

# Expériences d'imagerie médicale sur EGEE

Tristan Glatard, Sorina Pop  
CNRS – CREATIS, Lyon

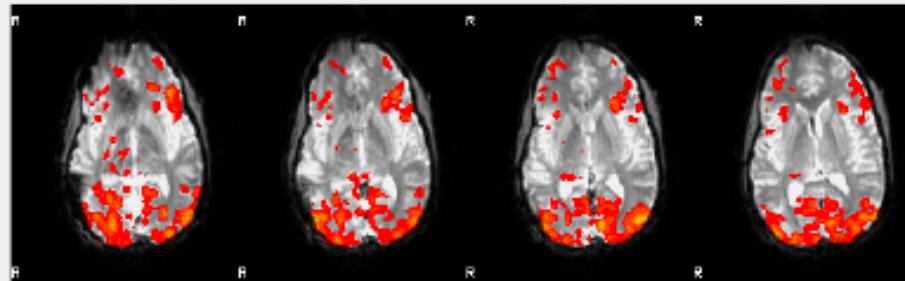


Institut de Recherche pour le Développement  
Montpellier  
26 mars 2010

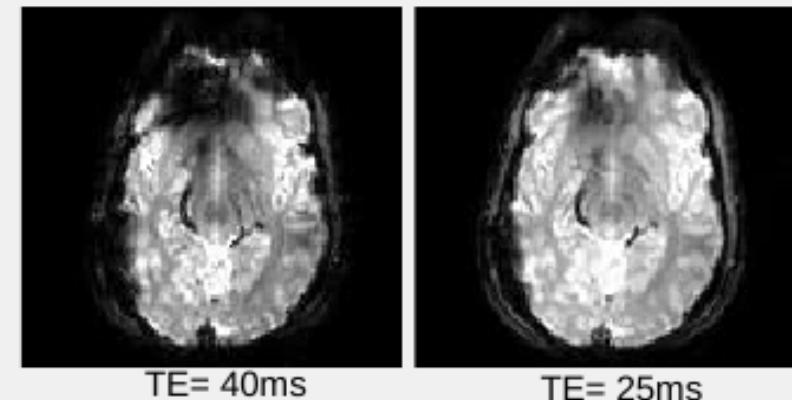
- **Part of the EGEE Life-Science cluster**
  - <https://twiki.cern.ch/twiki/bin/view/EGEE/LifeSciences>
  - Medical imaging, bioinformatics, drug discovery
- **Global VOs**
  - Biomed (100+ sites world-wide)
  - embrace, enmr.eu, moldyngrid.org
  - vo.neugrid.eu
- **Regional**
  - bio, gene, libi, tps.infn.it, vlemed, vo.renabi.fr, lsgrid
- **Multidisciplinary VO**
  - fkppi.kisti.re.kr, vo.iscpif.fr, vo.rhone-alpes.idgrilles.fr

# Parameter sweep in fMRI

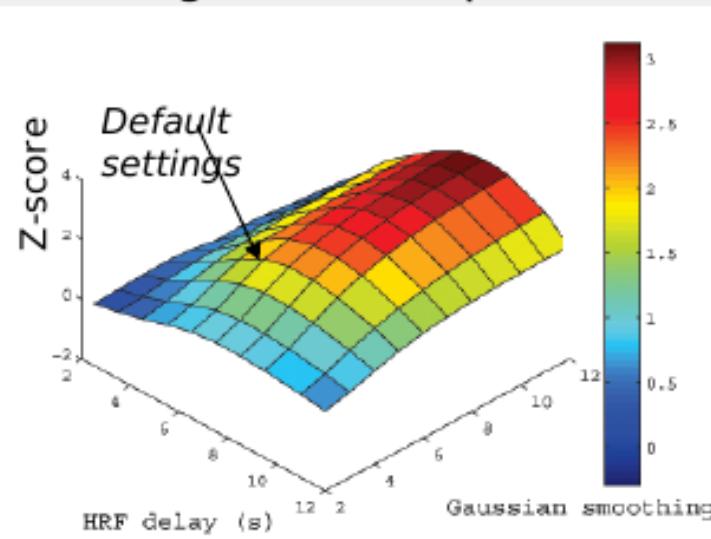
- Activation maps



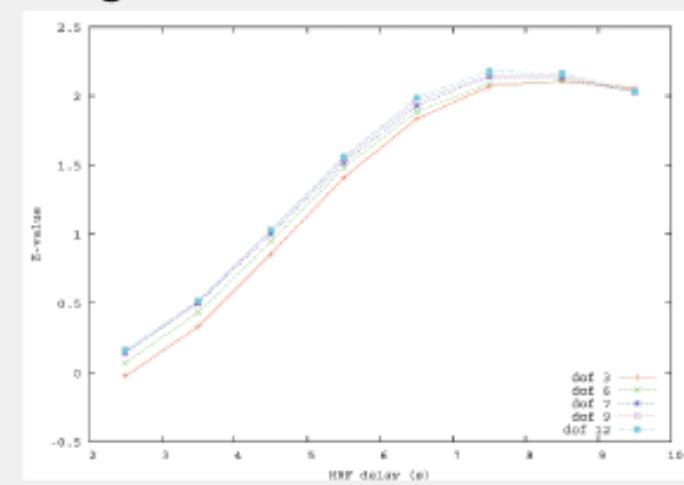
- Echo times comparison



- Smoothing and HRF optimization



- Registration evaluation



- 10,000 jobs - 1 CPU year - 1 week elapsed time - 1.5 TB out

# Grid execution

- Infrastructure
  - 4 sites (687 CPUs, 500TB) of the Dutch grid (part of EGEE, vlemed VO)
- Results

#P <sub>i</sub>	#T	#S	#D	#H	# Analyses	CPU (days)	Data (TB)	Elapsed (hours)	Speed -up	# Submit Jobs	Failure (%)
Individual Analyses											
batch 1	11	1	5	5	8	2200	74.9	0.31	14.9	120.5	2200
batch 2	11	1	6	5	8	2640	89.8	0.38	11.6	186.6	2642
batch 3	11	1	6	5	8	2640	89.8	0.38	32	67.38	2687
batch 4	11	1	5	5	8	2200	74.9	0.31	10.2	176.8	2203
total	11	2	11	5	8	9680	329.4	1.38	68.7	115	9732
Group Analyses											
batch 1	1	6	5	8	240	1.4	7.1	8.0	4.3	401	40.15
batch 2	1	6	5	8	240	1.4	7.1	9.5	3.6	240	0.00
batch 3	1	5	5	8	200	1.2	6	14.9	1.9	200	0.00
batch 4	1	5	5	8	200	1.2	6	11.3	2.5	600	66.67
total	2	11	5	8	880	5.2	26.2	43.7	2.9	1441	38.93
Group Difference Analyses											
batch 1	11	5	8	440	7	23.8	44.3	3.8	2650	83.40	



# Mean-Shift filtering optimization

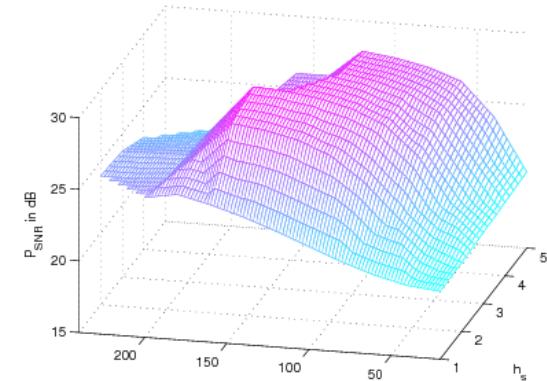
- Sweep on scale parameters of MS filter



