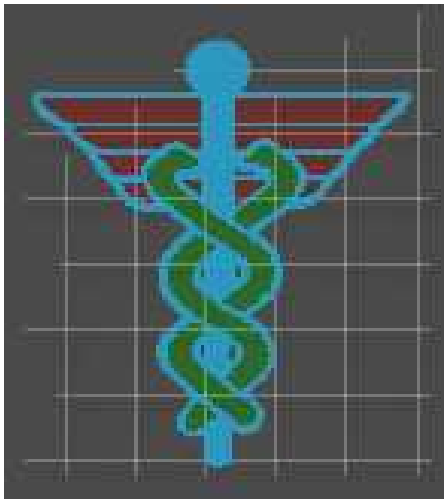
MEDIGRID

ACI-GRID project

French ministry of research



Medical image processing on grids

http://www.creatis.insa-lyon.fr/MEDIGRID

Johan Montagnat

Partners

French National Center for Scientific Research

CREATIS

Signal & image processing

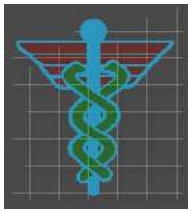
Radiology department

Communicating IS

Information system

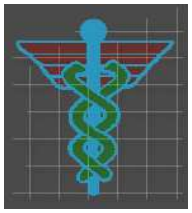
ERIC

Image processing



Medical applications on GRIDs

- Medical applications have specific requirements for grid computing:
  - Data:
    - Are heterogeneous
    - Have a strong semantic
    - Are distributed over medical sites
    - Are confidential (security issues)
  - Processings
    - Are often correlated (pipelines of processings)
    - Computation time is often important (physicians will accept to wait for minutes at most)
    - Computation time is sometimes critical (e.g. real time simulation)
    - Emergency situation: ambulance jobs
- Existing grids are not taking into account all these requirements today



MEDIGRID Objectives

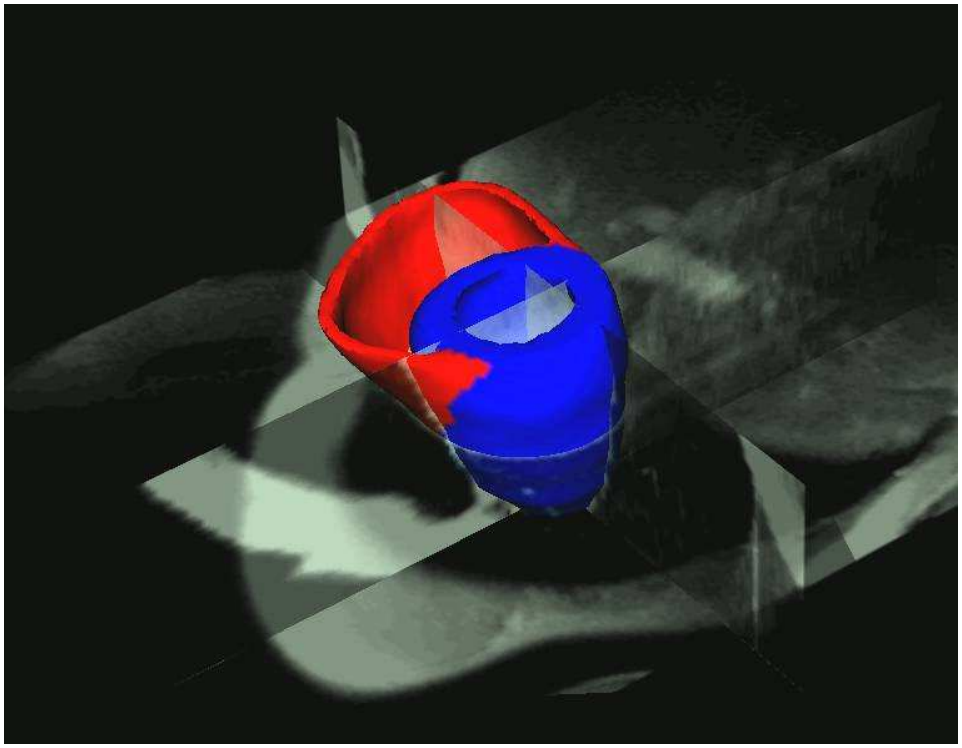
- Use computation GRIDs to face recent challenges in medical data analysis. We are focusing on two application kinds:
  - Computation intensive image processing algorithms
    - Parallelization
    - Reduced computation time
  - Management of very large datasets
    - Distributed storage
    - Massive distributed processing
    - Statistical analysis

