Increased Quality

GRID COMPUTING

Decreased Time to Market

Group3 Results

Use Case: MS Analysis

**Huge number of data**

```
 ...
700.003      9.000
700.015      3.000
700.026      2.000
700.038      3.000
700.050      1.000
700.062      2.000
700.073      2.000
700.085      4.000
700.097      3.000
700.109      2.000
700.121      5.000
700.132      8.000
700.144      7.000
700.156      12.000
```

Mass/Charge (MALDI-TOF)

Mass Spectra Analysis, from serum samples, produces many mass spectra.

From the Analysis, we can obtain desirable *Bio-Markers* for early detection of Cancer.
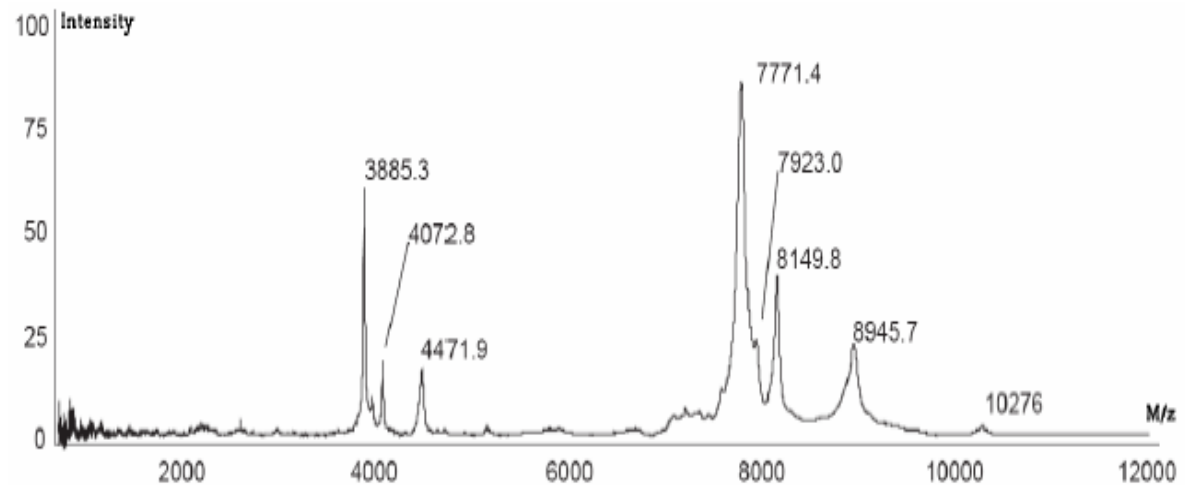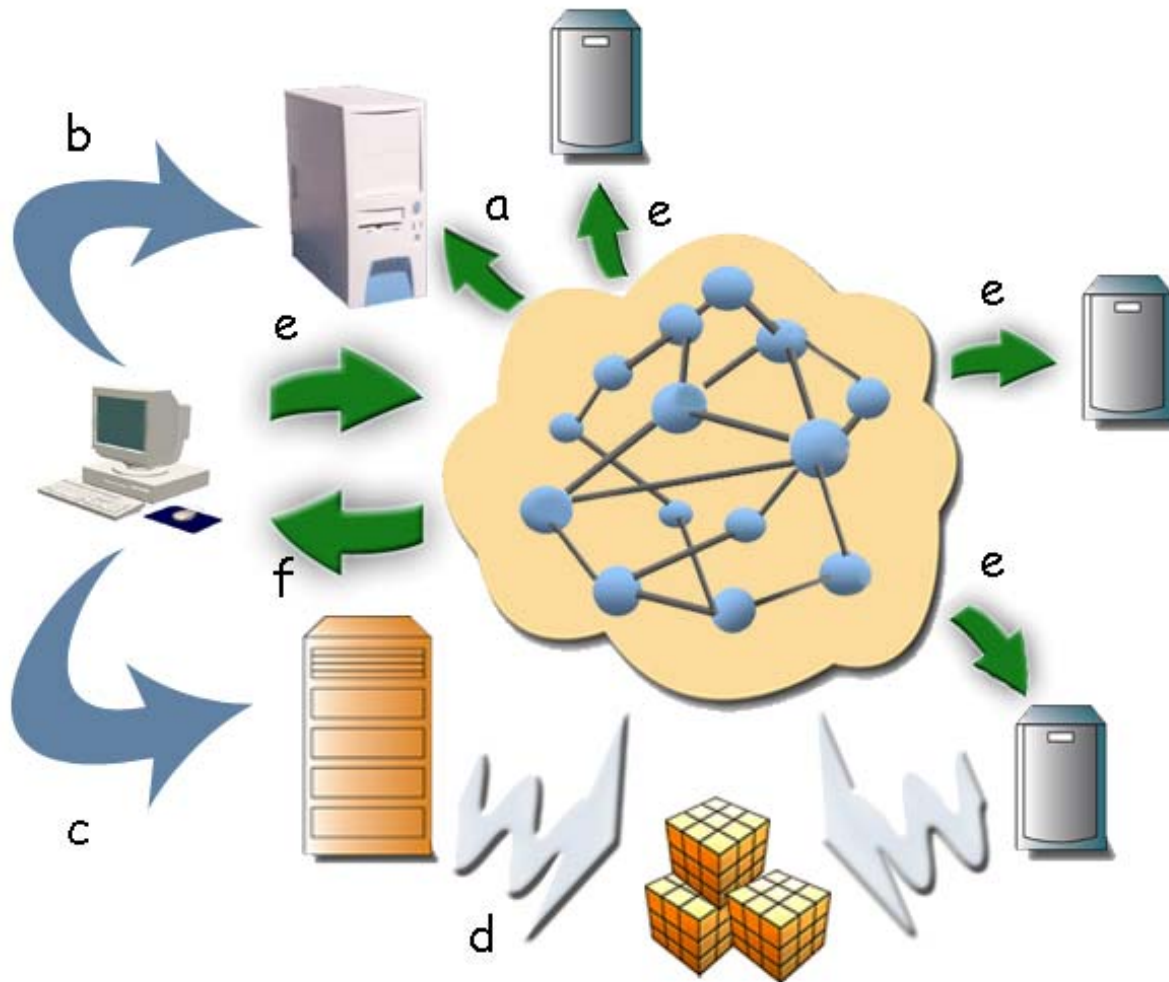
Handling Spectral Data

