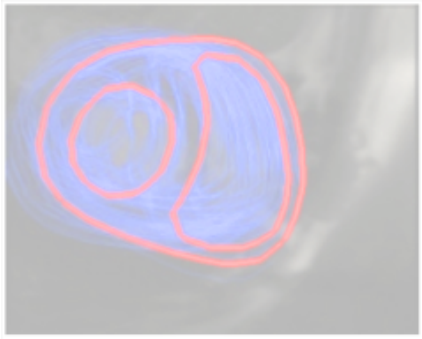
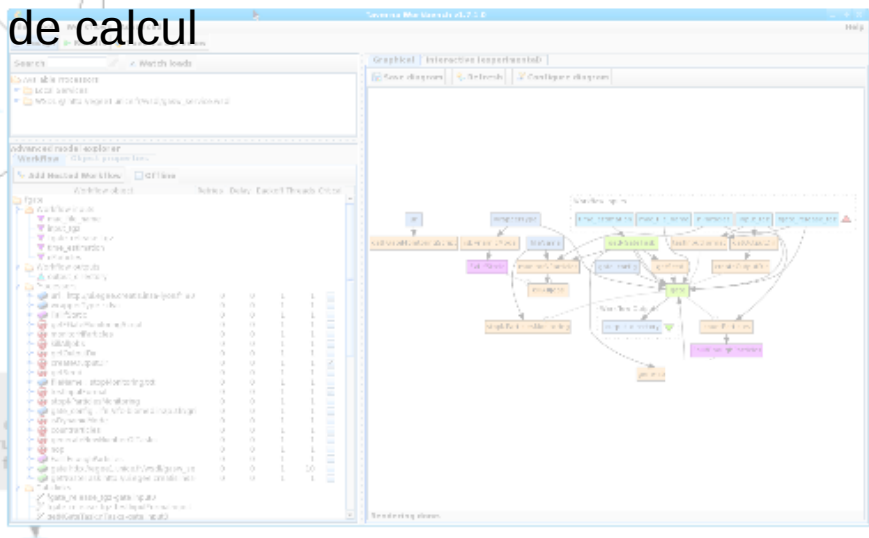


Expériences d'imagerie médicale sur EGEE

Tristan Glatard
CNRS – CREATIS, Lyon

Journée d'information grille de calcul
CPPM Luminy
Marseille
27 avril 2010



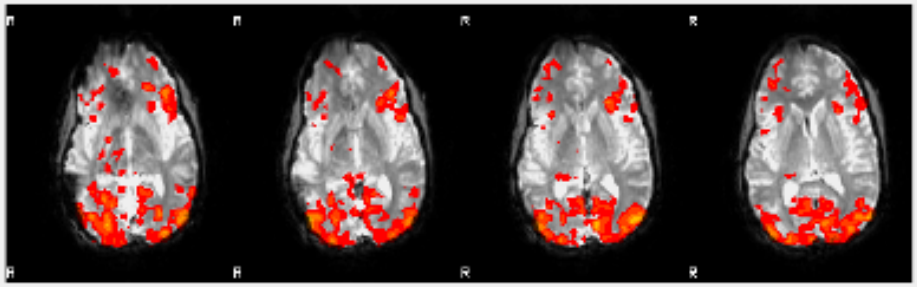
EGEE medical imaging activities

- **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 VOs**
  - fkppl.kisti.re.kr, vo.iscpif.fr, vo.rhone-alpes.idgrilles.fr

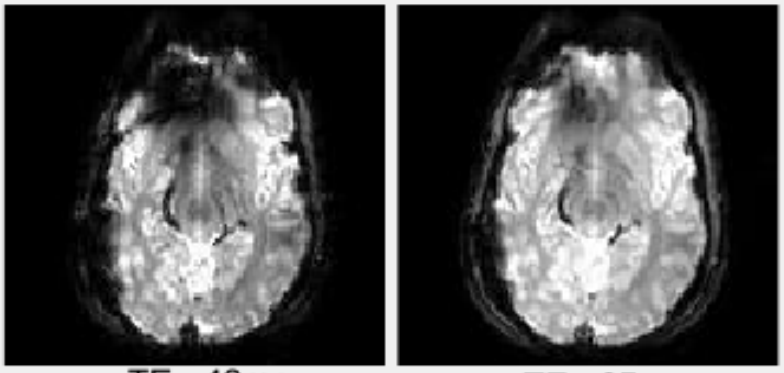
[miccai-G'08]