1. Complex modeling of anatomical structures

- Anatomical modeling for:
  - Segmentation
  - Quantitative analysis
- Linear Finite Element

  Modeling of biomechanics
  - Parallelization of large

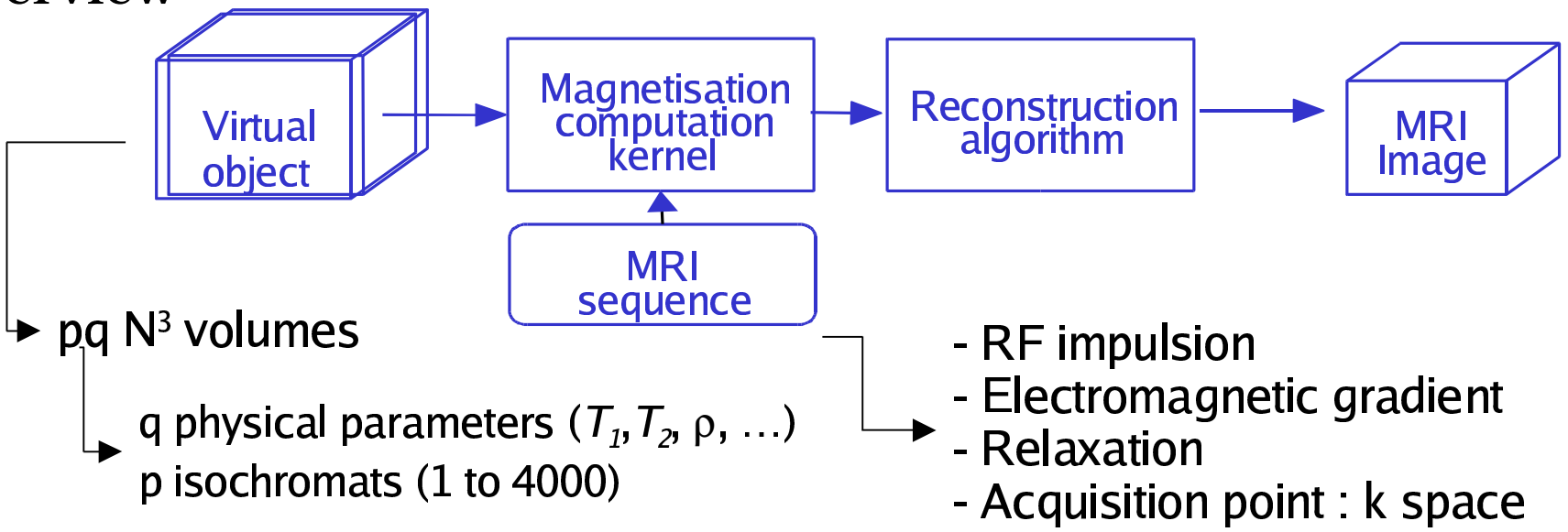
    linear systems



- Modeling / segmentation of 3D+T cardiac sequences in a reasonnably short amount of time (few minutes)

2. Simulation of MRIs

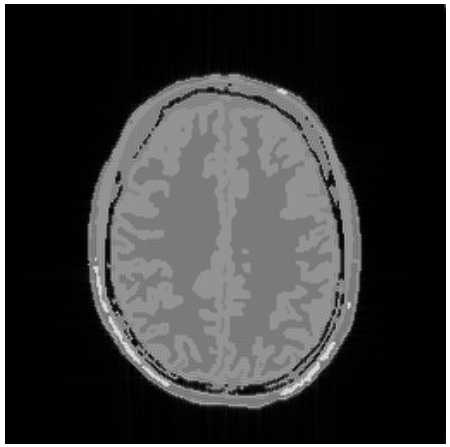
- Produce simulated images from a perfectly known model for:
  - Artifacts study and correction
  - Image processing evaluation
  - MRI sequences testing and design
- Overview

