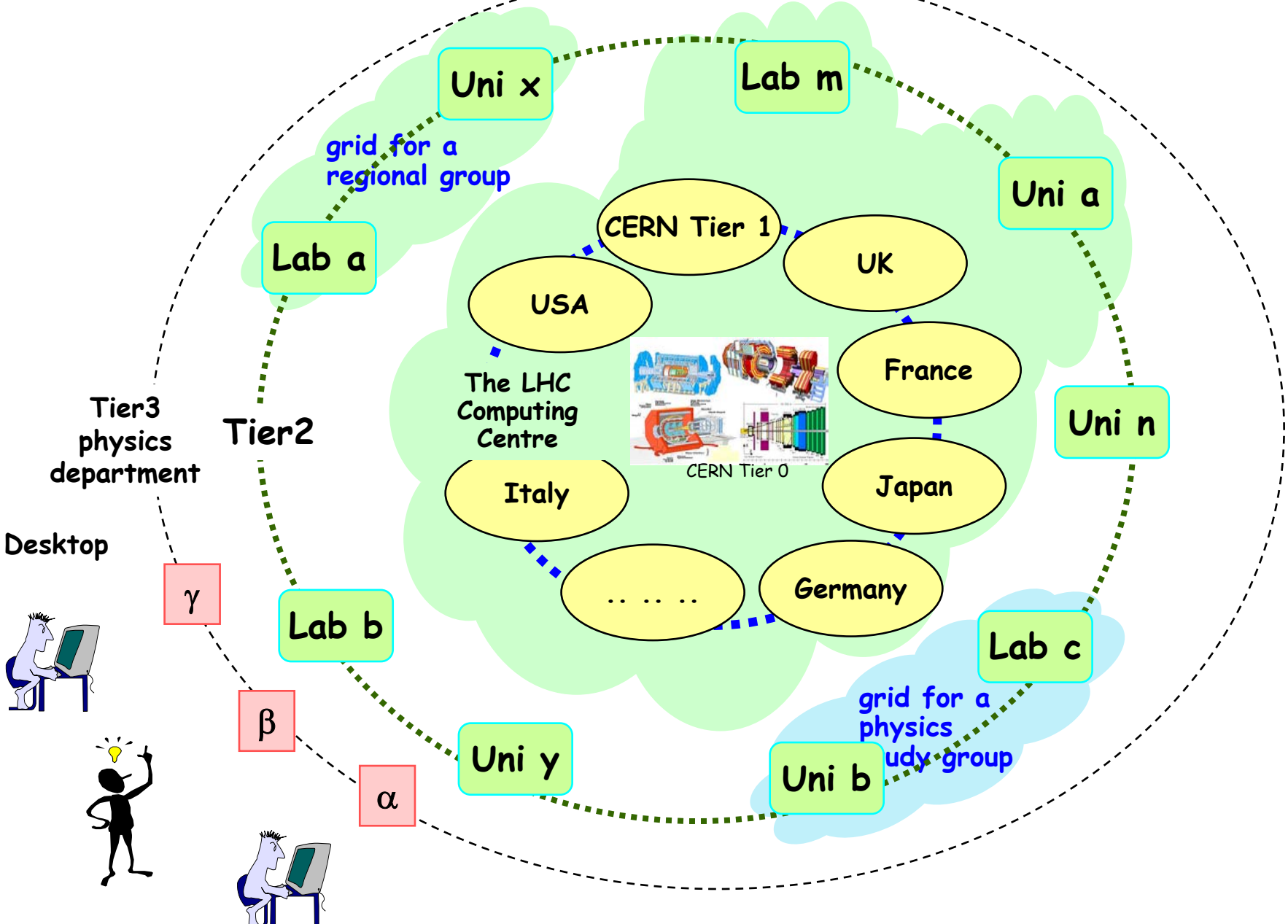
# CERN's Network in the World



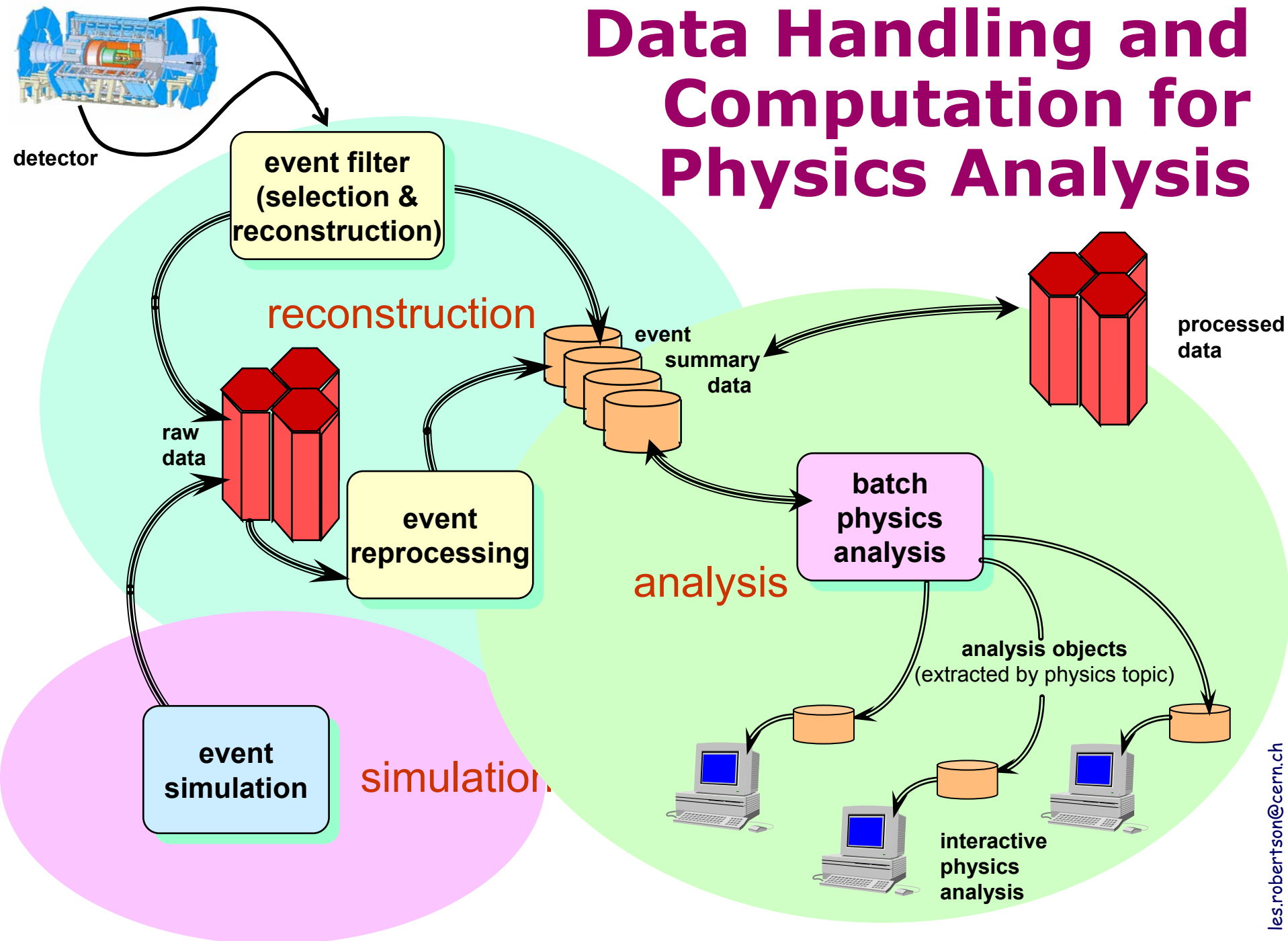
**Europe: 267 institutes, 4603 users**

**Elsewhere: 208 institutes, 1632 users**

# Deploying the LHC Global Grid Service



# Data Handling and Computation for Physics Analysis





## Earth Observation applications

- **Global Ozone (GOME) Satellite Data Processing and Validation by KNMI, IPSL and ESA**
- The **DataGrid testbed** provides a **collaborative processing environment** for 3 geographically distributed **EO** sites (Holland, France, Italy)

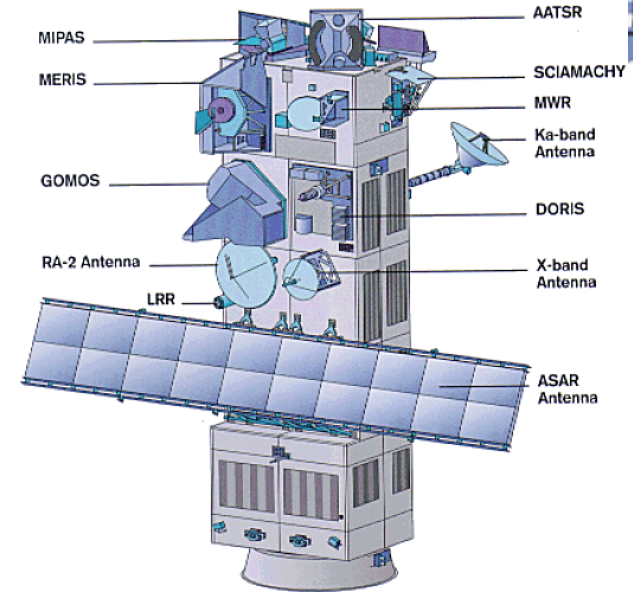


# Earth Observation

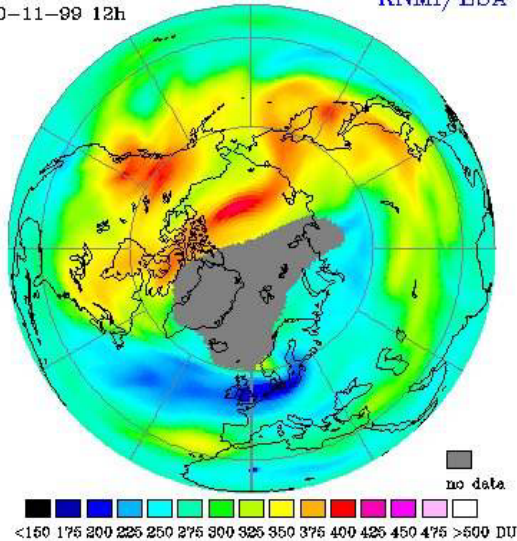


## ESA missions:

- about 100 Gbytes of data per day (ERS 1/2)
- 500 Gbytes, for the next ENVISAT mission (2002).



Assimilated GOME total ozone  
30-11-99 12h  
KNMI/ESA



## DataGrid contribute to EO:

- enhance the ability to access high level products
- allow reprocessing of large historical archives
- improve Earth science complex applications (data fusion, data mining, modelling ...)