- **Data Manipulation**
  - e.g., storage and querying

- **Data Preprocessing and preparation for further analysis**
  - Noise reduction;
  - Baseline Correction;
  - Etc…

- **Studying of Data Mining Techniques for Knowledge extraction**
  - e.g., disease early detections.
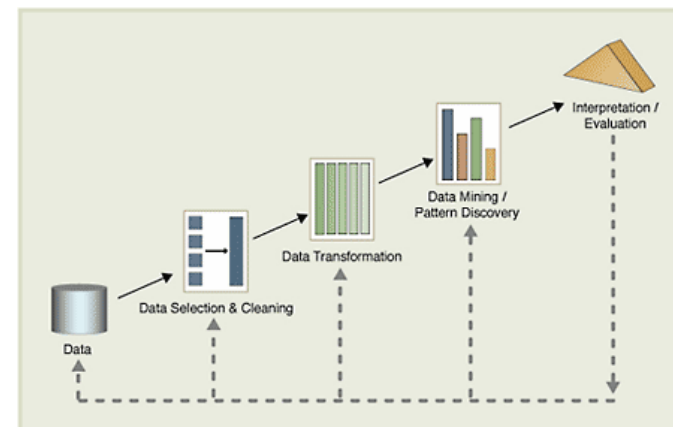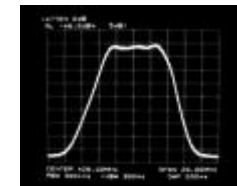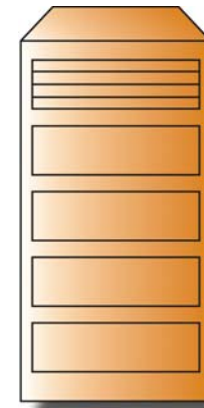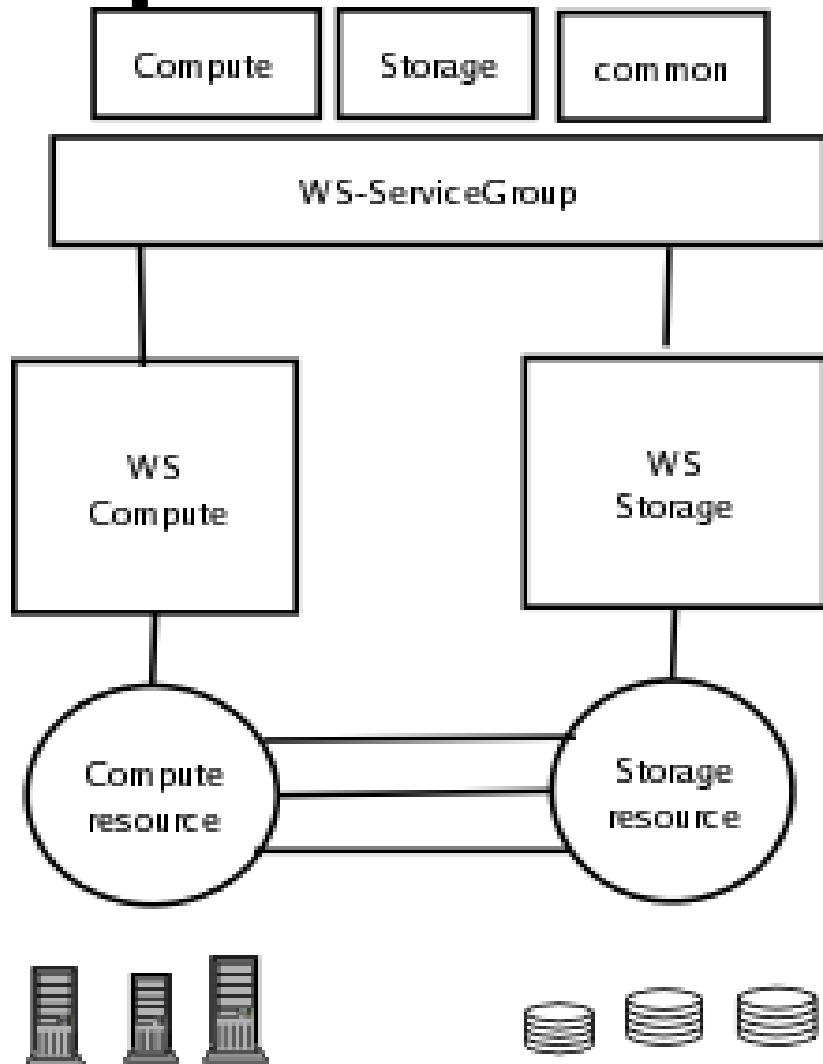
Current Architecture

a. Each server notifies its working status to a remote registry;

b. The client queries the registry to know every available service;

c. The client stores the dataset to be analyzed in a remote repository;

d. The dataset is stored in a database by a local server;

e. The client calls the repository to get a portion of dataset to be analysed;

f. Each service returns its results to client.

Proposed Grid Architecture

Compute	Storage	common

WS-ServiceGroup

WS Compute

WS Storage

Compute resource

Storage resource

Two kinds of services are needed: Storage and Compute, grouped by means of WS-ServiceGroup;

The client calls the two services by means of the grouped interface. More specialized interface is to be defined to use Compute and Storage service in a call.

The Storage WS is asked to store the dataset in a remote repository;

The Compute WS is asked to start the data analisys passing it the EPR pointing to the data;

Potential Technology for Deployment

- WSRF (globus probably)

- Storage: OGSA-DAI

- Computation: Condor-G, LCG-gLite

- Patience (to have all the staff working together)