Noisy image



Restored image



PSNR w.r.t scale parameters

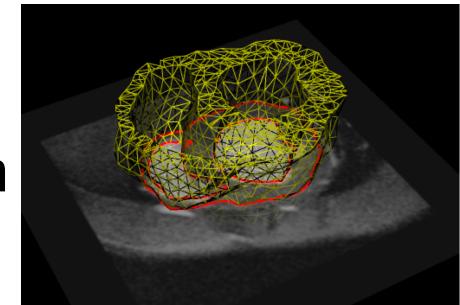
- Grid facts (Matlab code)

$h_s$ value	Total CPU time	Elapsed time	Speed-up	Total data transfer time	Produced data	Successful tasks	Total tasks	Error ratio
1	13.0 days	3h25min	92	89.8h	43GB	8,000	8,106	1.3%
2	50.4 days	18h36min	65	63.2h	41GB	8,000	8,929	10.4%
3	17.3 days	13h01min	32	34.8h	5GB	1,000	1,317	24%
4	29.0 days	13h23min	52	44.7h	5GB	1,000	1,089	8.2%
5	54.3 days	11h09min	117	51.9h	5.1GB	1,000	1,179	15.2%

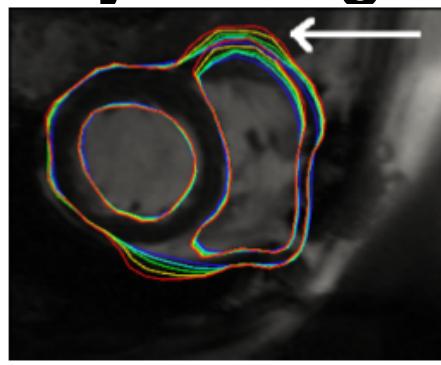
# Cardiac MRI segmentation

- **Cardiac segmentation with deformable models**

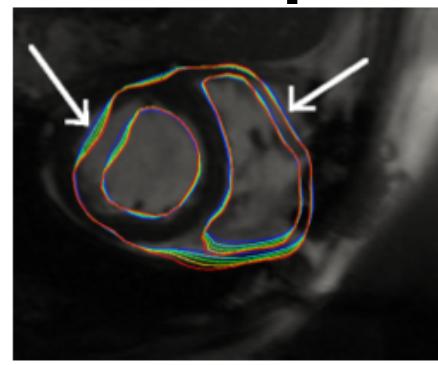
- Volumetric template mesh
- Initialized with rigid registration
- Image gradient ▶ force field ▶ deformation



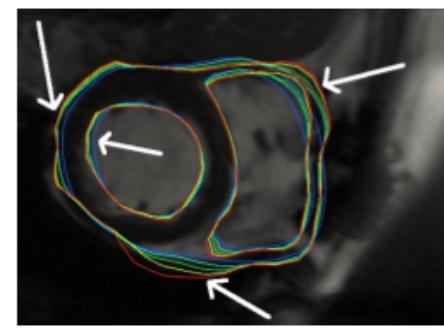
- **Adjust segmentation parameters**



init 0



init 1



init 2

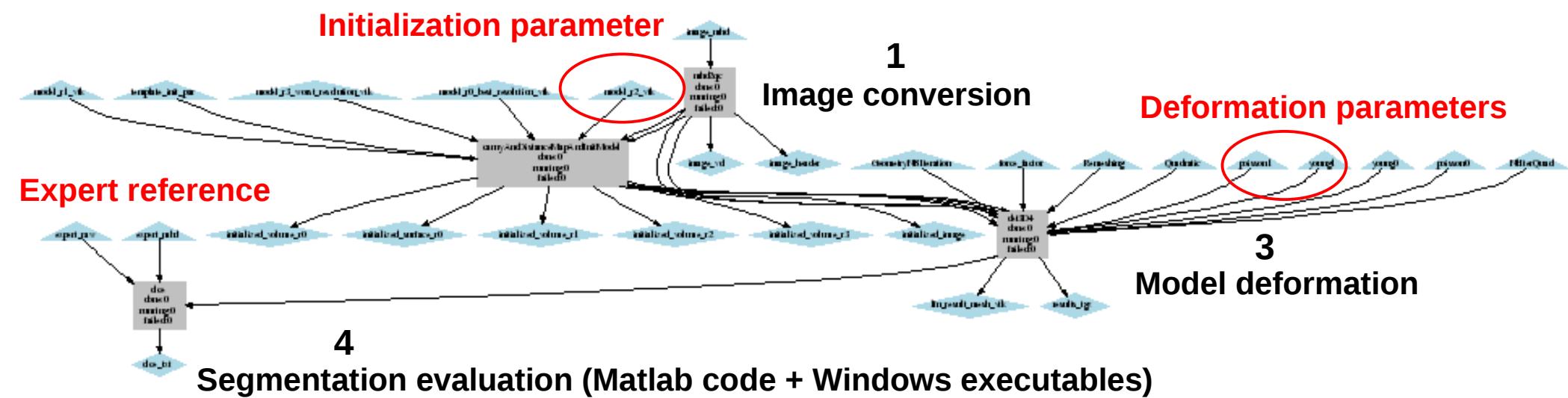
—	force factor = 0.1
—	force factor = 0.2
—	force factor = 0.3
—	force factor = 0.4
—	force factor = 0.5

- **Estimate myocardium physical parameters**

- Best segmentation ◀ most realistic parameters  
(e.g. Young modulus)

# Cardiac segmentation workflow

- Main steps:
  - Image conversion
  - Initialization
  - Deformation
  - Evaluation w.r.t expert reference



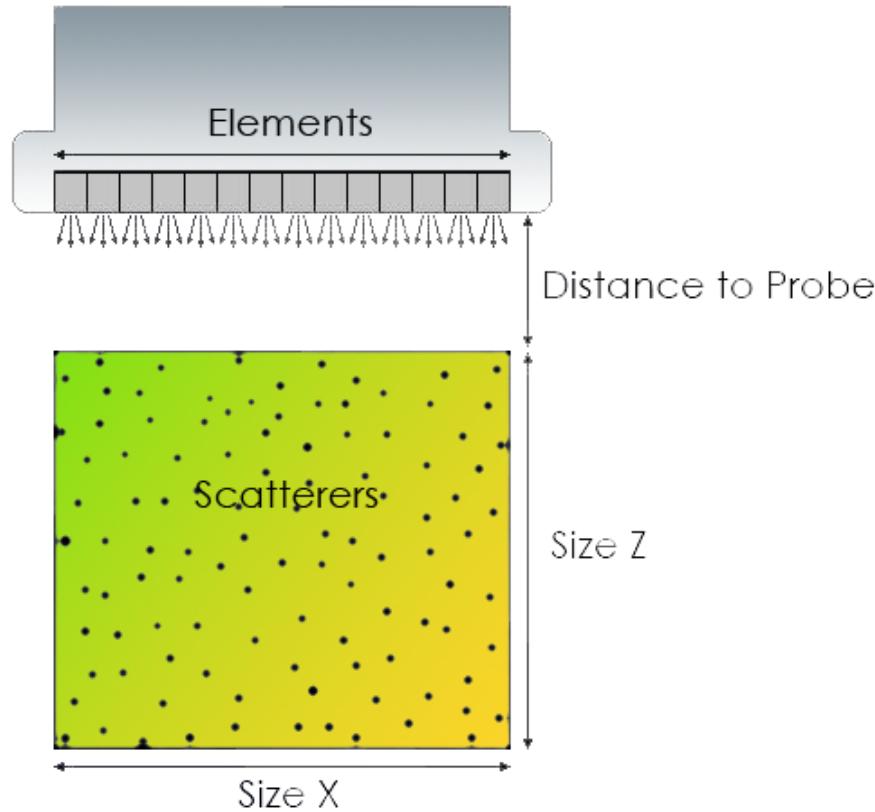
- Workflow iterated on parameter/image sets

