

Modelling Mixed Workflows between Grid and HPC in EUFORIA

Isabel Campos, Enol Fernández, Luis Cabellos,
Marcin Plociennik, Michal Owsiak,
Bernard Guillerminet and Alejandro Soba

Outline

- Introduction
- Grid Infrastructure
- HPC Infrastructure
- Workflow management
- Conclusions

Introduction

- EUFORIA (EU fusion for ITER applications)
- **ITER** is an experiment that aims to demonstrate that it is possible to produce commercial energy from fusion.
- Requires a high degree of physics modelling and simulation (even in current construction phase)
- Plasmas are **complex** systems
 - Very demanding from a computational point of view
 - Simulate 1ms takes **several days** in a 50-100TFlop machine
 - ITER expects to reach **5 minutes**

Introduction

EUFORIA Objective:

Develop a comprehensive framework and infrastructure for the european fusion modelling community.

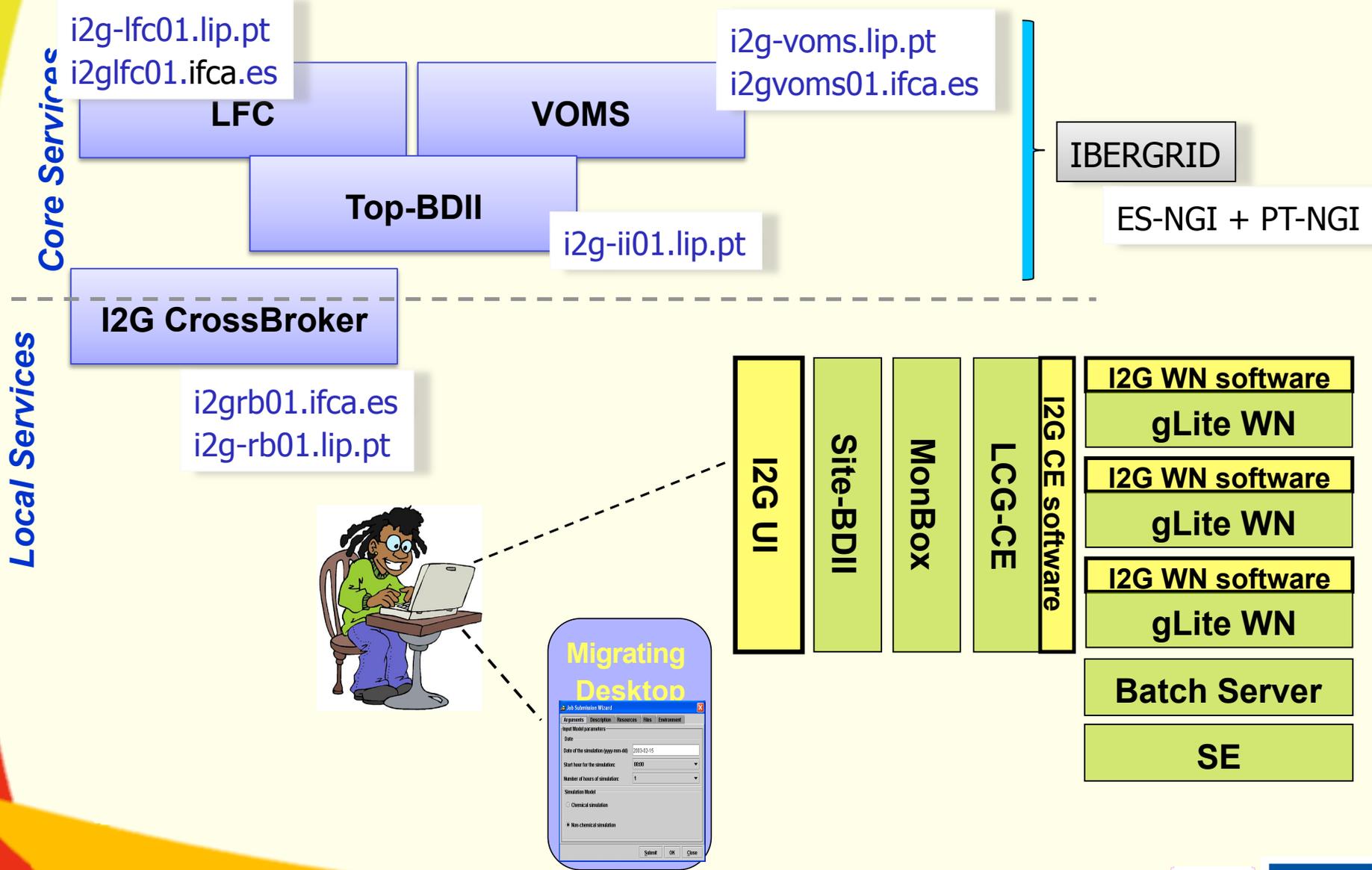
- Deploy **Grid** and **HPC** infrastructures
- Adaptation and optimization of fusion codes
 - Oriented to grid and/or HPC
- Development of advanced tools for:
 - Workflows
 - Visualization
 - Data Mining