Source: L. Fusco, June 2001

The image shows the Earth Observing Satellite (ENVISAT) in orbit above the Earth. The satellite is a complex, gold-colored structure with various instruments and antennas. A long, thin solar panel array extends from the main body of the satellite. The Earth's surface is visible below, showing blue oceans, white clouds, and brown landmasses. The title 'ENVISAT' is written in large, bold, yellow letters on the left side of the image.

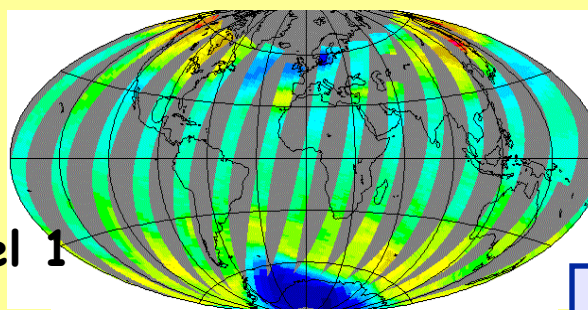
# ENVISAT

- 3500 Meuro programme cost
- Launched on February 28, 2002
- 10 instruments on board
- 200 Mbps data rate to ground
- 400 Tbytes data archived/year
- ~100 'standard' products
- 10+ dedicated facilities in Europe
- ~700 approved science user projects

# EDG EO challenge: Processing / validation of 1y of GOME data



Raw satellite data from the GOME instrument (~75 GB - ~5000 orbits/y)

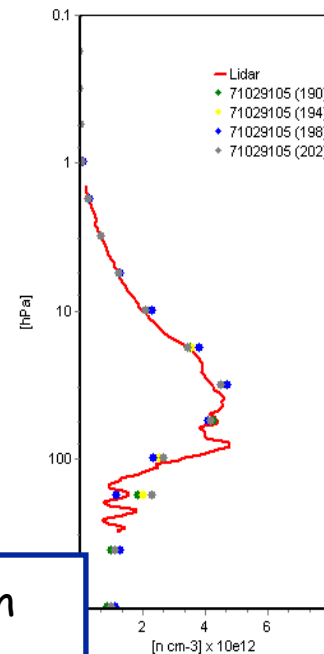


(example of 1 day total O<sub>3</sub>)