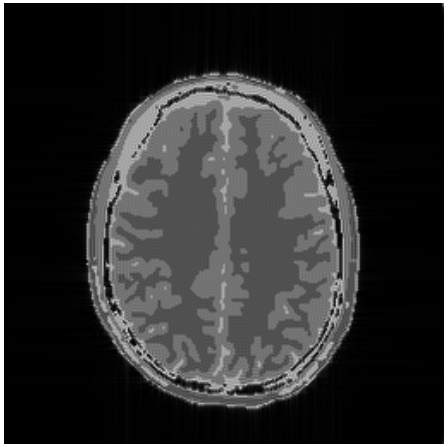


First results

- Synthetized images

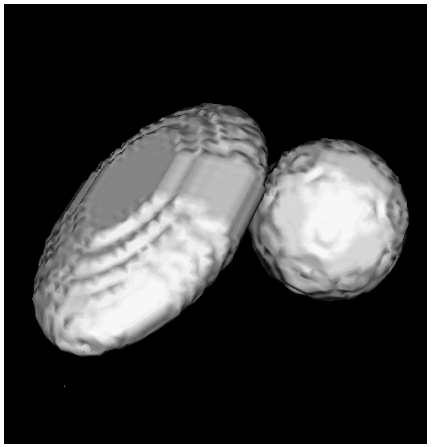
2D (256^2) brain MRI





T1 contrast T2 contrast

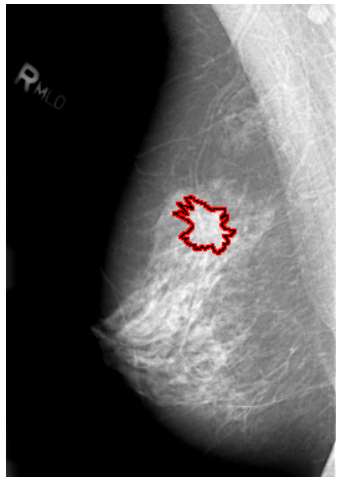
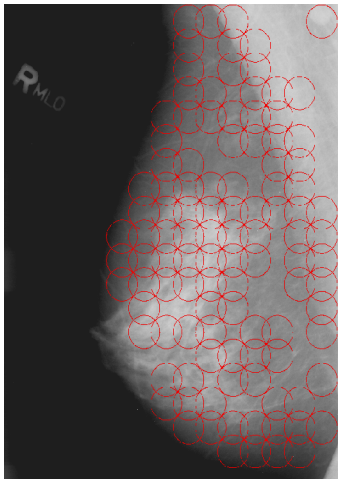
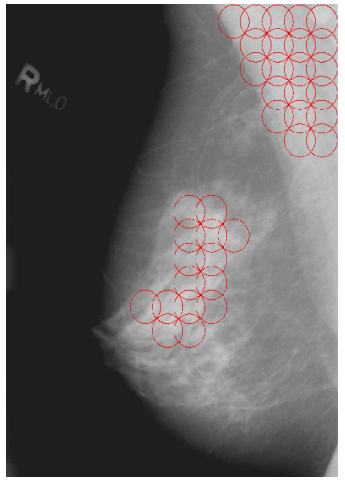
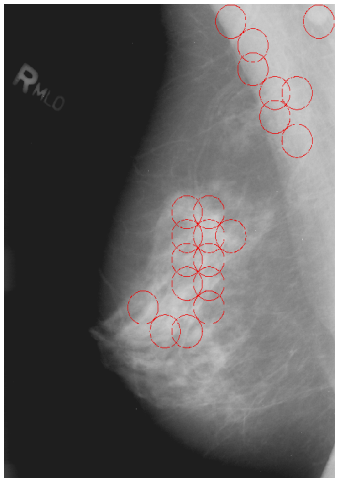
3D (64^3)