Grid Infrastructure

- Largely based on the developments of the *Interactive European Grid* project (<http://www.i2g.eu>)
 - Advanced Scheduling Capabilities (**CrossBroker**)
 - Support to **parallel** jobs with MPI (mpi-start)
 - Support to **interactive** jobs via i2glogin
 - **RAS** in combination with **Migrating Desktop**
- Full **interoperability** with other gLite based infrastructures
 - Like **EGEE**
- Access through the **EUFORIA VO**



Grid Infrastructure



Grid Infrastructure

- Centers with infrastructure in Euforia
 - **IFCA** in Santander, Spain
 - **FZK** in Karlsruhe, Germany
 - **Chalmers** University, Sweden
 - **Ciemat** in Trujillo, Spain
- From the Grid Information System (~2700 CPUs, 40 TB online)

#CPU|Free|Total|Jobs|Run|Waiting|ComputingElement

```
-----  
1488 702 0 0 0 0 i2gce01.ifca.es:2119/jobmanager-lcgpbs-euforia  
452 324 0 0 0 0 iwrce2.fzk.de:2119/jobmanager-lcgpbs-i2gpar  
640 113 0 0 0 0 svea-gl2.c3se.chalmers.se:2119/jobmanager-lcgpbs-euforia  
56 56 0 0 0 0 ce-euforia.ceta-ciemat.es:2119/jobmanager-lcgpbs-euforia
```

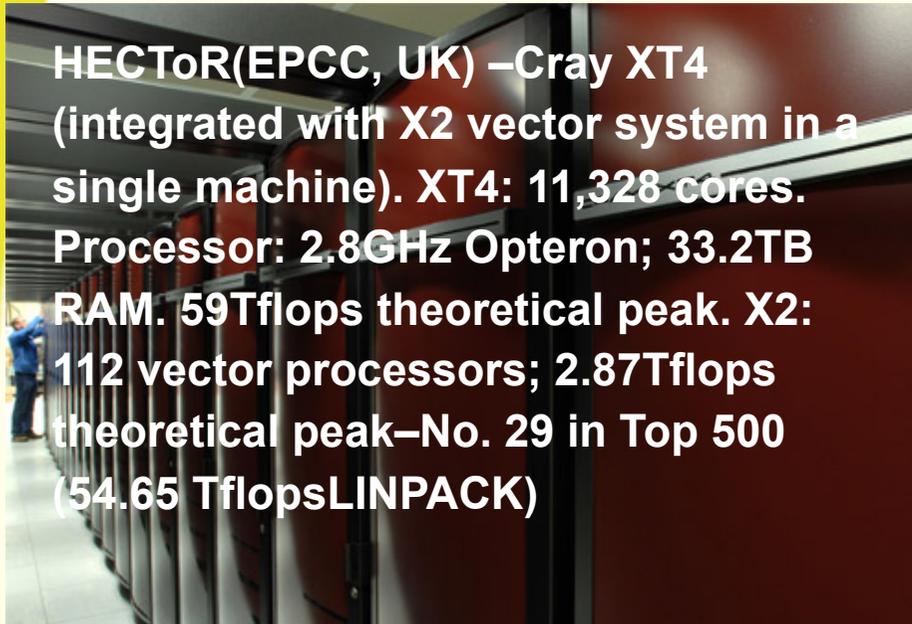
Avail Space(Kb) Used Space(Kb) Type SEs

```
-----  
38 TB 6TB n.a storm.ifca.es  
1.8 TB 129931 n.a iwrse2.fzk.de  
142 GB -- n.a. svea-gl3.c3se.chalmers.se
```

Grid Infrastructure: towards EGI

- Central services are **guaranteed** by IBERGRID (LIP and CSIC)
- Main middleware pieces have a **continuity path** in EGI
 - Migrating Desktop and RAS (PSNC)
 - CrossBroker and MPI support (CSIC)
- **Infrastructure** itself is already a part of the NGIs resources
 - CSIC and Ciemat in Spain
 - FZK in Germany
 - Chalmers in Sweden