# FIELD US simulation

<http://server.oersted.dtu.dk/personal/jaj/field/>

- Principle

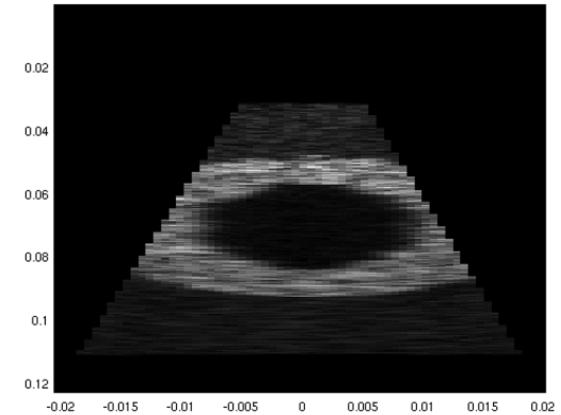
PROBE



MEDIUM

- Parallelism on lines
- Parallelism on mediums

- Example on 2D beating heart

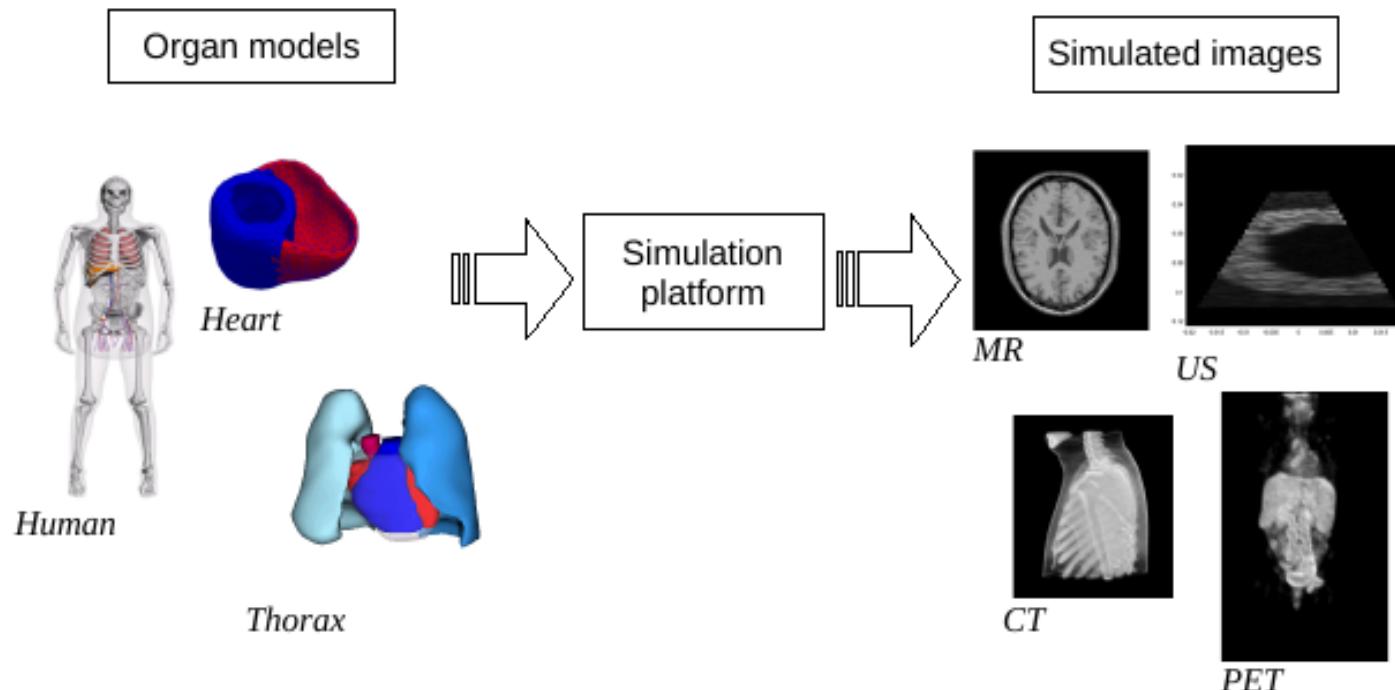


- 1920 lines (30 mediums x 64 lines)
- > 16h => < 3h
- 12% error (first try) => 2% (fine-tuning)

# Virtual Imaging Platform

<http://www.creatis.insa-lyon.fr/vip>

- **Simulation of medical images from organ models**
  - Includes semantic information
  - Linked to computing platforms



# Hadrontherapy simulation with GATE

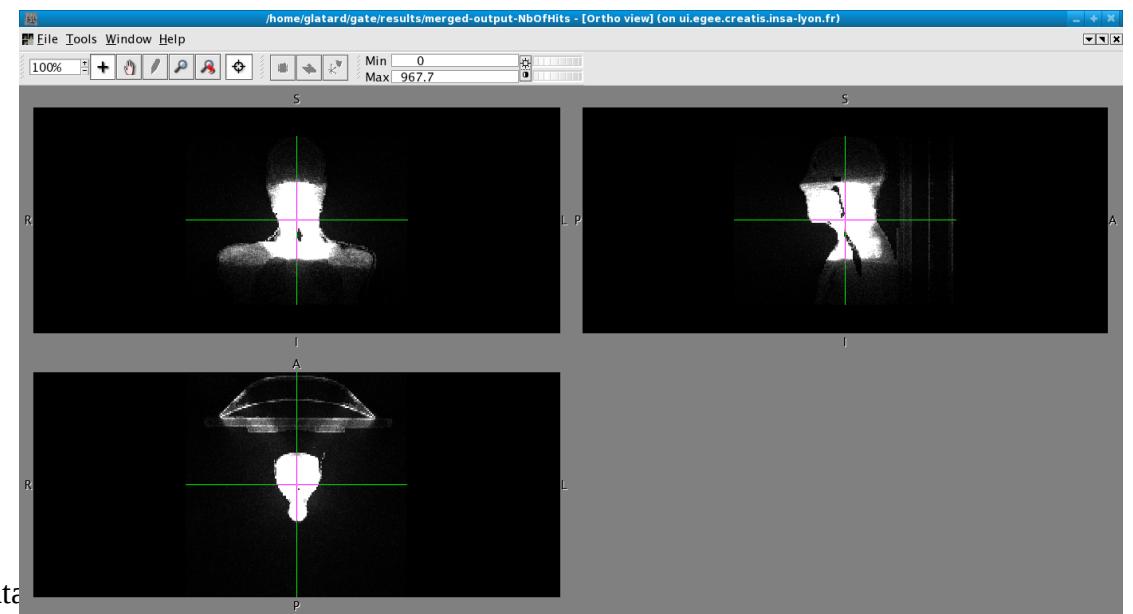
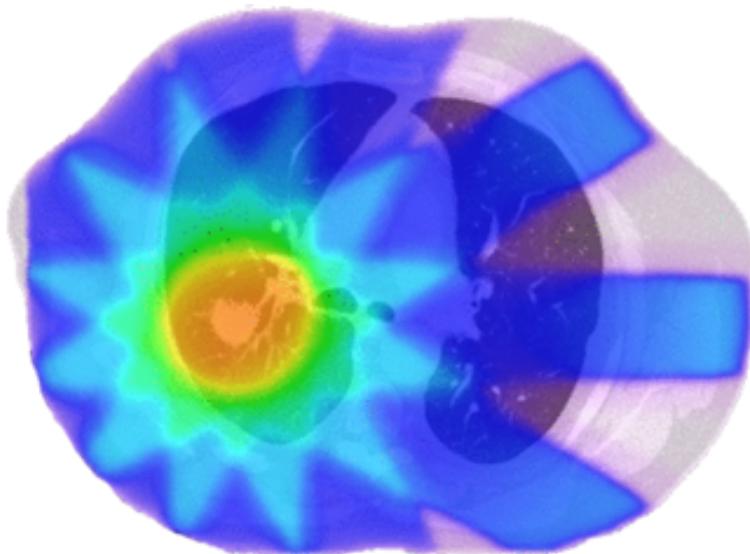
<http://opengatecollaboration.healthgrid.org/>