the globus® alliance

Work completed

- Local Explorer (1.2)
- Regular Explorer of WS (3.2)
- 2 pillars found by human-grid
  - we distributed points between the members of group
  - we did lot of scanning and recursive exploring
  - we found 2 pillars
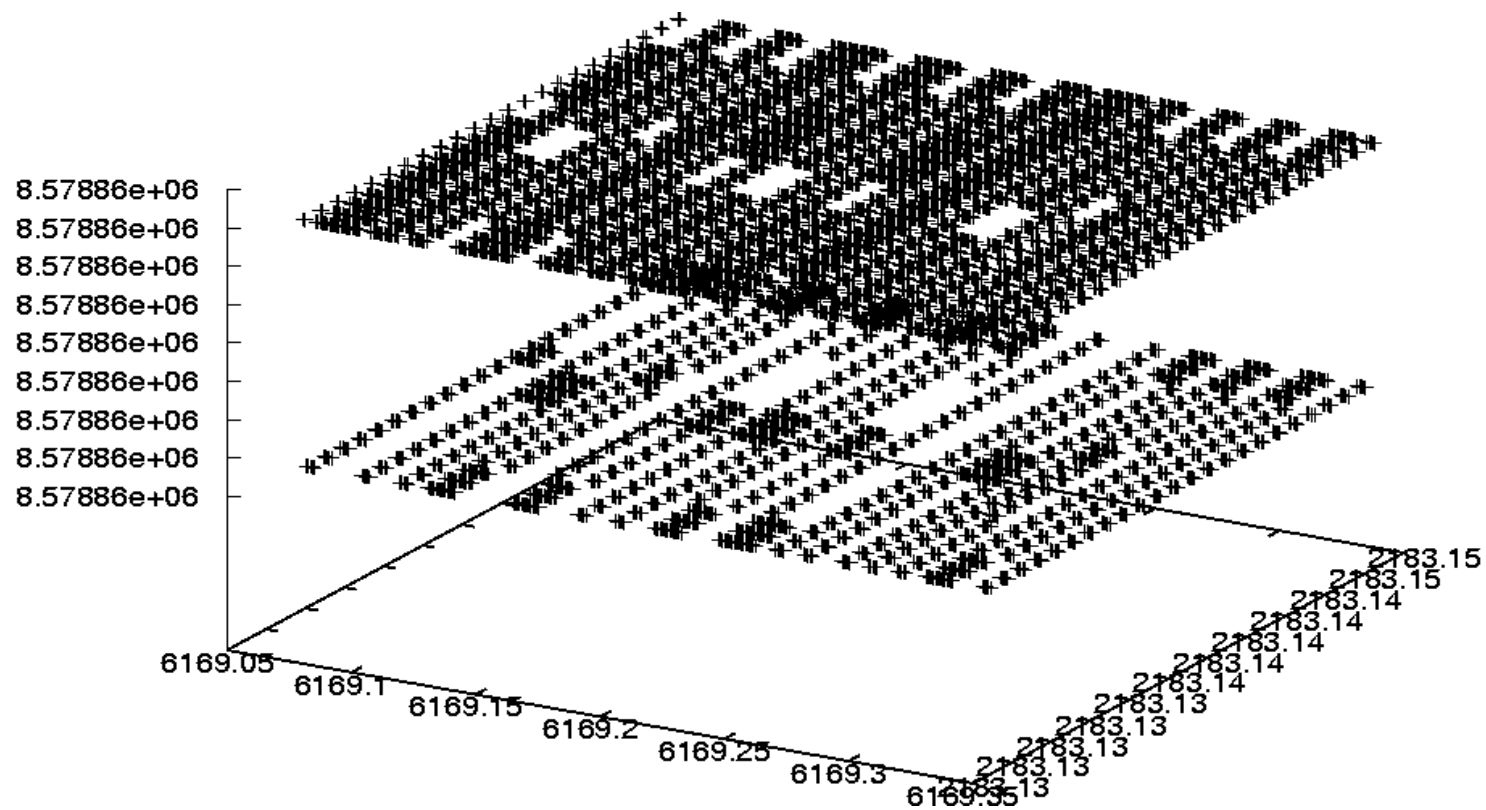 - no grid technology used except WS clients