HPC Infrastructure



HECToR(EPCC, UK) –Cray XT4
(integrated with X2 vector system in a single machine). XT4: 11,328 cores. Processor: 2.8GHz Opteron; 33.2TB RAM. 59Tflops theoretical peak. X2: 112 vector processors; 2.87Tflops theoretical peak–No. 29 in Top 500 (54.65 TflopsLINPACK)



Louhi(CSC, Finland)–Cray XT4.
4,048 cores. 4.5TB RAM. 37.68Tflops peak.–No. 70 in Top 500 (26.80 TflopsLINPACK)



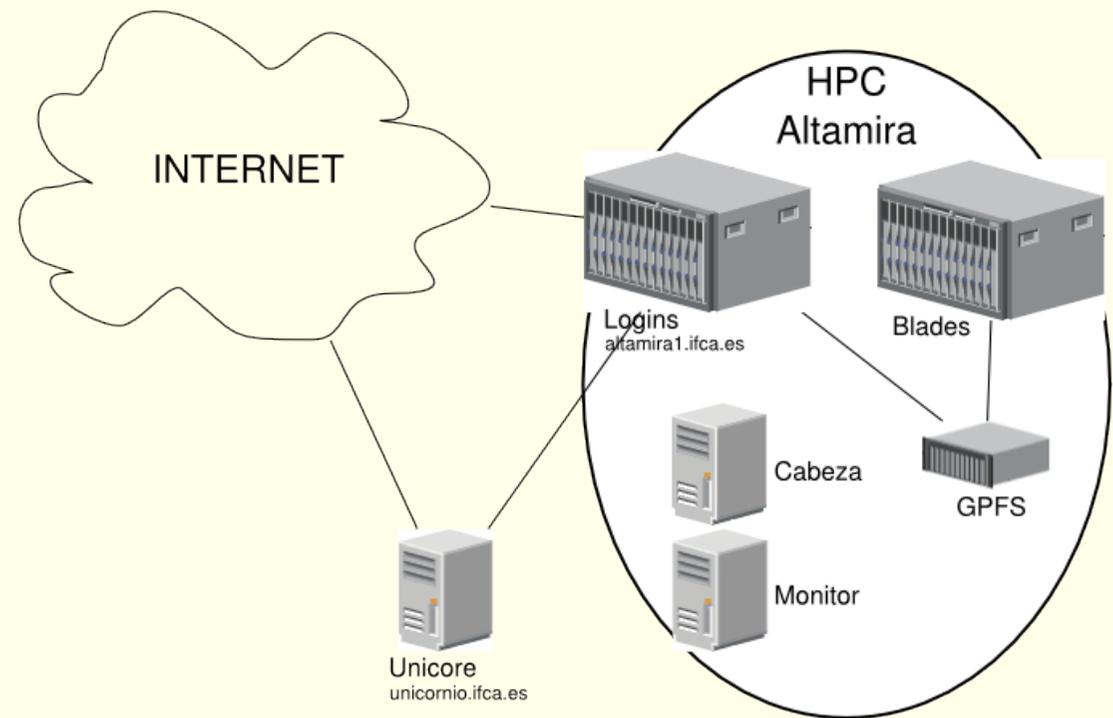
MareNostrum(BSC, Spain)
–IBM Cluster. 10240 cores. Processor: 2.3GHz PPC 970; 20TB. 94.21Tflops peak. –No.26 in Top 500 (63.83 TflopsLINPACK)

HPC Infrastructure

- HPC ecology more diverse than Grid
- Codes **optimized** for running in the facilities
 - Installation/compilation and profiling done until June 2008
 - Now users can execute their applications (2M CPU Hours available)
- Access through **DEISA**

HPC Infrastructure

- **Unicore** interface for **Altamira** (HPC facility at IFCA)
 - Access with User certificates
 - Job execution
 - File staging