LIDAR data (7 stations, 2.5MB per month)

**ESA(IT) KNMI(NL)**  
Processing of raw GOME data to ozone profiles.  
2 alternative algorithms  
~28000 profiles/day

**IPSL(FR)**  
Validate some of the GOME ozone profiles (~10<sup>6</sup>/y)  
Coincident in space and time with Ground-Based measurements



Level 2

**DataGrid environment**

Visualization & Analyze

# EO WebMap Portal



ESA - GRID on demand - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Refresh Home Search Favorites History Mail Print Edit Discuss AIM

Address http://giserver.esrin.esa.int/grid-demo/ Go Links >>

**GRID on-Demand**  
Ozone Application

European Space Agency

Work Package 9

Start Date (1997):  -

End Date (1997):  -

Dataset:

ESA Catalogue Server:

Select GRID:

[acknowledgments]

ESA Storage(AMS)
LIDAR stations
GRID Storage
Processed Files
Web Mapping

1

- 71003124.LV1
- 71001121.LV1
- 71002114.LV1
- 71003111.LV1
- 71001102.LV1

2

- 71003124.LV1
- 71001121.LV1
- 71002114.LV1
- 71003111.LV1
- 71003092.LV1

3

- 71003124
- 71001121
- 71002114
- 71003111
- 71003092

4

- 71003124
- 71002114
- 71003092
- 71003111

ATSR/2 (Orbit 1283...)

FileName = 71003111.LV1

Orbit = 12831

StartDate = 1997-10-03

StartTime = 11:09:59.99

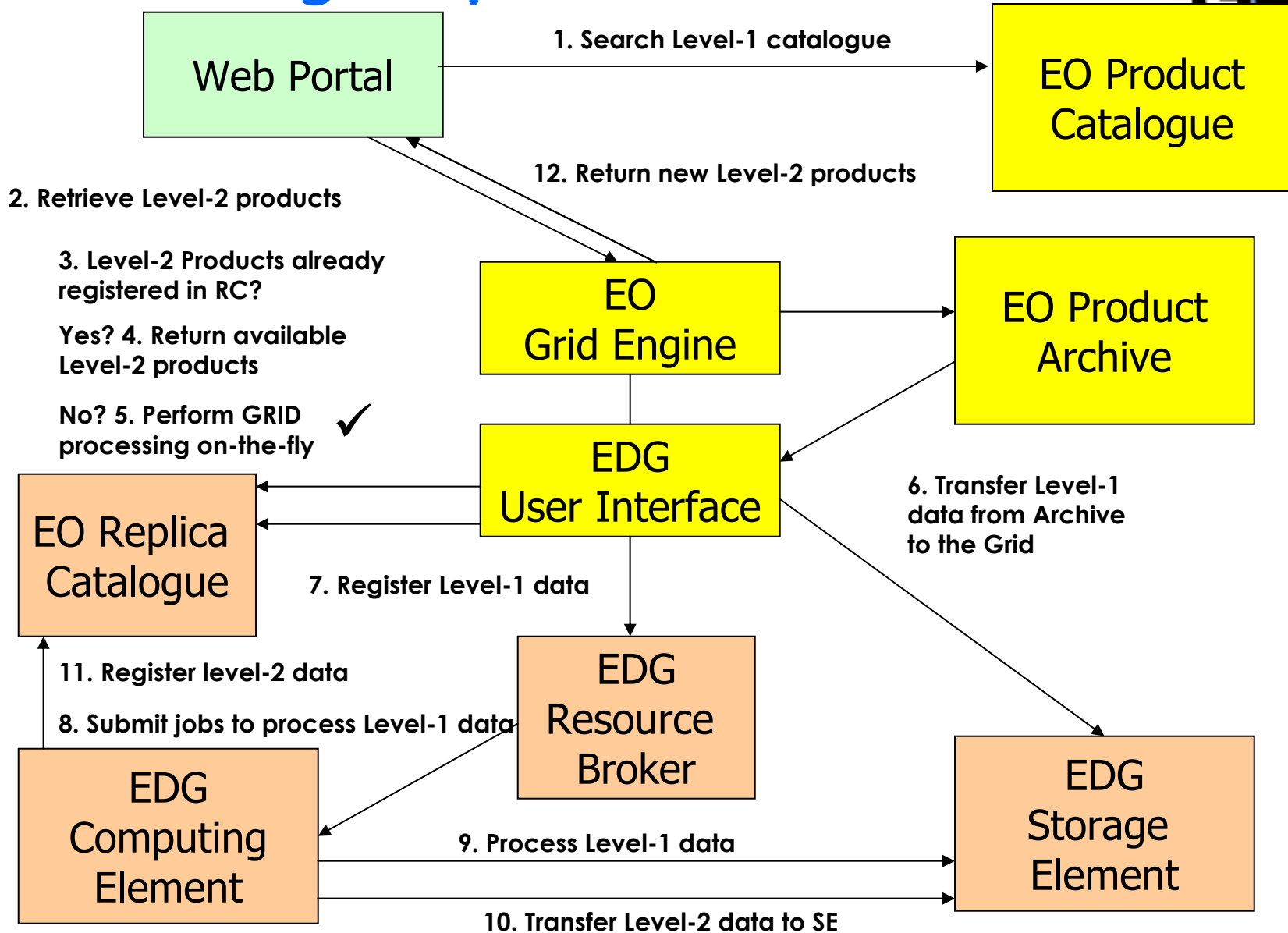
EndDate = 1997-10-03

EndTime = 12:00:17.99

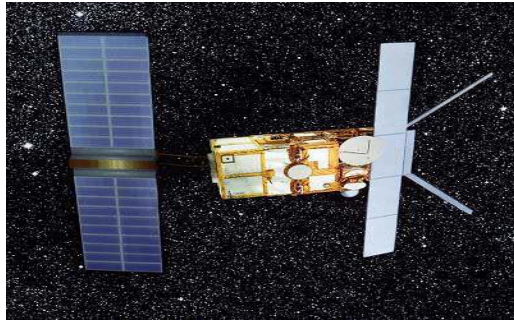
71002114.LV1,12817,1997-10-02,11:39:59.99,1997-10-02,12:30:17.99