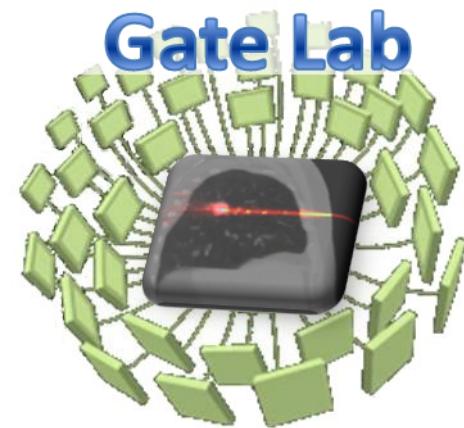
- **Simulation principle**

- Patient scan + source model
- Particle (photon, hadron) tracking through matter
- Hit / dose maps

- **Involved effort**

- High number of particles (~ 20,000,000)
- Monte-Carlo simulation ► divisible load

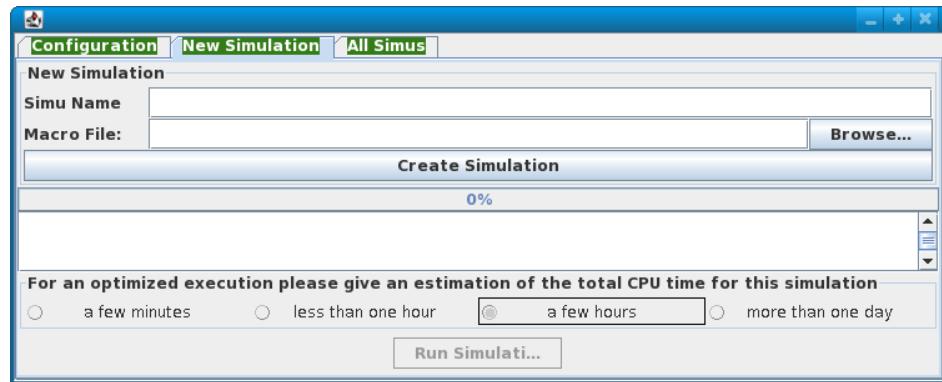




# GATE-Lab interface

- **Dedicated VBrowser plugin**

- Simulation file parsing
- Parameter checking
- Input files bundling and upload
- Time estimation
- History management (+cleanup)



- **Customized server interface**

- #simulated particles
- Current status
- Link to results directory
- Confirm/retry

Info for simulation "09\_12\_09\_09\_46\_Test11\_57" (inactive)

**Inputs**  
Submission time: December 09 2009 09:55:04  
Input directory

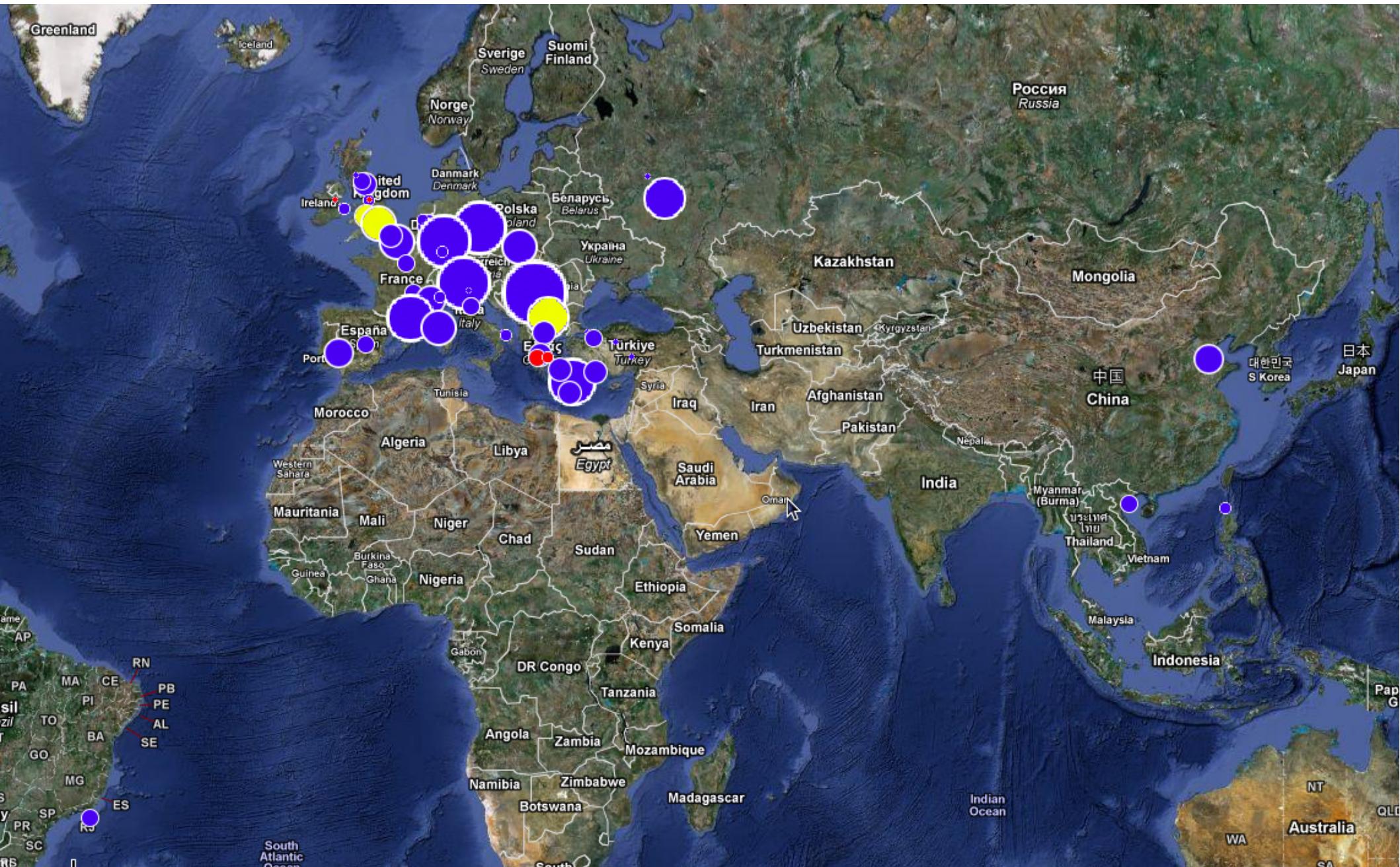
**Simulation status**  
Results directory  
Simulated particles: 995999751 - (99.6%)  
Jobs: Waiting: 0 - Running: 0 - Successfully completed: 251 - Failed: 19 - Timeout: 0 - Cancelled: 0  
Status: Simulation completed (validated)  
Logs

- stdout
- stderr

Total elapsed time: 173767.648s

Advanced

# A (hadrontherapy) experiment



VBrowser[0]:<https://ws1.grid.sara.nl/~glatard/workflow/workflow-PTlv75/workflow-PTlv75.html>

**Location Edit View Tools Windows Help**

**Resource**

- My Vle
  - /home/glatard
  - biomedLFC
  - SRB@sara
  - viz-login
  - gwendia\_cardiac
  - vlemed\_LFC
    - alex
    - amy
    - amy-12
    - amy-ndc
    - back-ups
    - demo-VLfmRI
    - fokke
    - generated
    - glatard
    - jalkemade
    - jeroene
    - kbolebiar
    - martin
    - matthan
    - mdm
    - piter.t.de.boer
    - remi
      - data\_storage
      - db
      - joblogs
      - masks
      - output
      - output-may-2008
      - output\_26-08
      - scripts
    - workflows
      - group
      - individual
        - laps
        - nback
      - roi
        - groupAnalyses
        - individualAnalyses
          - inputs
  - http
    - results
    - silvia
    - testVFLFC
    - testVFLFC2
    - tristan
    - wibisono
    - garbage.sh
    - hello-1228916611960604456.t
  - ccUI
  - Desktop
  - applisCreatis

**Grid files browsing**

**CobraViewer**

Status Services Input Results Info

```

graph TD
    roi[roi  
done:0  
running:10  
failed:0] --> indivAnalysis[indivAnalysis  
done:2  
running:0  
failed:0]
    indivAnalysis --> zstat2standard[zstat2standard]
  
```

**Workflow monitoring**

**JOB STATUS:workflow-PTlv75**

N#	JobID	JobStatus	link Out	S...
1	https://rb.grid.sara.nl:9000...	DONE (SUC...)	Not yet...	
2	https://rb.grid.sara.nl:9000...	DONE (SUC...)	Not yet...	
3	https://rb.grid.sara.nl:9000...	SCHEDULED	Not yet...	
4	https://rb.grid.sara.nl:9000...	READY	Not yet...	
5	https://rb.grid.sara.nl:9000...	READY	Not yet...	
6	https://rb.grid.sara.nl:9000...	READY	Not yet...	
7	https://rb.grid.sara.nl:9000...	WAITING	Not yet...	