Parameter sweep in fMRI

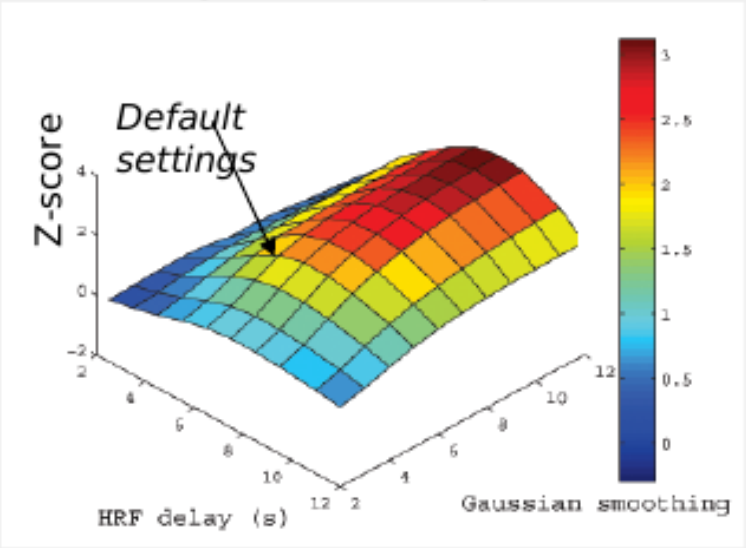
- Activation maps



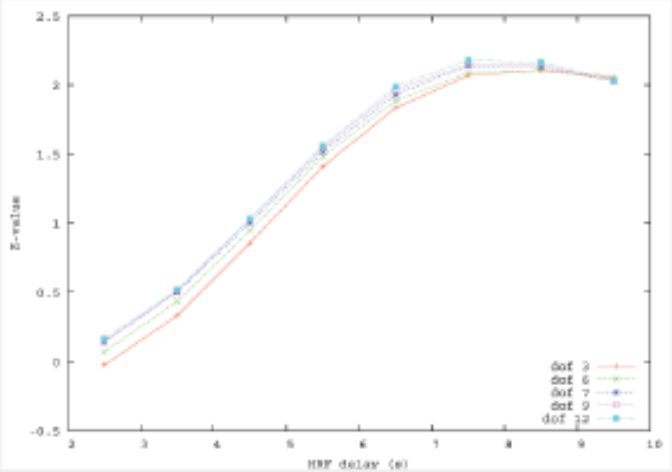
- Echo times comparison



- Smoothing and HRF optimization



- Registration evaluation



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

With R. Soleman and A. Nederveen

vl·e

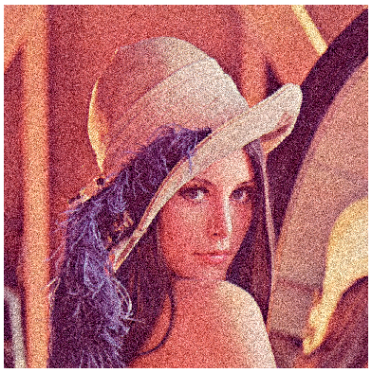
Grid execution

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

	$\#P_i$	#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	0.00
batch 2	11	1	6	5	8	2640	89.8	0.38	11.6	186.6	2642	0.08
batch 3	11	1	6	5	8	2640	89.8	0.38	32	67.38	2687	1.75
batch 4	11	1	5	5	8	2200	74.9	0.31	10.2	176.8	2203	0.14
total	11	2	11	5	8	9680	329.4	1.38	68.7	115	9732	0.53
Group Analyses								(MB)				
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								(MB)				
batch 1			11	5	8	440	7	23.8	44.3	3.8	2650	83.40