Internet

# Processing Sequence



# GOME Ozone Profile Validation



ERS/GOME satellite

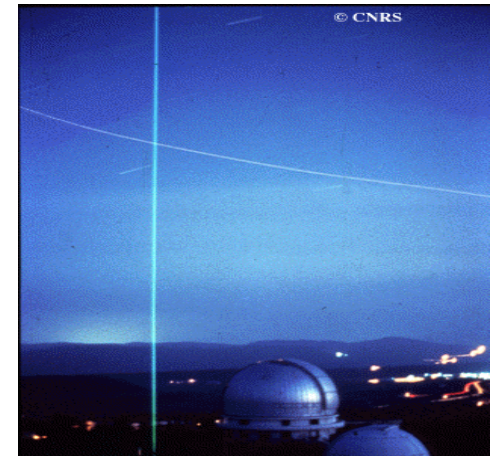
50 km



10 km

OZONE LAYER

- ✓ Goals of the DataGrid application
  - validate satellite data with all ground based data available in an easy way:
    - Comparison of ozone profiles provided by satellite with lidar data in different locations and times (see the web portal)
    - Statistical comparison and analysis in order to improve algorithms.

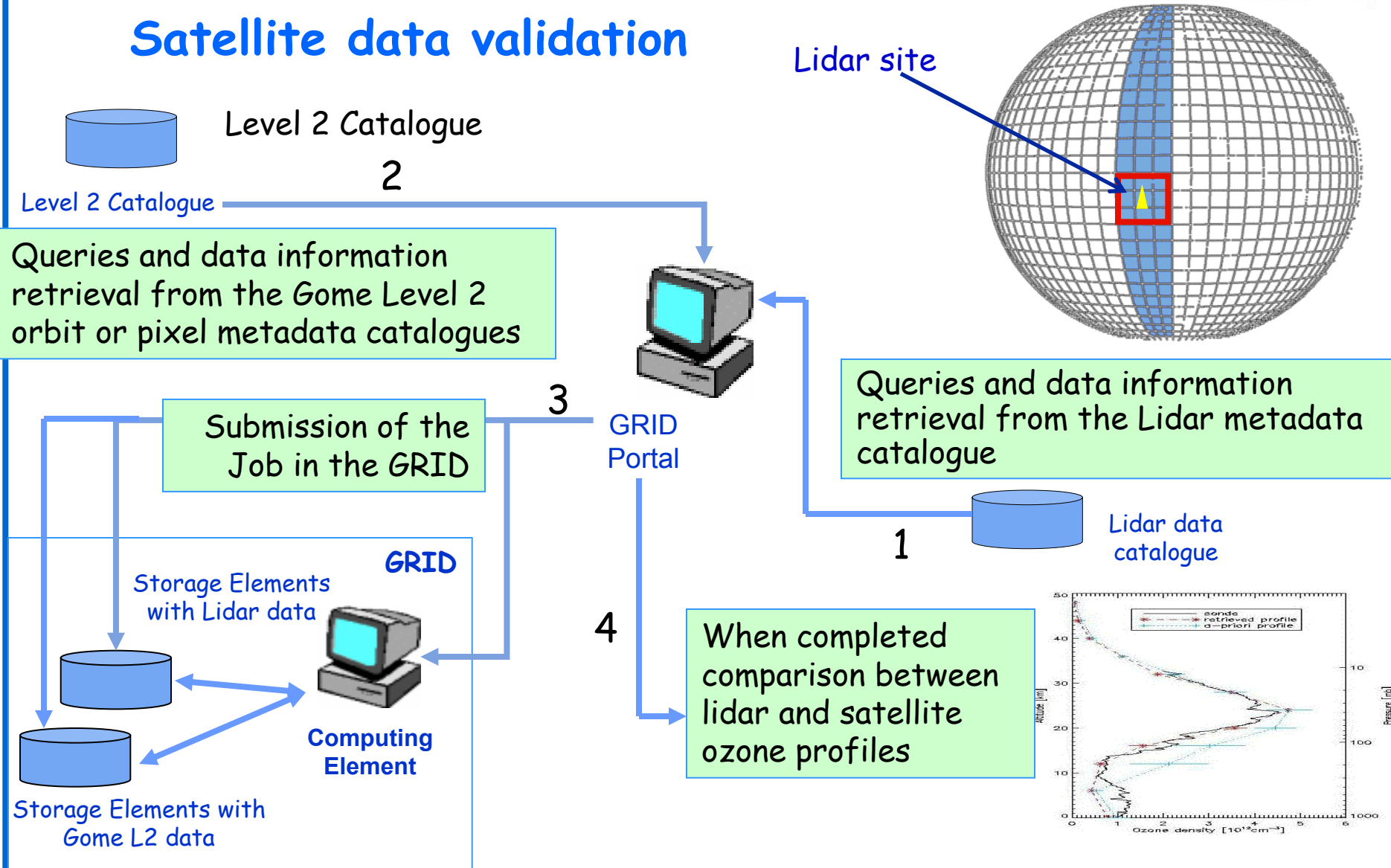


Lidar at the Haute Provence Observatory

# Validation Processing Sequence



## Satellite data validation



# Validation Output

Figure 1:

Estimation of the bias between Gome and Lidar using one month of data.