Codes	Grid	Ported to (centre)	Libraries needed
CENTORI	No	Cray XT4 (EPCC)	---
GEM	Yes	Cray XT4 (EPCC-CSC) PPC970 (BSC)	FFTW 3.1.1
ERO	No	Cray XT4 (EPCC)	IBM ESSL and PMAPI
EIRENE	Yes	PPC970 (BSC)	---
ELMFIRE	Yes	PPC970 (BSC) Cray XT4 (BSC)	GSL, BLAS, PETSC, PESSL
GENE	Yes	PPC970 (BSC)	ESSL, MKL or FFTW BLAS/LAPACK/ PETSC/SLEPC OpenMP
BIT1	Yes	PPC970 (BSC) Cray XT4 (CSC)	X11
Esel	No	CSC	HDF4(version 4.2r2)
Isdep	Yes	CSC	---
SOLPS	No	CSC	NCARG (version 5.0.0) NetCDF (version 3.6.2) ACML (version 4-0-1) (or some other BLAS/ LAPACK library) X11

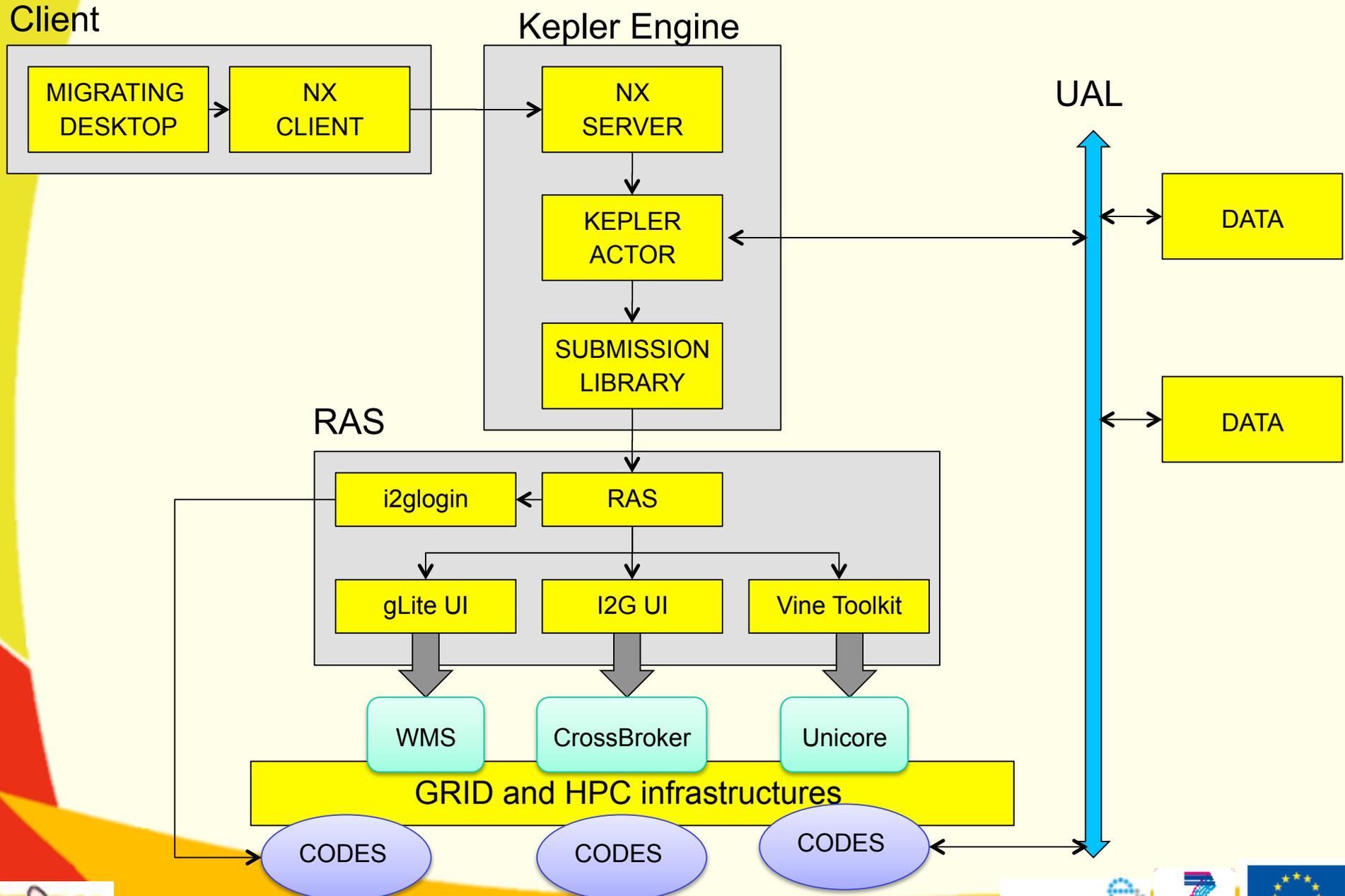
Workflow management

- End-to-end simulation of a fusion device requires **integration** and **coupling** of several complex codes
- Workflows started on a central facility
- Computation done elsewhere (mainly)
 - On the GRID
 - Serial and small scale parallel
 - For scenario scans
 - On an HPC
 - For larger runs
 - Chosen by the workflow on the basis of
 - Suitability
 - Minimum cost
 - Minimum time to completion
- Results (or a “reference”) are brought back to the central facility for permanent storage and further analysis