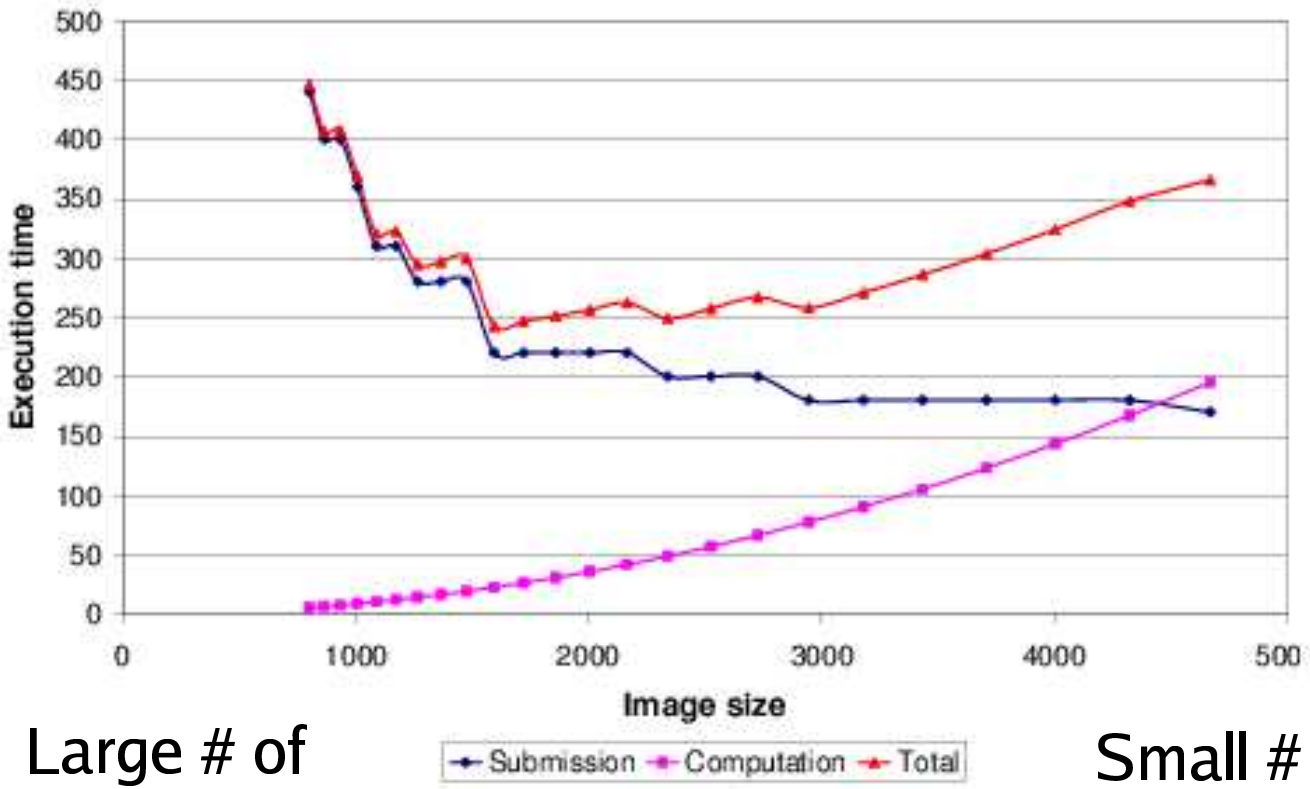
- Computation time
  - 2D: small cluster (1024^2 = 2.3 days)
  - 3D: full scale grid (128^3 = 2.3 days, 512^3 = 104 years)

3. Mammographies analysis

- More than 10000 images, 450 Gbytes
- 400 sub regions (e.g.) per image
- About 250 variables extracted on each region for training and for CBIR
  - Texture, gray-levels and shape analysis
  - Image indexation
- Indexing requires about 30 minutes of computations per image (Sun Ultra-10, 440 MHz), no optimization