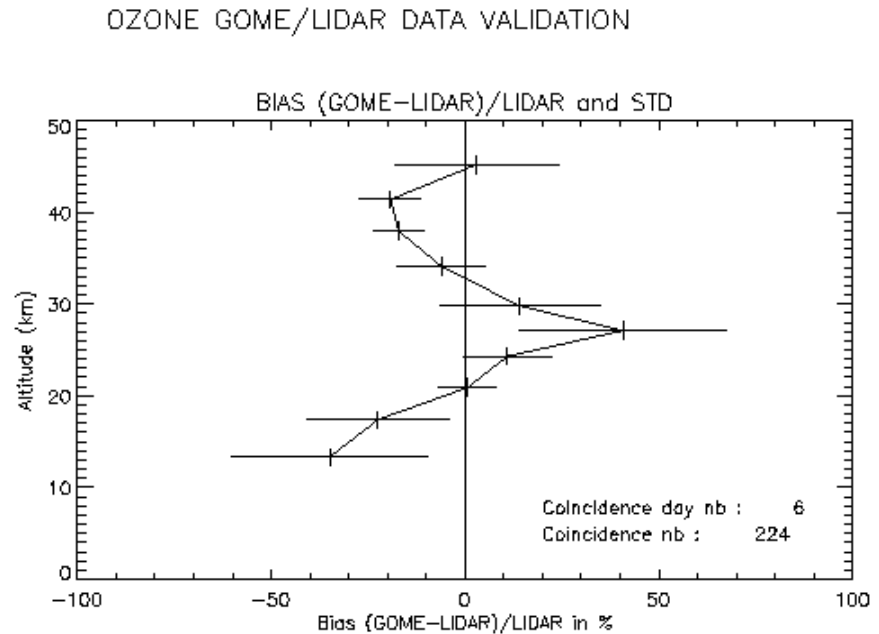
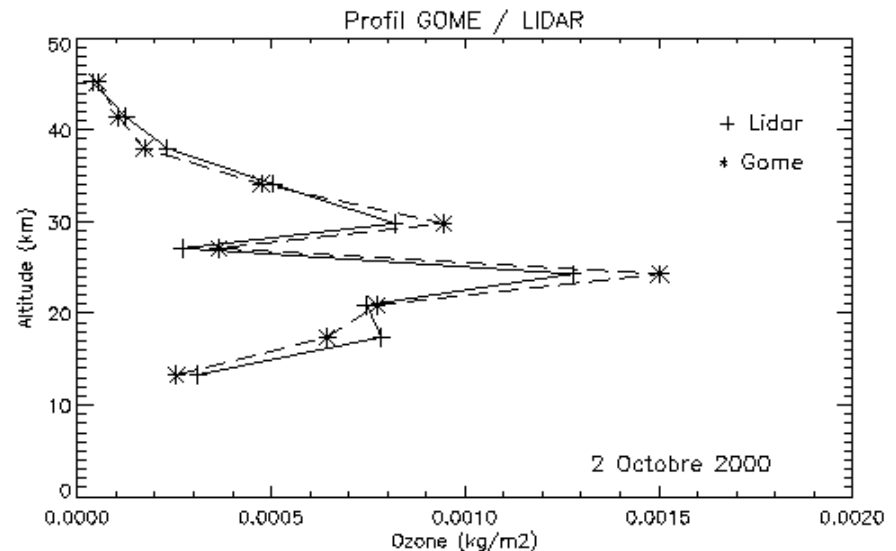


Figure 2 :

example of 2 profiles :  
Comparison between Gome profile and lidar profile for the 2nd October 2000.

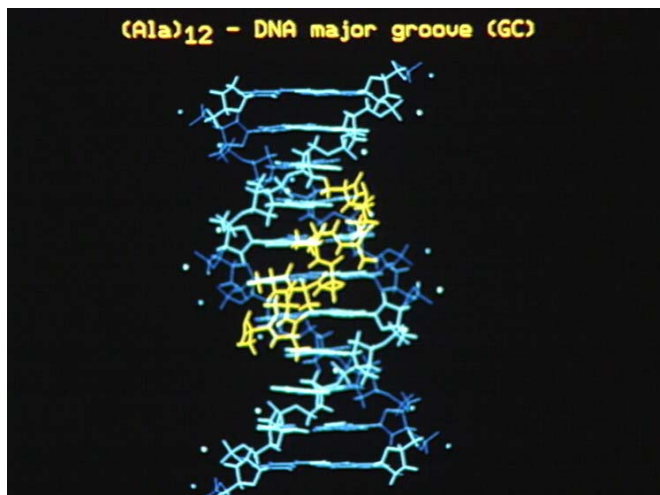




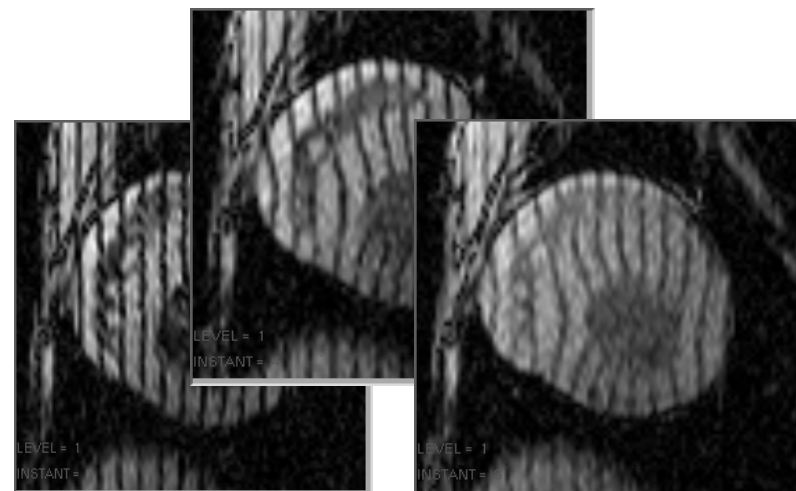
# Biomedical Applications



Genomics, post-genomics,  
and proteomics



Medical images  
analysis






Explore strategies that facilitate the sharing of genomic databases and test grid-aware algorithms for comparative genomics

Process the huge amount of data produced by digital imagers in hospitals.

# Biomedical Applications

## ◆ Bio-informatics

- **Phylogenetics : BBE Lyon (T. Sylvestre)**
- **Search for primers : Centrale Paris (K. Kurata)**
- **Statistical genetics : CNG Evry (N. Margetic)**
- **Bio-informatics web portal : IBCP (C. Blanchet)**
- **Parasitology : LBP Clermont, Univ B. Pascal (N. Jacq)**
- **Data-mining on DNA chips : Karolinska (R. Médina, R. Martinez)**
- **Geometrical protein comparison : Univ. Padova (C. Ferrari)**

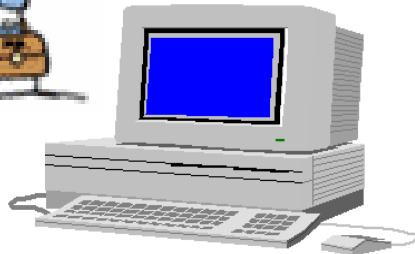
- |   |                                |
|---|--------------------------------|
|  | Applications deployed          |
|  | Applications tested on EDG     |
|  | Applications under preparation |

## ◆ Medical imaging

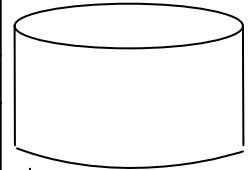
- **MR image simulation : CREATIS (H. Benoit-Cattin)**
- **Medical data and metadata management : CREATIS (J. Montagnat)**
- **Mammographies analysis ERIC/Lyon 2 (S. Miguët, T. Tweed)**
- **Simulation platform for PET/SPECT based on Geant4 : GATE collaboration (L. Maigne)**

# Medical Imaging

H



LFN	image	patient	hospital	...