**Job monitoring**

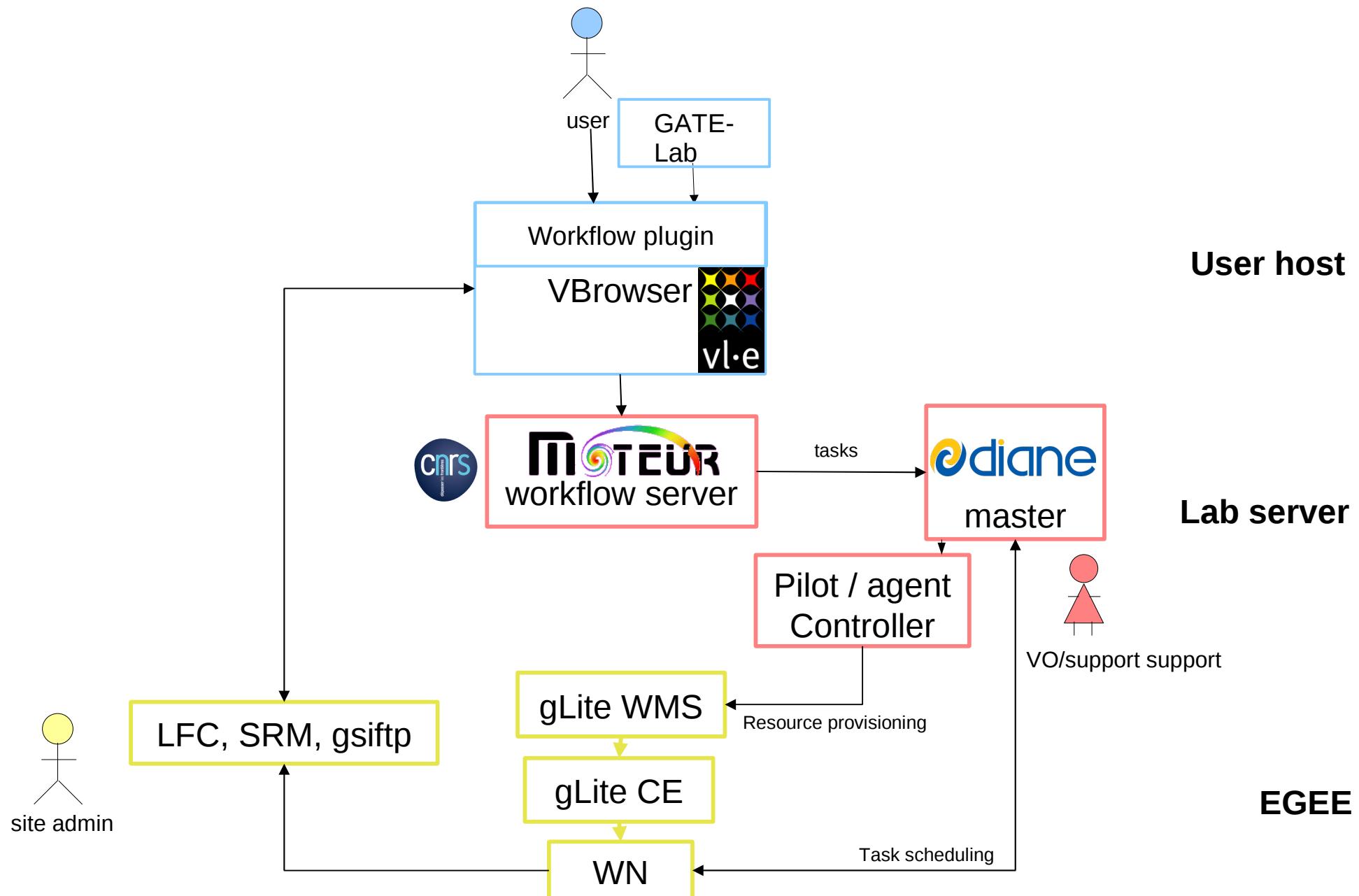
**Workflow inputs**

Load from file Save to file Add Parameter List Add Parameter Range Add Parameter Tag Path Delete selected

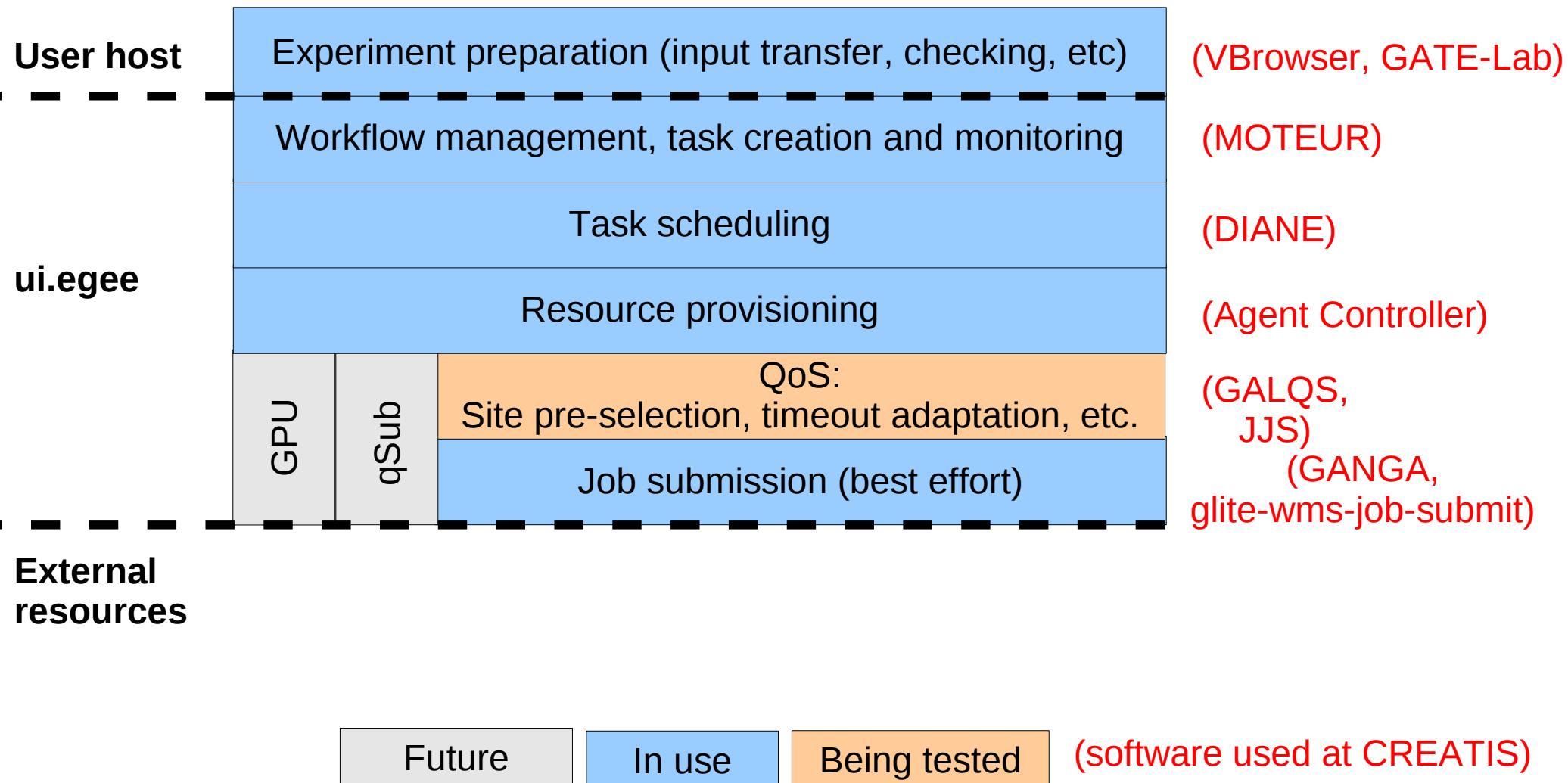
Name: indivAnalysis	Group	Value: Ifn://lfc.grid.sara.nl/grid/vlemed/remi/output/feat-dofhigh-12
Name: roi	Group	Value: Ifn://lfc.grid.sara.nl/grid/vlemed/remi/amygdLR_bin.nii.gz