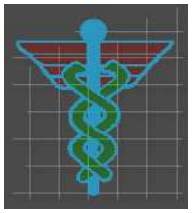





First results



Large # of small jobs

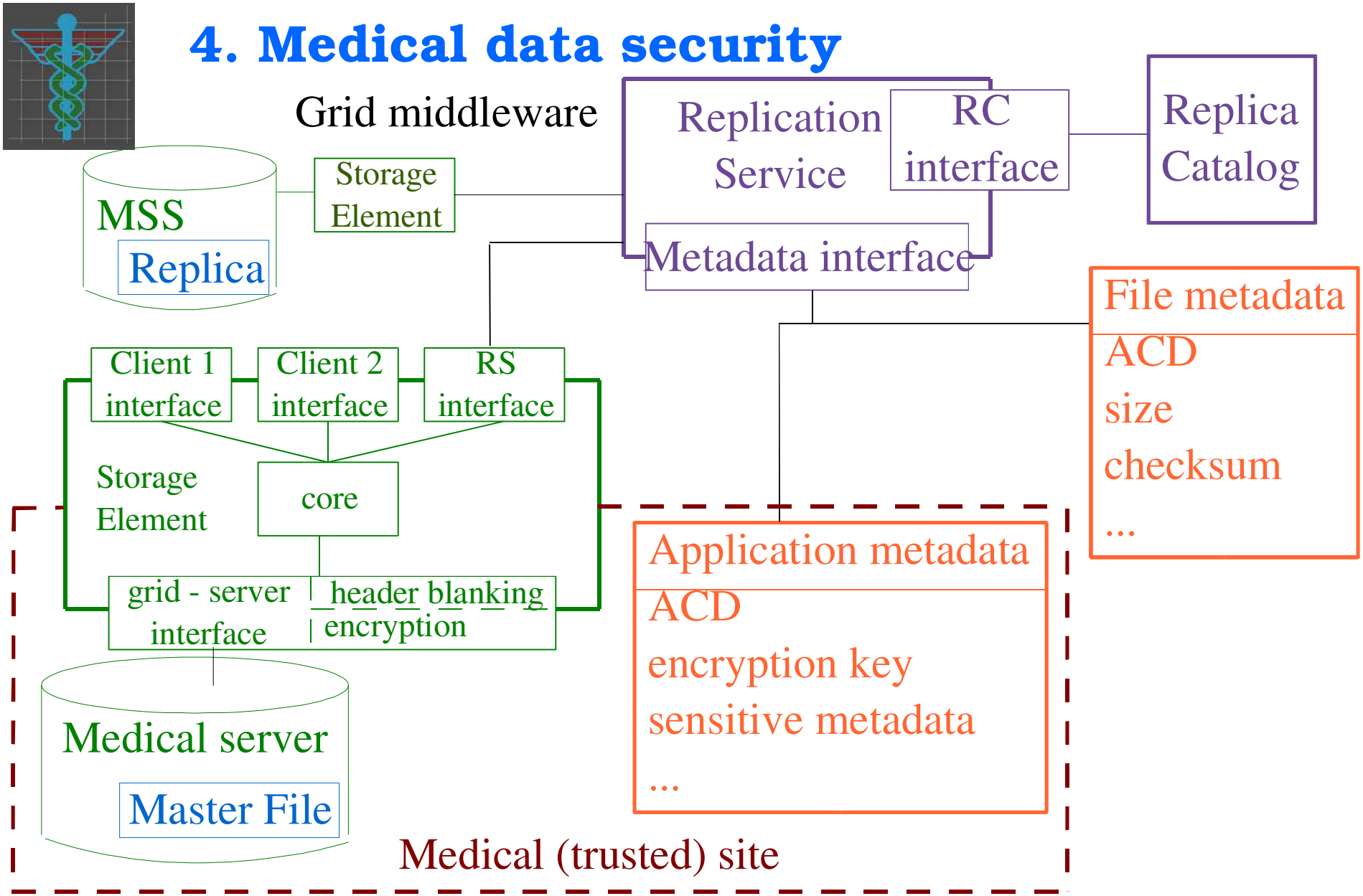
Small # of large jobs



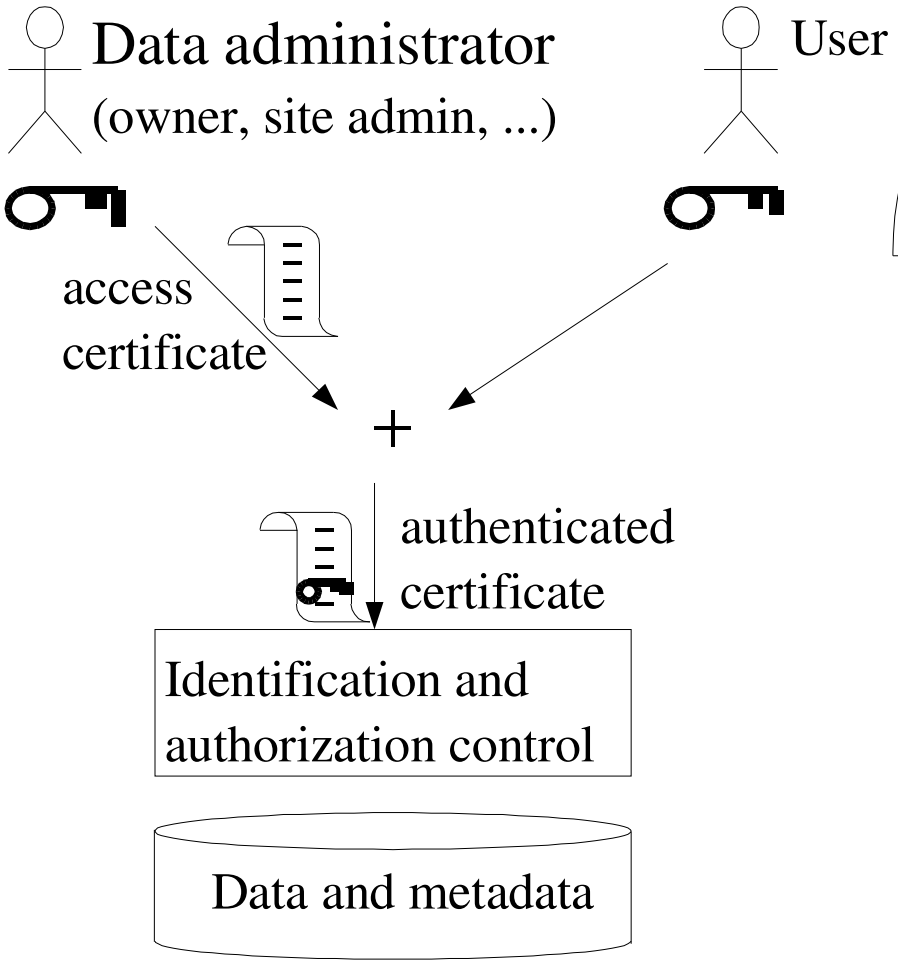
4. Shared and distributed data management

- Distributed data and distributed metadata
  - Metadata Distribution/Location Service (similar to GRID replication services for metadata)
  - Metadata and data should be synchronized (same lifetime, access authorization…)
  - Data traceability (How was data B produced? Which result was obtained from data A?)
- High level layer
  - Intelligent proxy hierarchy
  - Distributed dynamic indices for queries
  - Optimisation / caching of search requests

4. Medical data security

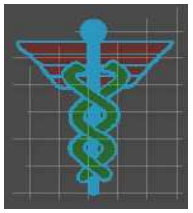


Fine grain data control



Access certificate

User ID
Admin ID
Target data ID
Authorized access modes
Certificate validity
Admin signature



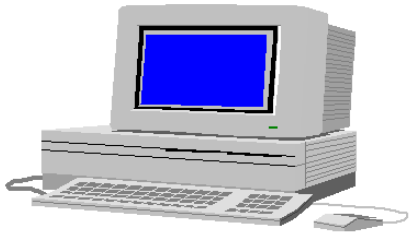
5. Hybrid (content-based and metadata) queries

- Content-based queries
  - Queries images over their content (medical images indexing and research)
  - Mixed content-based and metadata-based queries

- Job submission / data / metadata synchronisation
  - Use queries over metadata to describe input datasets for jobs
    - A job should be able to process a set of files (data)
    - A job should be able to process a set of files corresponding to some metadata (query + processing)

First experiment on the EDG testbed

1. Query the medical image database and retrieve a patient image



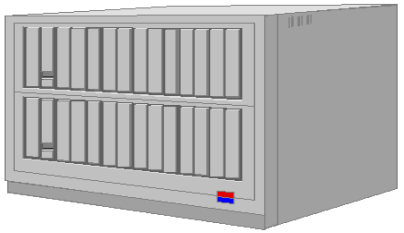
Exam image patient key ACL ...