Metadata

Medical images

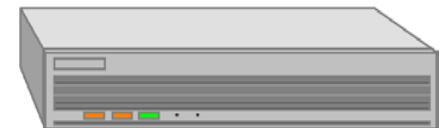
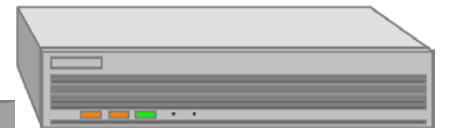
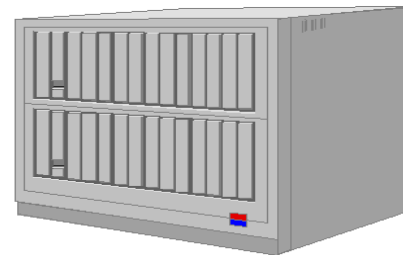
1. query

2. visualisation

5. best results visualisation

3. similarity search

4. scores



# Graphic layer



Job	Status	Target
27499 (similarity)	Terminated	localhost:0/noqueue
27503 (similarity)	Terminated	localhost:0/noqueue
27507 (similarity)	Terminated	localhost:0/noqueue
27511 (similarity)	Terminated	localhost:0/noqueue
27515 (similarity)	Terminated	localhost:0/noqueue
27520 (similarity)	Terminated	localhost:0/noqueue
27524 (similarity)	Terminated	localhost:0/noqueue
27528 (similarity)	Terminated	localhost:0/noqueue
27532 (similarity)	Terminated	localhost:0/noqueue
27536 (similarity)	Terminated	localhost:0/noqueue
27540 (similarity)	Terminated	localhost:0/noqueue
27544 (similarity)	Terminated	localhost:0/noqueue
27548 (similarity)	Terminated	localhost:0/noqueue
27552 (similarity)	Terminated	localhost:0/noqueue
27556 (similarity)	Terminated	localhost:0/noqueue
27560 (similarity)	Output ready	localhost:0/noqueue
27564 (similarity)	Running	
27568 (similarity)	Submitted	
27572 (similarity)	Submitted	
New similarity	Sending to UI	

Job Monitoring

Path: /medical/thorax

..	00000.inr
	00001.inr
	00002.inr
	00003.inr
	00004.inr
	00005.inr

Create Remove Delete View

Target SE: No default SE selected  
localhost

Grid File Browsing

Nimice

Path: /home/johan/RM/medical/heart

..	002.inr
	003.inr
	004.inr
	005.inr
	006.inr
	007.inr
	008.inr
	009.inr
	010.inr
	011.inr
	012.inr
	013.inr
	014.inr
	015.inr

Path: /medical

..	brain
	heart
	thorax

Target SE: gppse05.gridpp.rl.ac.uk  
grid005.pd.infn.it  
grid007g.cnaf.infn.it

Cancel

File registration and retrieval

## Image registration

The screenshot shows three main components of the image registration interface:

- Local files:** A file list on the left showing files named 002.inr through 018.inr. File 010.inr is selected.
- Grid files:** A file list on the right showing 'brain', 'heart', and 'thorax'. 'heart' is selected. Below it, a 'Target SE' dropdown menu shows options like 'gppse05.gridpp.rl.ac.uk', 'grid005.pd.infn.it', and 'grid007g.cnaf.infn.it'.
- Metadata:** A form on the right with fields for:
  - Source file: /home/johan/RM/medical/heart/237.inr
  - Destination: grid005.pd.infn.it/medical/237.inr
  - Type: 8 bits unsigned, Vectorial dim: 1
  - Size: 256 x 256 x 1 x 1
  - Voxels Size: 1.000 x 1.000 x 1.000 x 1.000
  - Patient name: Dupond, Françoise
  - Sexe: Female, Birth date: 21/03/1964
  - Hospital: Lyon Cardiology Hospital, Radiologist: Dr André Dussolie
  - Acquisition date: 16/10/1999
  - Modality: MRI, Region: Heart
  - Orientation: (empty dropdown)
  - Diagnosis: (empty text box)

Local files

Grid files

Metadata

## Image retrieval

The screenshot shows a search form for image retrieval with the following fields:

- Patient name: Dupont
- Sexe: (dropdown menu)
- Birth date: (text input)
- Hospital: (text input)
- Radiologist: (text input)
- Acquisition date: (text input)
- Modality: MRI (dropdown menu)
- Region: (dropdown menu)
- Orientation: (dropdown menu)

Buttons at the bottom include 'Query', 'Select all', and 'Cancel'.

Query over metadata

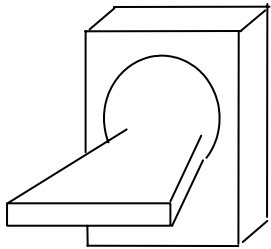
Patient	Medical	Hospital	Diagnosis	Image
Family name	First name	Sexe	Birth date	
Dupont	Peter	M	1944-03-12	
Dupont	Denise	F	1970-12-04	
Dupont	John	M	1966-11-18	
Dupont	Marc	M	1975-12-25	
Dupont	Denise	F	1970-12-04	
Dupont	Denise	F	1970-12-04	
Dupont	Denise	F	1970-12-04	
Dupont	Marc	M	1975-12-25	
Dupont	Marc	M	1975-12-25	
Dupont	Peter	M	1944-03-12	
Dupont	Jean	M	1978-06-02	
Dupont	Sandra	F	1962-12-27	
Dupont	Denise	F	1970-12-04	

Buttons at the bottom include 'View' and 'Dismiss'.

