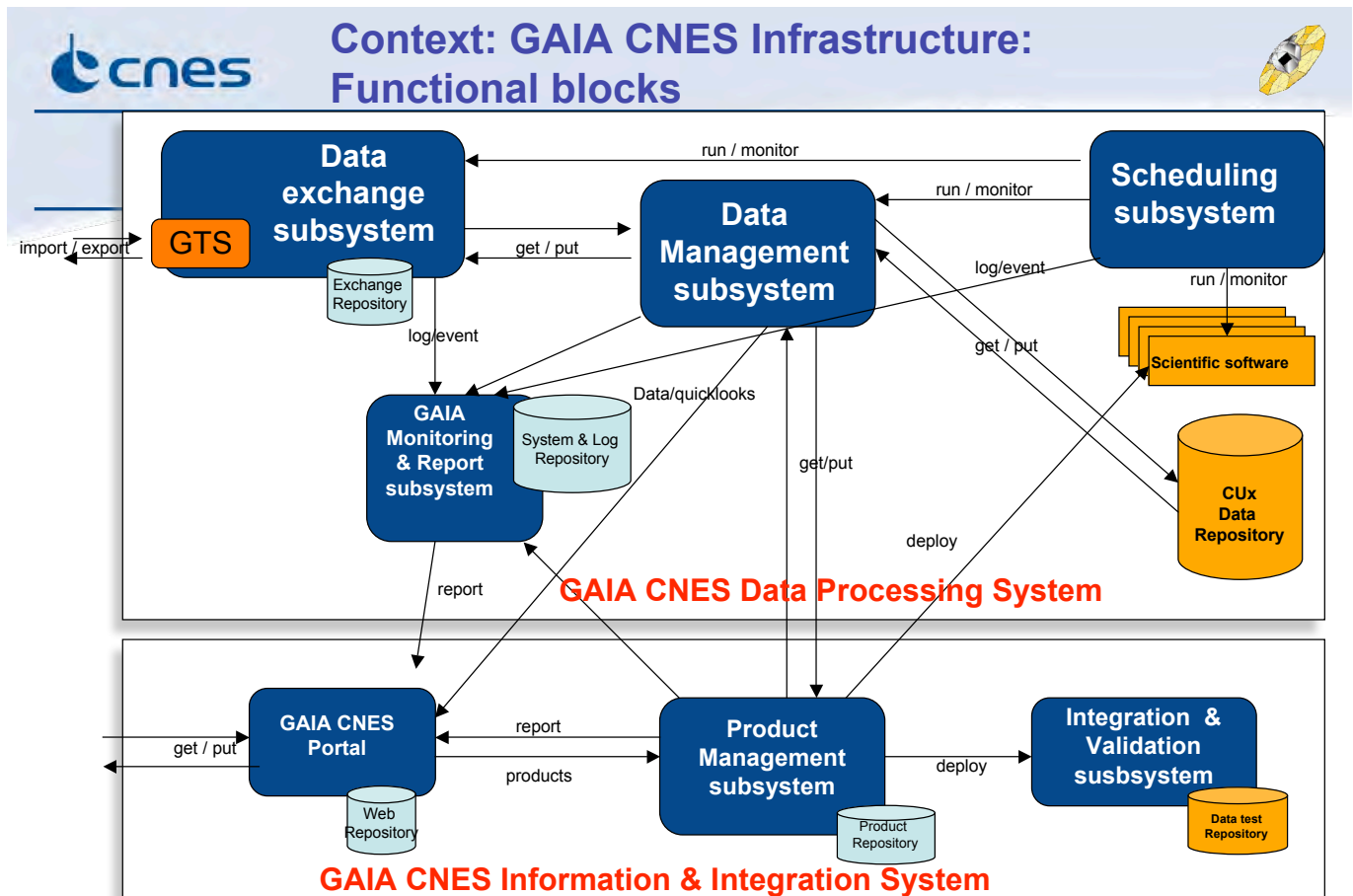
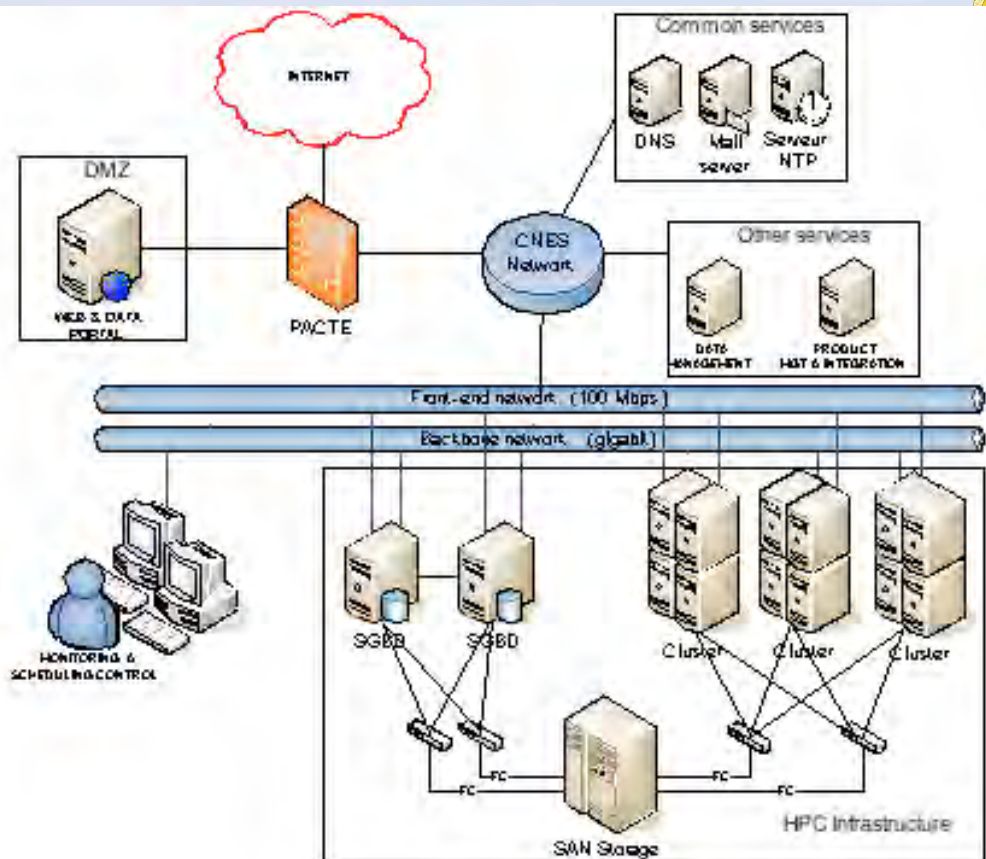