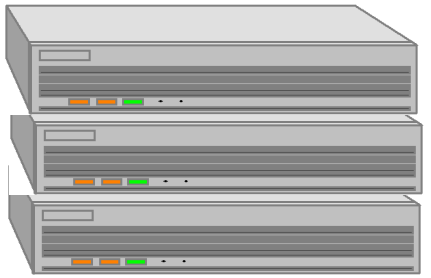
Metadata

Medical images

2. Compute similarity measures over the database images

Submit 1 job per image

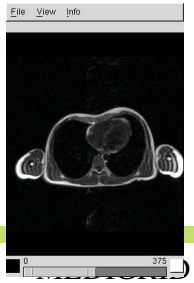


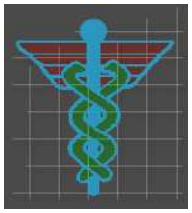


3. Retrieve most similar cases

Similar images

Low score images



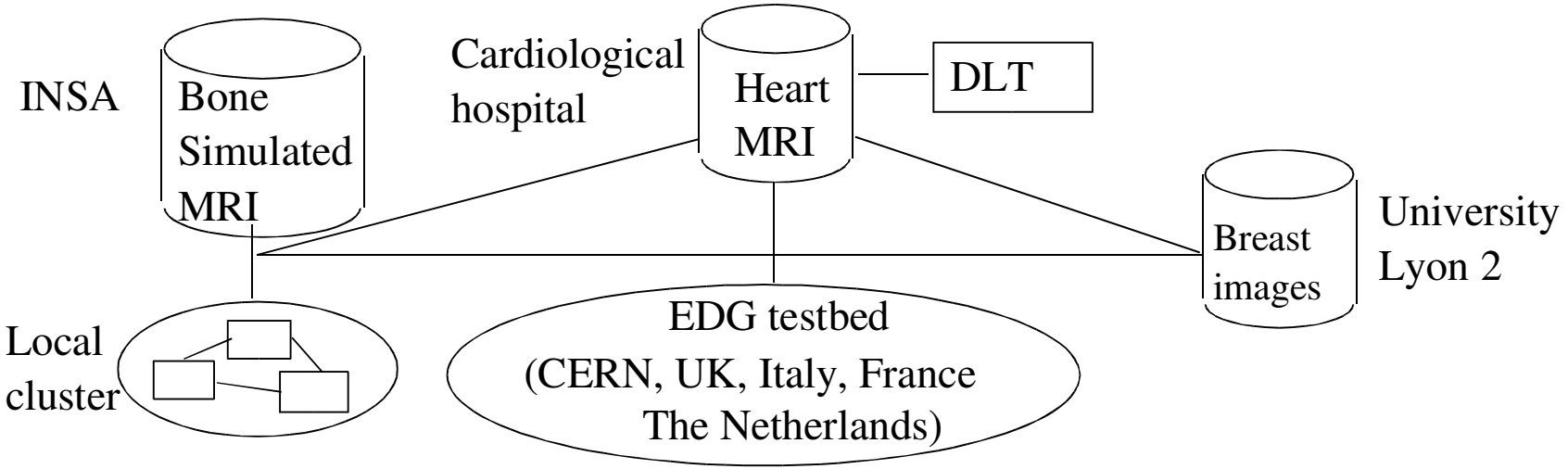


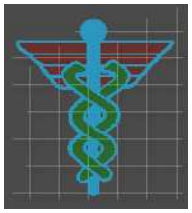
Data and users

- Medical Data
  - Images and metadata
  - Nominative (critical) and non-nominative (private) data
  - DICOM3 standard compliance for medical images
- Users
  - Patient: has free access to its medical data.
  - Physician: has complete read access to his/her patients data. Few persons have read/write access.
  - Researchers: may obtain read access to anonymous medical data for research purposes. Nominative data should be blanked before transmission to these users.

Testbed

- 4 image sources, 3 sites:
  - Heart sequences acquired at the Lyon cardiological hospital
  - Bone structure database from ESRF Grenoble
  - Mammographies from the DDSM
  - Simulated MRI images





Conclusions

- Significant growth of the grid awarness in the medical imaging community
  - Healthgrid'04, Clermont-Ferrand, France, January 2004

    `http://clermont04.healthgrid.org`
  - EU projects (DataGrid, CrossGrid...), e-Science, BIRN...
- Limitations of existing middlewares for biomedical applications
  - Complex datasets management
  - Security
  - Interactivity
  - ...