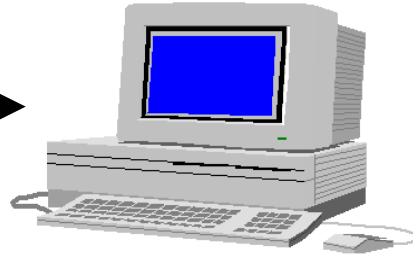
Query result



# Image Registration



Imager

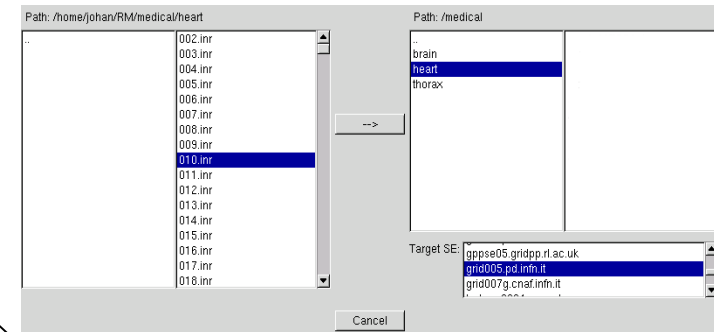


Source file: /home/johan/RM/medical/heart/237.inr  
Destination: grid005.pd.infn.it/medical/237.inr  
Type: 8 bits unsigned, Vectorial dim: 1  
Size: 256 x 256 x 1 x 1  
Voxels Size: 1.000 x 1.000 x 1.000 x 1.000

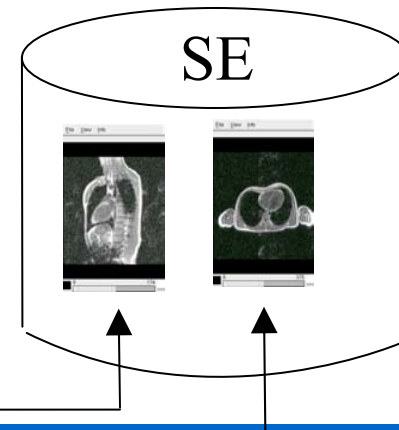
Patient name: Dupond François  
Sex: Female Birth date: 21/03/1964  
Hospital: Lyon Cardiology Hospital Radiologist: Dr André Dussole  
Acquisition date: 16/10/1999

Modality: MRI Region: Heart  
Orientation:  
Diagnosis:

Random Register Cancel



LFN	image	patient	hospital	...



# Similarity search



## Similarity computation

Job	Status	Target
27499 (similarity)	Terminated	localhost:0/noqueue
27503 (similarity)	Terminated	localhost:0/noqueue
27507 (similarity)	Terminated	localhost:0/noqueue
27511 (similarity)	Terminated	localhost:0/noqueue
27515 (similarity)	Terminated	localhost:0/noqueue
27520 (similarity)	Terminated	localhost:0/noqueue
27524 (similarity)	Terminated	localhost:0/noqueue
27528 (similarity)	Terminated	localhost:0/noqueue
27532 (similarity)	Terminated	localhost:0/noqueue
27536 (similarity)	Terminated	localhost:0/noqueue
27540 (similarity)	Terminated	localhost:0/noqueue
27544 (similarity)	Terminated	localhost:0/noqueue
27548 (similarity)	Terminated	localhost:0/noqueue
27552 (similarity)	Terminated	localhost:0/noqueue
27556 (similarity)	Terminated	localhost:0/noqueue
27560 (similarity)	Output ready	localhost:0/noqueue
27564 (similarity)	Running	localhost:0/noqueue
27568 (similarity)	Submitted	localhost:0/noqueue
27572 (similarity)	Submitted	localhost:0/noqueue
New similarity	Sending to UI	

Job monitoring

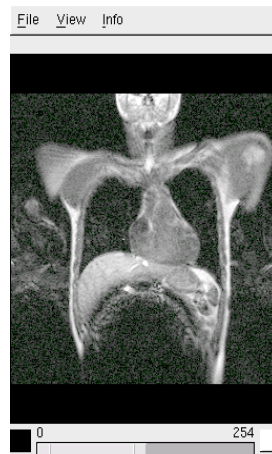
File	Similarity	About
Source image: Jones Jean    Cardiology Center of Monaco    Dr Jina Carlson    1997-11-18		
Results:		
0.904684	Durand Jean	Lyon Cardiology Hospital    Dr Alain Deloin    2002-02-21
0.743148	Dupont Marc	Cardiology Center of Monaco    Dr Francis Black    1998-01-18
0.219428	Durand Jean	Cardiology Center of Monaco    Dr Jina Carlson    2000-10-08
0.217490	Jones Linda	Montreal Neurological Institut    Dr Fany Anderson    2000-12-21
0.193847	Jones Sandra	Cardiology Center of Monaco    Dr Francis Black    2000-12-25
0.003237	Dupont Denise	Montreal Neurological Institut    Dr Norbert White    1998-10-22
0.003084	Dupont John	Montreal Neurological Institut    Dr Norbert White    1998-04-22
0.002636	Smith Marc	Cardiology Center of Monaco    Dr Jina Carlson    1997-04-04
0.001778	Durand Sylvie	Lyon Neurology Hospital    Dr Martine Follet    2001-02-14
0.001515	Smith Marc	Montreal Neurological Institut    Dr Norbert White    2001-02-09
0.001023	Durand Jean	Cardiology Center of Monaco    Dr Jina Carlson    2000-02-24