Workflow management

- EUFORIA integrates those codes and existing tools as Kepler workflows:
 - Kepler already used by the fusion community
 - Workflows seen as a set of **actors** with I/O ports for communication
 - Actors can be nested (**hierarchical workflows**)
- New actors developed in EUFORIA
 - Submission to i2g and gLite
 - Submission to HPC (via Unicore)

Workflow management



Submission Library

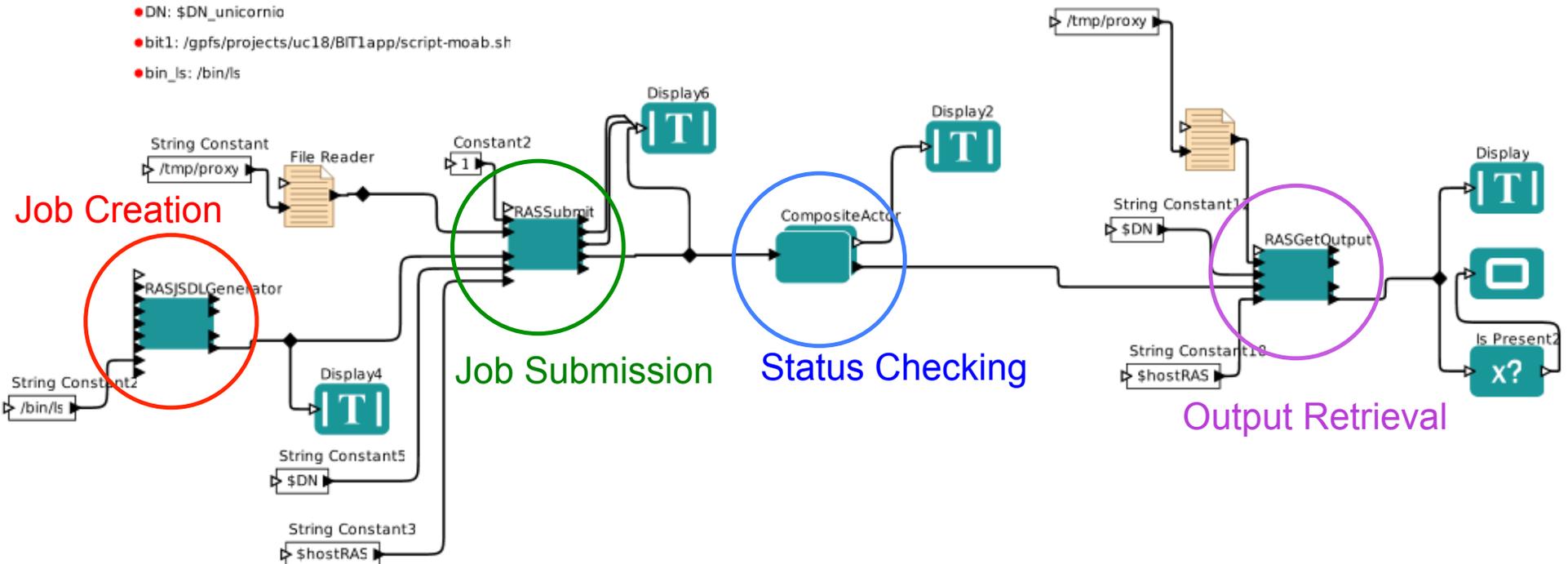
- Java API designed submit jobs
 - Using standard JSDL
 - Manages grid credentials (including VOMS)
 - Submission to different infrastructures
 - Manages input and output staging
- RAS+Vine
 - Performs actions on the grid on behalf of the MD
 - Vine toolkit (from OMII-EU) Allows submission to several middlewares