# GAIA – CU6 Bruxelles Meeting (12-13 october 2006)

## Preparation of CNES DPC Infrastructure Technology studies

prepared by F. Jocteur Monrozier





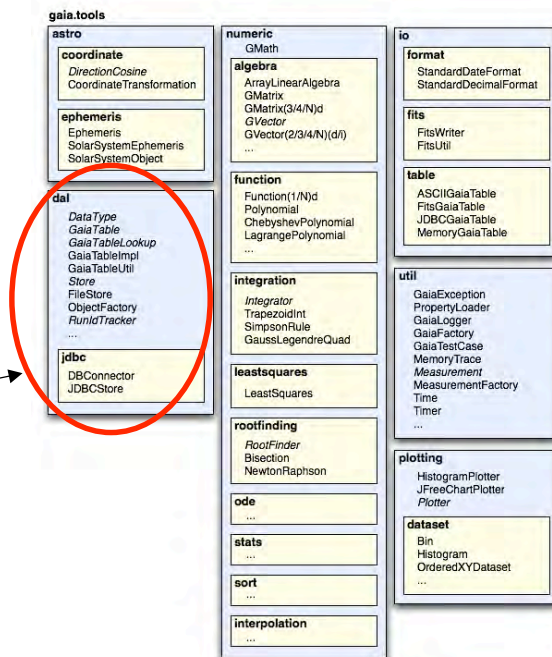
- Scientific processing will be managed as **workflows**
- A workflow is the description of how **a set of programs** can be executed (dependencies, chain, constraints ...).
- A program execution is **a job** that is executed on a cluster node
- HPC problematic (a summary) is to allocate jobs on cluster nodes in order to optimize the infrastructure use (i.e : job scheduler & resource manager)
  - ♦ To do that, each job is described with:
    - Commands to be executed
    - Description of resources needed (estimation of CPU time execution, memory used)



## Context : Data access optimisation

### ■ One of the CNES DPC problematic is to optimise the integration of algorithms within programs in order to optimise data access

- ♦ Optimisation of data access will be done during integration of Soft. Products at CNES
- ♦ Data structure access must be managed (see ICD @ CU level)
- ♦ A unique Data access interface must be used : **DAL implemented by ESAC**
  - see GaiaTools : DAL package (Data access Layer)



### ■ (Another problematic is Algorithm integration : Algorithms should follow some design rules in order to be easily integrated (TBD))



## Technology activities

### ■ Infrastructure :

- ♦ acquisition and installation of 3 dedicated servers (Linux RHES4– bi-processor 2.6Ghz, 4GB RAM)
- ♦ connected to a SAN storage (1TB).