Ranked list of images

## Results visualization



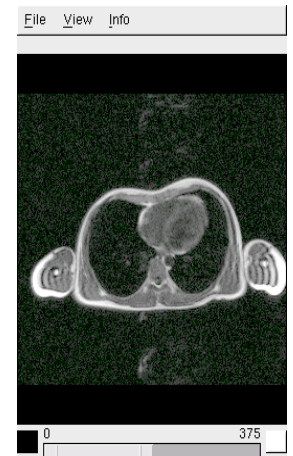
Source image



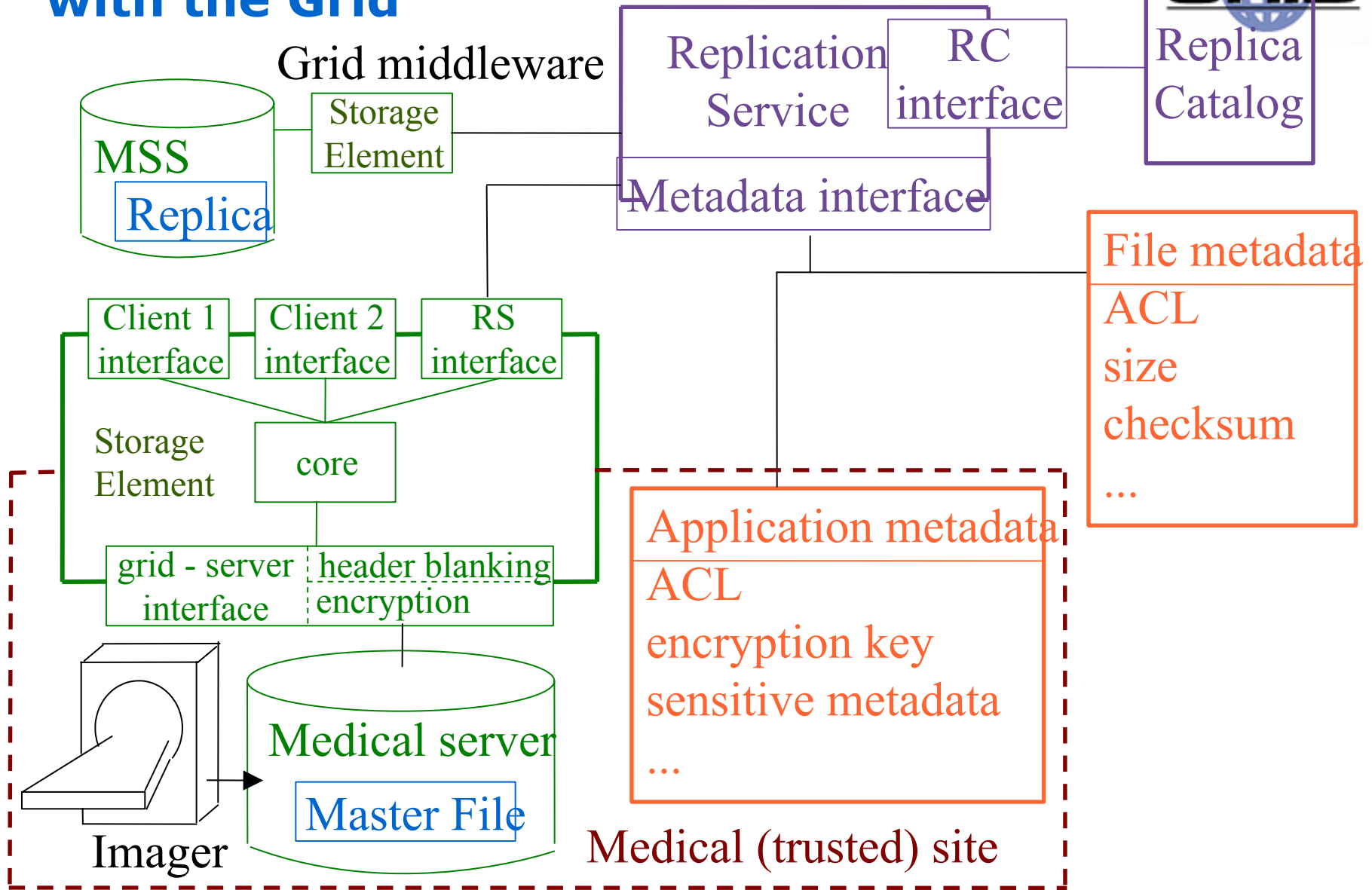
Most similar images



Low score images



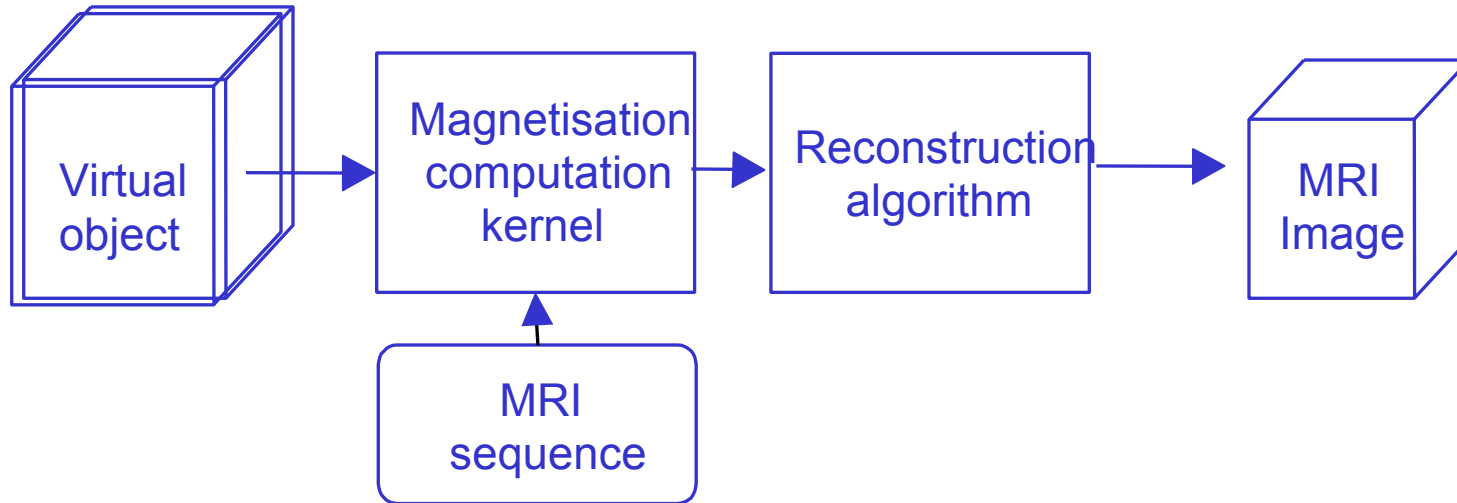
# Future: Interfacing medical data with the Grid





# Parallel Processing

- Magnetic Resonance Images simulation using the grid



- 3 levels of parallelism:
  - Parallel isochromat computations
  - Multi-slice MRI computation
  - Parallel magnetization kernel

# Summary



- ◆ Use Cases
  - High Energy Physics
  - Earth Observation
  - Biomedical Applications

# Further Information



- ◆ High Energy Physics

<http://datagrid-wp8.web.cern.ch/DataGrid-WP8/>

- ◆ Bio-Informatics

<http://marianne.in2p3.fr/datagrid/wp10/index.html>

- ◆ Earth Observation

<http://styx.esrin.esa.it/grid/>