Workflow management

DDF Director

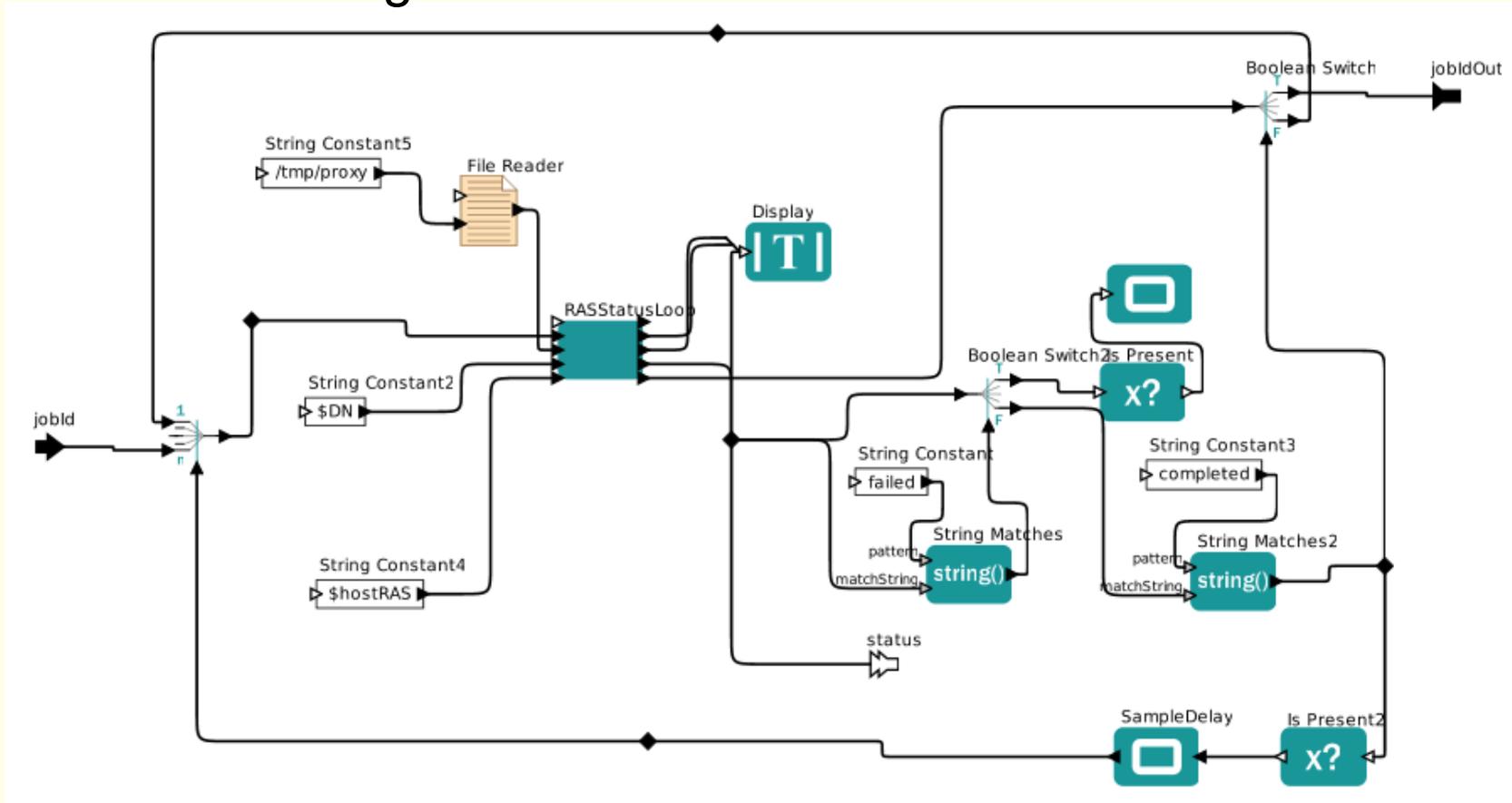


- hostRAS: http://gridui01.ifca.es:8090
- DN_seagrass: seagrass.man.poznan.pl
- DN_unicornio: unicornio.ifca.es
- DN: \$DN_unicornio
- bit1: /gpfs/projects/uc18/Bit1app/script-moab.sh
- bin_ls: /bin/ls



Workflow management

Status Checking Actor



Conclusions

- EUFORIA develops an integrated environment for fusion modelling
 - Grid and HPC infrastructure
 - Use of existing tools from i2g and fusion community
 - RAS/Migrating Desktop, CrossBroker, Vine, Unicore
 - Kepler, UAL
- Scientists able to perform end-to-end simulations of fusion devices with workflows
 - Executing codes in grid and HPC