Run Workflow Web service URL: <https://ws1.grid.sara.nl/~glatard/workflow/workflow-PTlv75/workfl>

# Grid set up used at Creatis



- **Applications described as workflows**
  - Parallel language
  - Middleware-independent
  - Provides structure to data (provenance logs)
- **Codes installed on the fly on the grid nodes**
  - Dependencies (e.g. libs) bundled in tgz
  - Only assumes that grid clients are installed
- **Case of matlab applications**
  - Compiled with toolboxes on a representative machine (license)
  - Deployed with Matlab Compiler Runtime (no license)



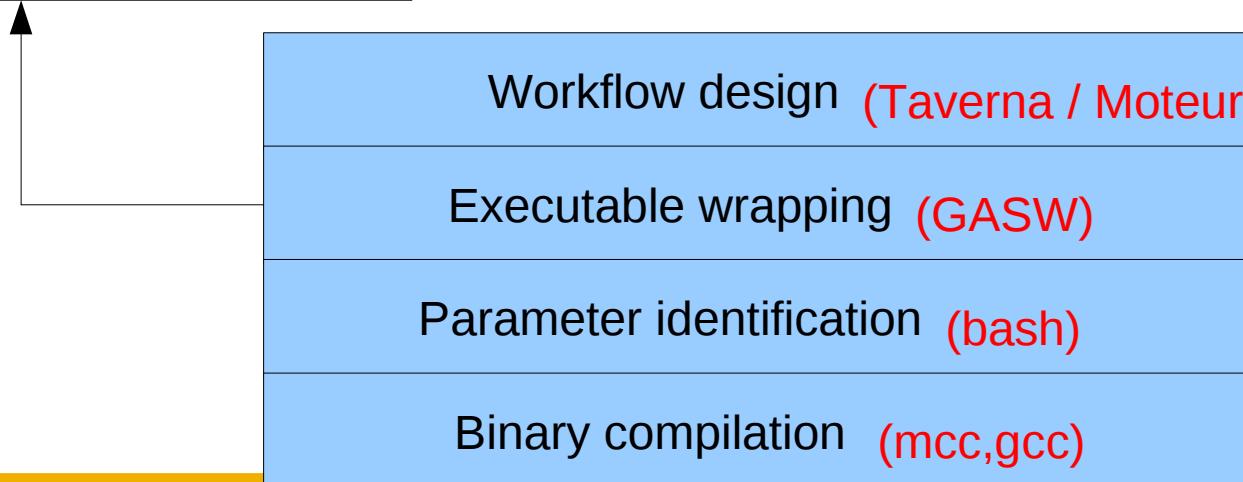
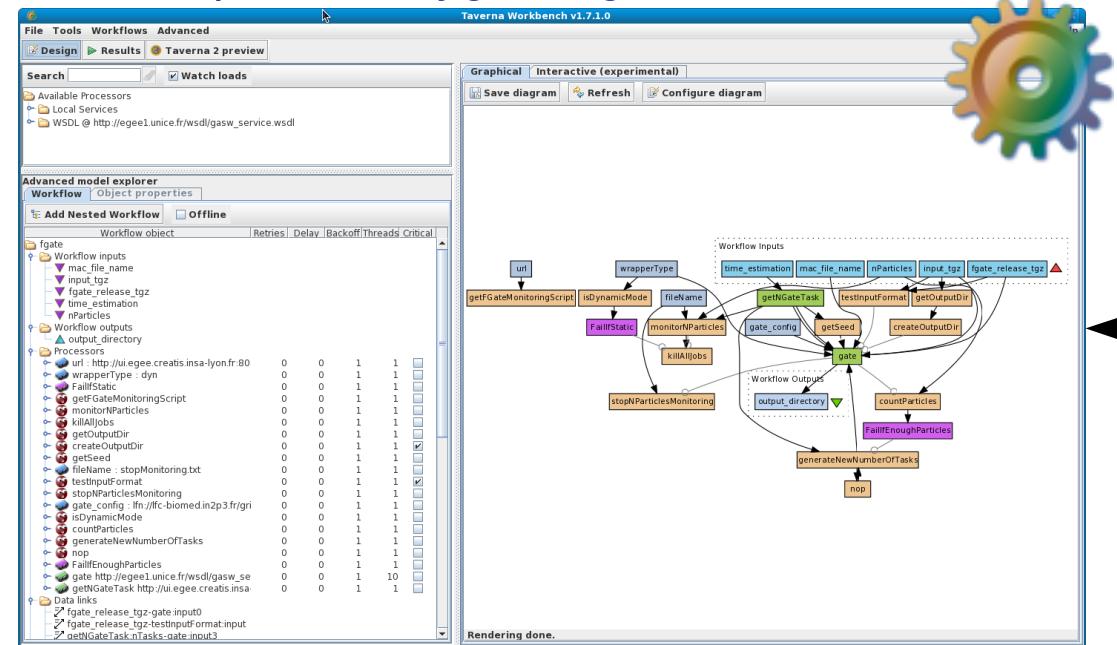
# Application porting stack