First pillar :GILDA

Coordinates: 6169.069 2183.128 6169.35 2183.146
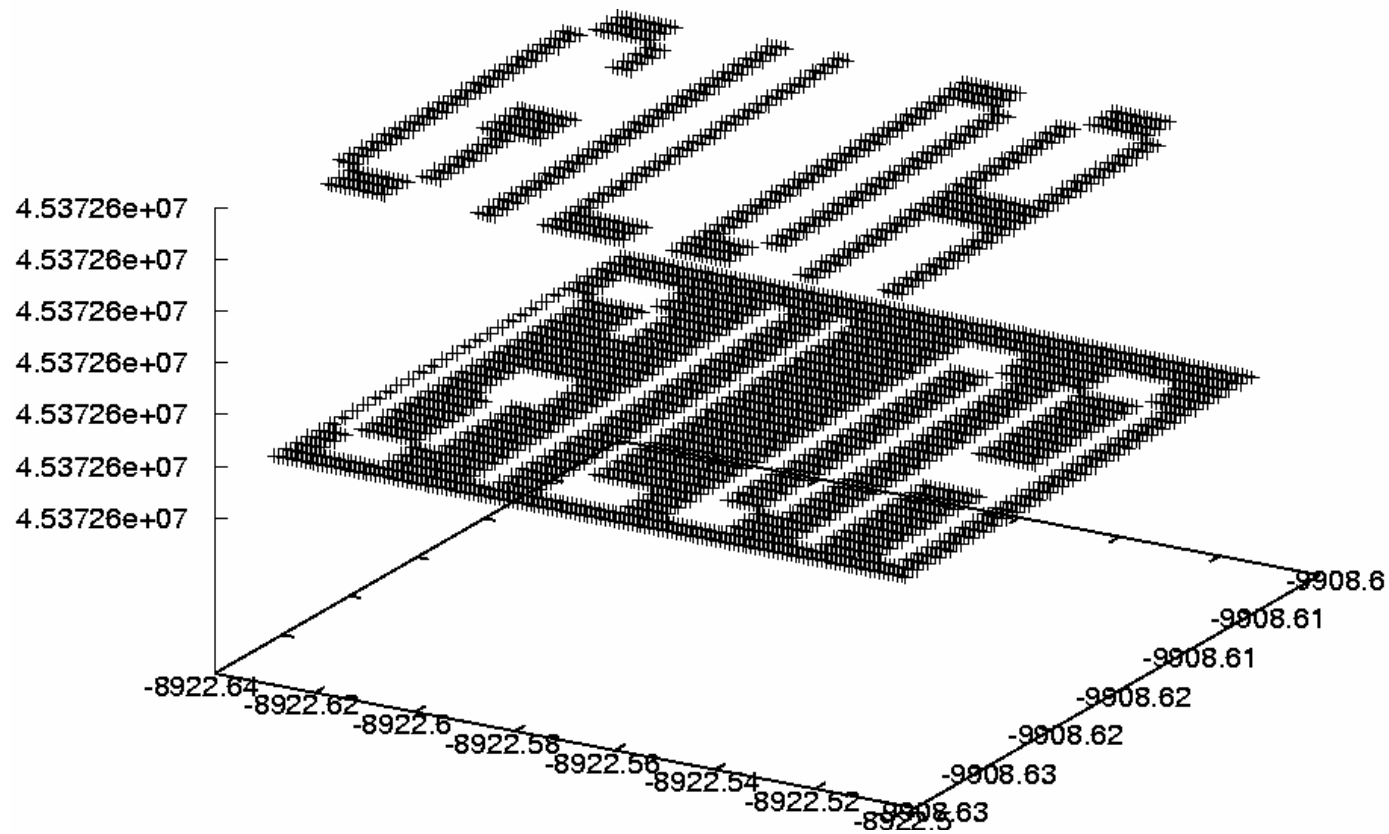
Second pillar :FAB GAGLIARDI

Coordinates: -8922.639 -9908.627 -8922.512 -9908.60

Great plans we had

- Write WS client which gets 2D file and returns 3D file
- Write 2D file generators
  - Scanner
  - Fast Fourier Transformation to detect edges
  - Genetic algorithm to find max points
- Run everything parallel using Condor and LCG
- Client to get data from OGSA-DAI

Members roles

- **Valerio** – chief Debugger
- **Asli** – chief Explorer
- **Tommaso** – chief DataMan
- **Giovanni** – chief Distributer
  Yoonhee – chief Scanner
- **Dario** – chief FourierMan
- **Marko** – chief DagMan

Feedback on the School

+ Lots of Information

+ Good idea to have an exercise that integrates all the knowledge

+ We appreciate the hard work that was done for preparing the school

- Too much useless Java debugging

 Enough fish and pasta