### ■ On going studies and experimentations (3Q–4Q 2006)

- ♦ **Job allocation optimisation - BSC GridSuperscalar toolkit**
  - Report available
  - Test development achieved
- ♦ **Database benchmarks - Oracle RAC** (load balancing)
  - SRS (data model, ...)
  - Configuration of Oracle
  - Test development not started
- ♦ **Data access Layer - DAL study**
  - Draft report
  - Test development not started
- ♦ **Data transfer service - GTS solution**
  - Initial SRS
  - Draft report
  - Test development started
- ♦ **Performance - Java benchmarks**
  - Draft report
  - Test development started
- ♦ **Meta-scheduling (integration of job clusters – i.e several clusters)**
  - Draft report (criteria definition)

All reports will be soon available on the Gaia Wiki site.

Any comments are welcome.

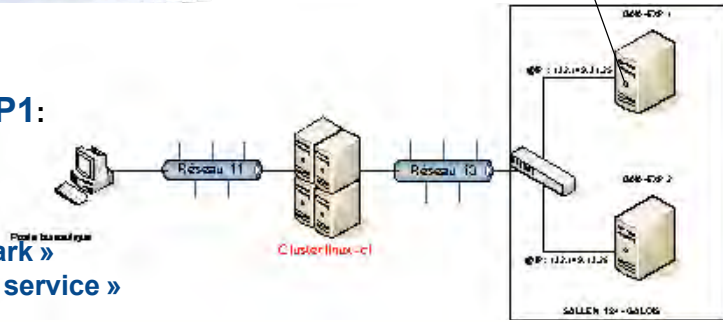
SUN X4200  
bi-proc  
Opteron AMD  
2,6Ghz – 4Go



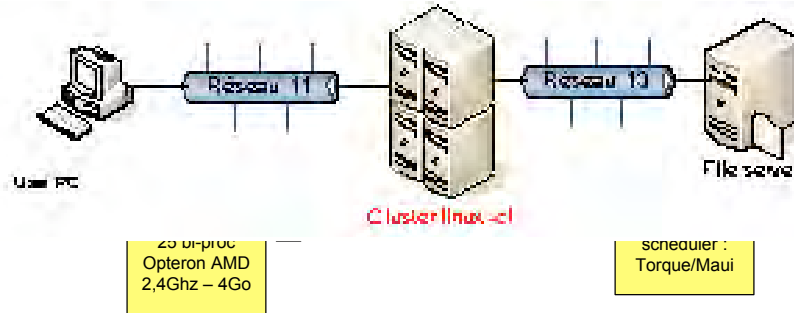
## Architectures for benchn

### Configuration GAIA-INFRA-EXP1:

- experimentation « Java benchmark »
- experimentation « Gaia Transfer service »



### Configuration GAIA-INFRA-EXP2: experimentation « BSC toolkit »



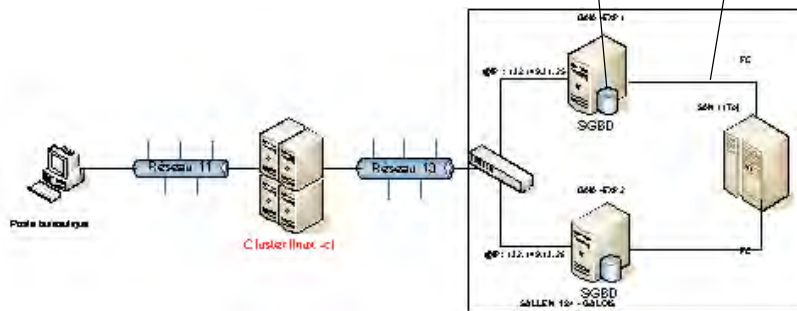
## Architectures for benchmarking experimentations

Oracle 10g configuration  
Single & RAC

FC : 2Gb/s

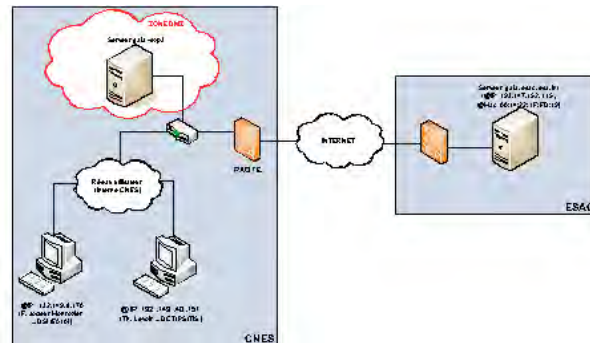
### Configuration GAIA-INF

- experimentation « Database (mode single & RAC)
- experimentation « Data acce



### Configuration GAIA-INFRA-EXP4:

- experimentation « GTS »





# Schedule