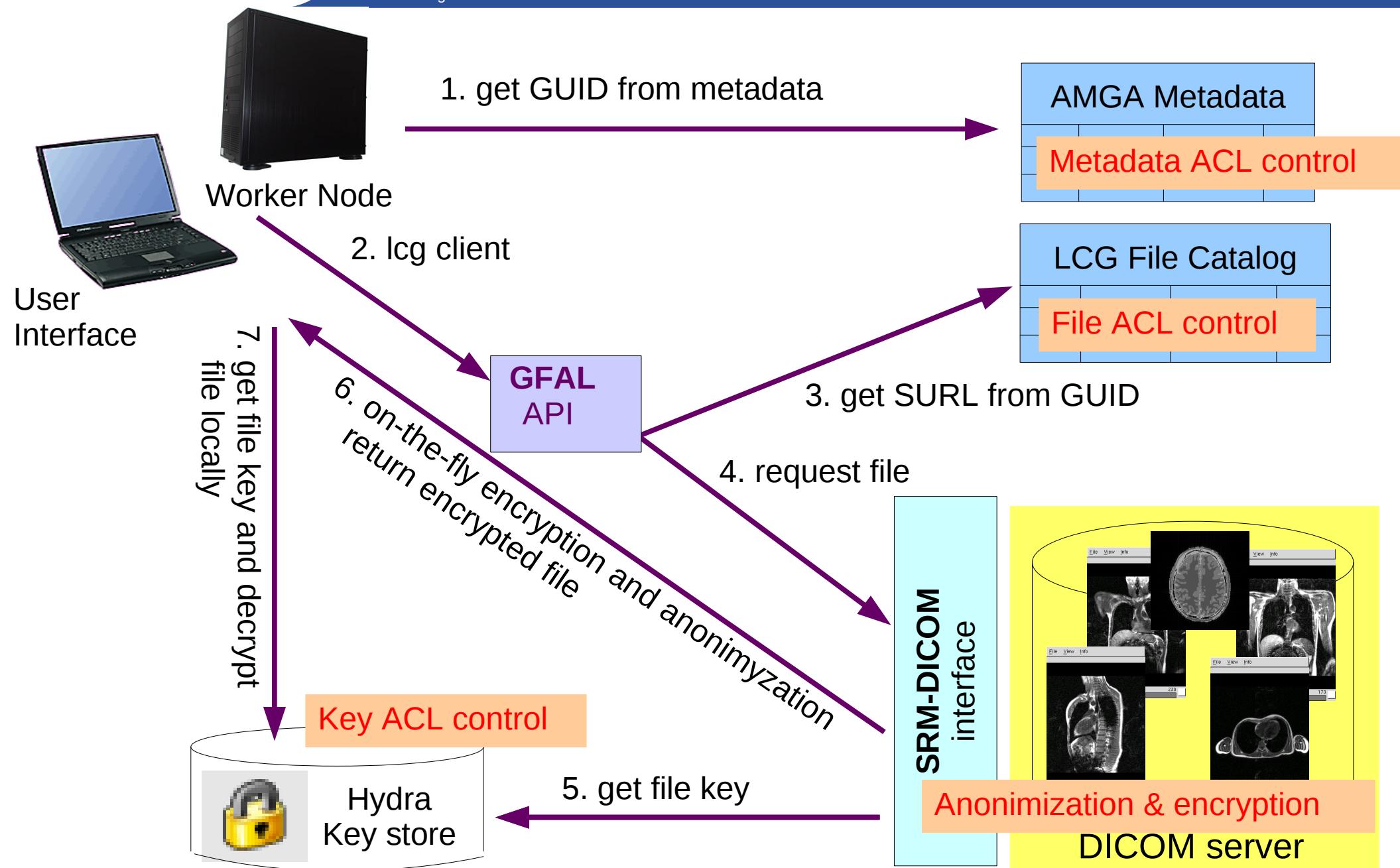
```

<description>
    <executable name="CrestLines.pl">
        <access type="URL">
            <path value="http://colors.unice.fr:80/">
        </access>
        <value value="CrestLines.pl"/>
        <input name="image" option="-im1">
            <access type="LFN" />
        </input>
        <input name="scale" option="-s"/>
        <output name="crest_lines" option="-c2">
            <access type="LFN" />
        </output>
        <sandbox name="convert8bits">
            <access type="URL">
                <path value="http://colors.unice.fr:80/">
            </access>
            <value value="Convert8bits.pl"/>
        </sandbox>
    </executable>
</description>

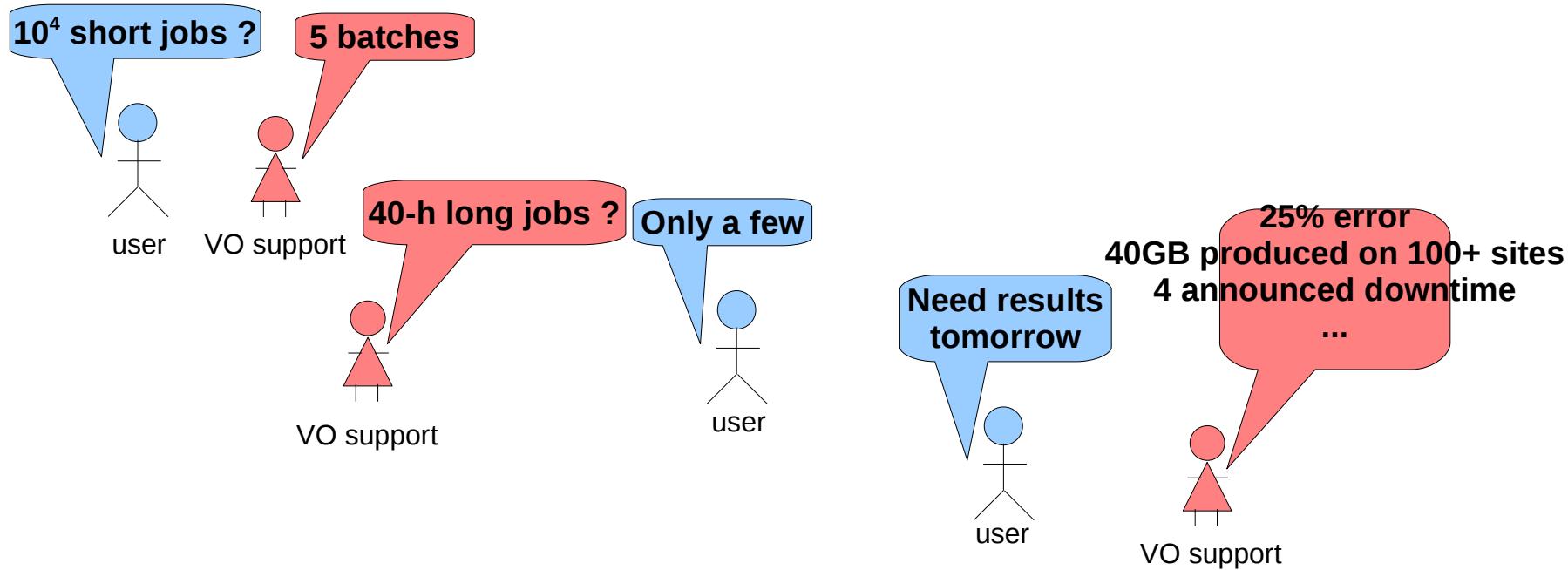
```

<http://www.mygrid.org.uk/tools/taverna/>





# Experiment planning

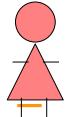


## => experiment planning effort

- plan experiment with user
- react to operational issues
- provide time-to-result estimates

# Data management issues

- **experiment level**

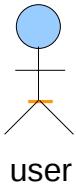


“Jobs are failing due to file transfer issues”

VO/user support

=> Data placement and replication of volatile data ?

- **user level**

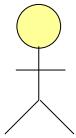


“The file I have uploaded to the grid in December is not available”

user

=> Data placement and replication of permanent data ?

- **VO level**



site admin

“Storage Element is full / being decommissioned ; please organize migration”

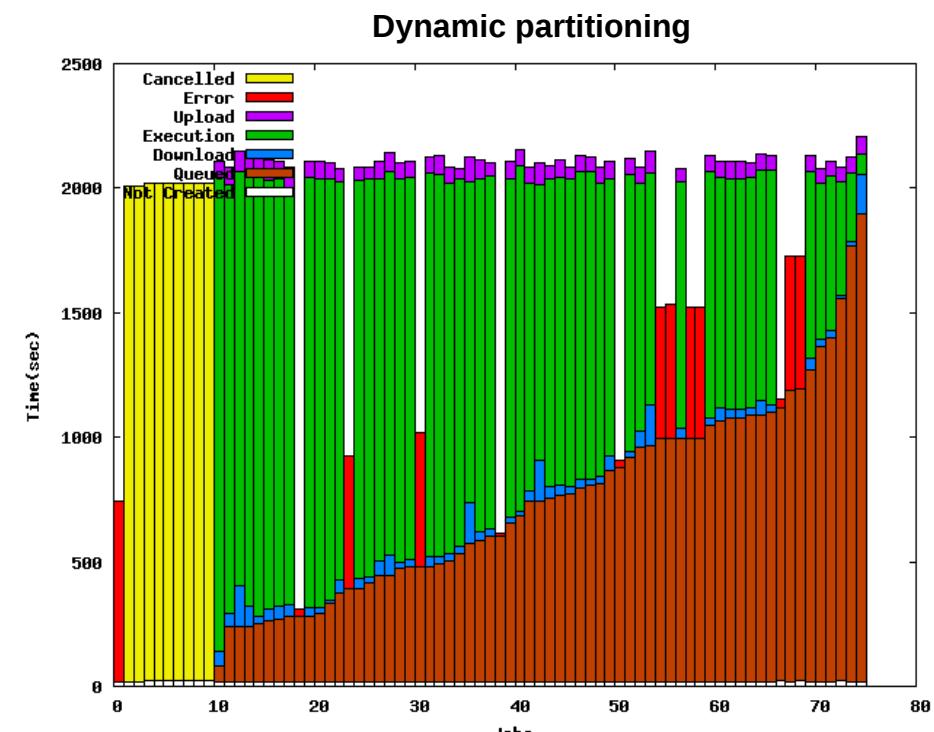
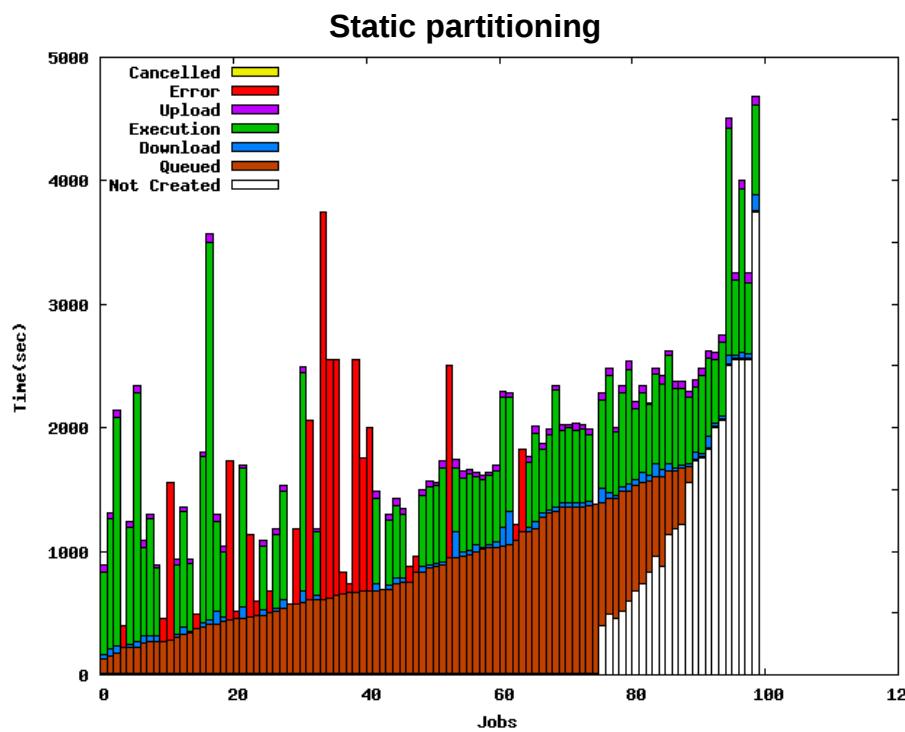
=> Data placement and replication of VO data ?

# Robust task scheduling



“99% of your experiment has completed ; the last 3 jobs will be available in 12 hours”

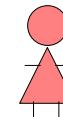
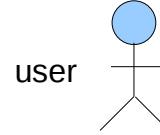
- **adjustments for GATE simulations**



# Reliability

- **Job error handling**

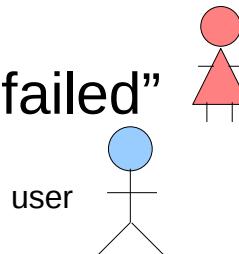
- “I made a typo my file name”
  - “200 jobs were resubmitted 10 times”
  - => detect permanent VS temporary errors ; local VS grid-wide errors ; user VS system errors



VO/user support

- **Recovery**

- “only two of your 5,000 jobs failed”
  - “which ones ?”

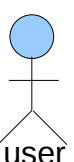
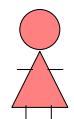


VO/user support

- **(week-)long experiments**

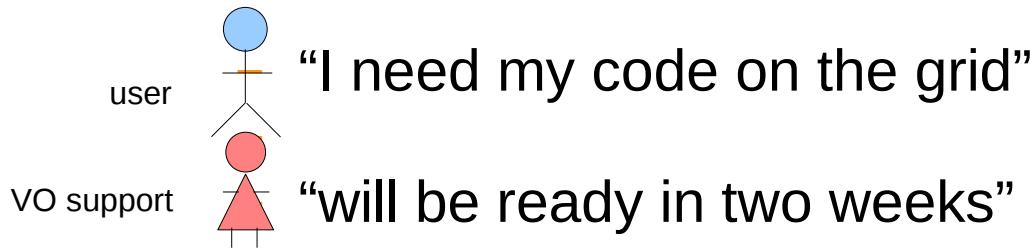
- “The server hosting your experiment had to reboot: your workflow has been killed”
  - “It had been running for 3 days and was 53% complete...”
  -

VO/user support



user

# Application porting



- **users are not autonomous with workflow programming**

=> an IDE for grid workflows ?

- Design assistance
- Workflow versioning
- Prototyping, debugging, testing environment
- Benchmarking / profiling tool

# Resource monitoring



“Monitor 100 Storage Elements => 1.3 ticket / day”  
“Monitor 10,000 SE/CE pairs ?”

- **Tools for error identification**
  - cron, Nagios, Hudson, etc
  - technical teams on shift
- **Error handling**
  - Ticketing systems
  - Mostly manual, hardly scalable

- **EGEE Life-Science cluster**
  - <https://twiki.cern.ch/twiki/bin/view/EGEE/LifeSciences>
  - (until the end of April)
- **biomed Virtual Organization**
  - <https://voms-biomed.in2p3.fr:8443/voms/biomed/>
  - [egee-biomed-vo-manager@healthgrid.org](mailto:egee-biomed-vo-manager@healthgrid.org)
- **Other Life-Science Vos**
  - embrace, enmr.eu, moldyngrid.org, vo.neugrid.vo, bio, gene, libi, tps.infn.it, vlemed, vo.renabi.fr, lsgrid, fkppi.kisti.re.kr, vo.iscpif.fr, vo.rhone-alpes.idgrilles.fr
  - <https://cic.gridops.org/index.php?section=vo>

# Credits

## FR National projects

Cardiac segmentation: Gwenda (2007-2010)  
Radiotherapy simulation: hGATE (2010-2012)  
Image simulation: VIP (2010-2012)



Institut national  
de la santé et de la recherche médicale



*Creatis*

## EGEE-III Life-Science cluster

### Creatis

Hugues Benoit-Cattin ; MRI simulation  
Sorina Camarasu-Pop  
Patrick Clarysse ; Cardiac segmentation  
Christopher Casta ; Cardiac segmentation  
Denis Friboulet ; US simulation

Carlos Gines Fuster ; FIELD grid porting, monitoring tools  
Carole Lartizien ; PET & CT simulation  
Thomas Grenier ; Mean-Shift optimization  
Ting Li ; Mean-Shift optimization  
Hervé Liebgott ; US simulation  
Simon Marache ; PET & CT simulation  
David Sarrut ; GATE simulation

### Cardiac workflow, MOTEUR2

Johan Montagnat ; CNRS I3S  
Ketan Maheshwari ; CNRS I3S  
Benjamin Isnard ; INRIA LIP

### VL-e medical software

Silvia D. Olabarriaga ; AMC Amsterdam  
Piter T. de Boer ; Universiteit Van Amsterdam  
Spiros Koulouzis ; Universiteit Van Amsterdam

### Pilot jobs (DIANE)

Jakub T. Moscicki ; CERN

### Grid support

glatard@creatis.insa-lyon.fr  
<https://gus.tz.k.de>