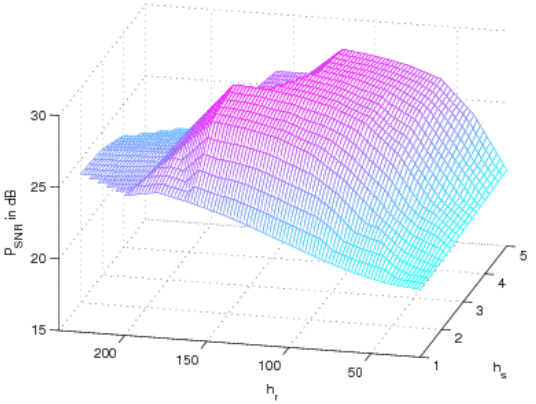
Mean-Shift filtering optimization

- **Sweep on scale parameters of MS filter**



Noisy image

Restored image



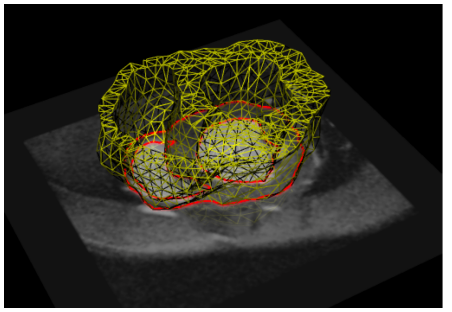
PSNR w.r.t scale parameters

- **Grid facts (Matlab code)**

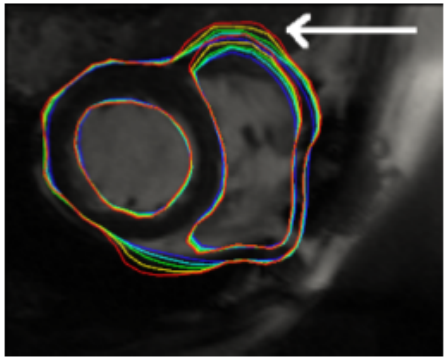
h_s value	Total CPU time	Elapsed time	Speed-up	Total data transfer time	Produced data	Successful tasks	Total tasks	Error ratio
1	13.0 days	3h25min	92	89.8h	43GB	8,000	8,106	1.3%
2	50.4 days	18h36min	65	63.2h	41GB	8,000	8,929	10.4%
3	17.3 days	13h01min	32	34.8h	5GB	1,000	1,317	24%
4	29.0 days	13h23min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

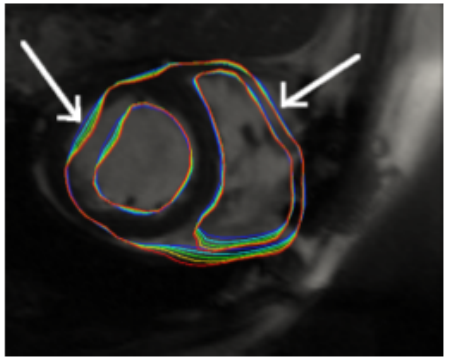
- **Cardiac segmentation with deformable models**
  - Volumetric template mesh
  - Initialized with rigid registration
  - Image gradient ► force field ► deformation



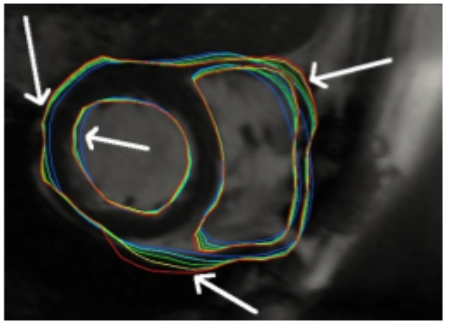
- **Adjust segmentation parameters**



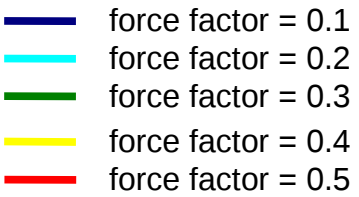
init 0



init 1



init 2

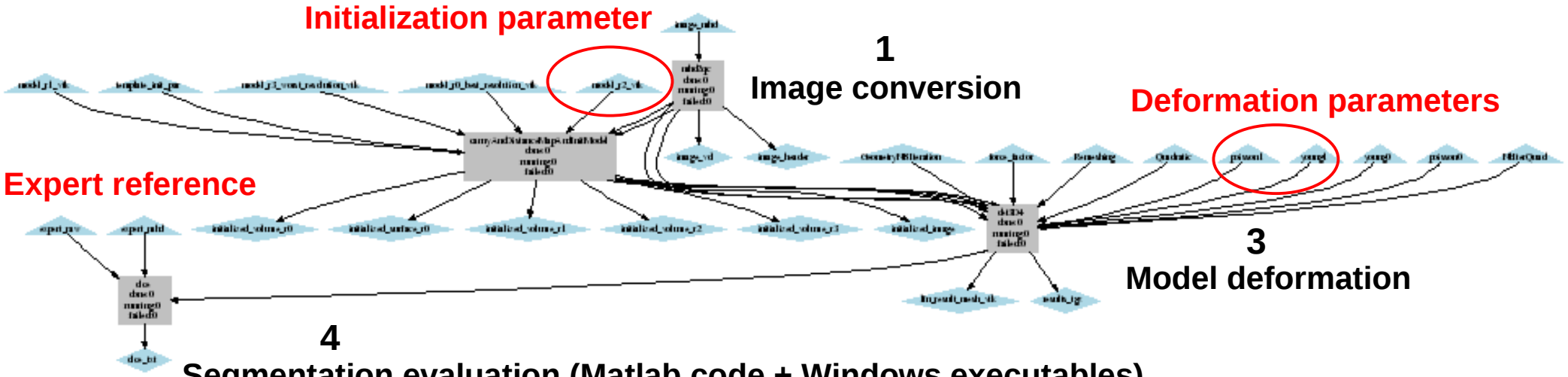


- **Estimate myocardium physical parameters**
  - Best segmentation ◄ most realistic parameters (e.g. Young modulus)

Creatis

Cardiac segmentation workflow

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

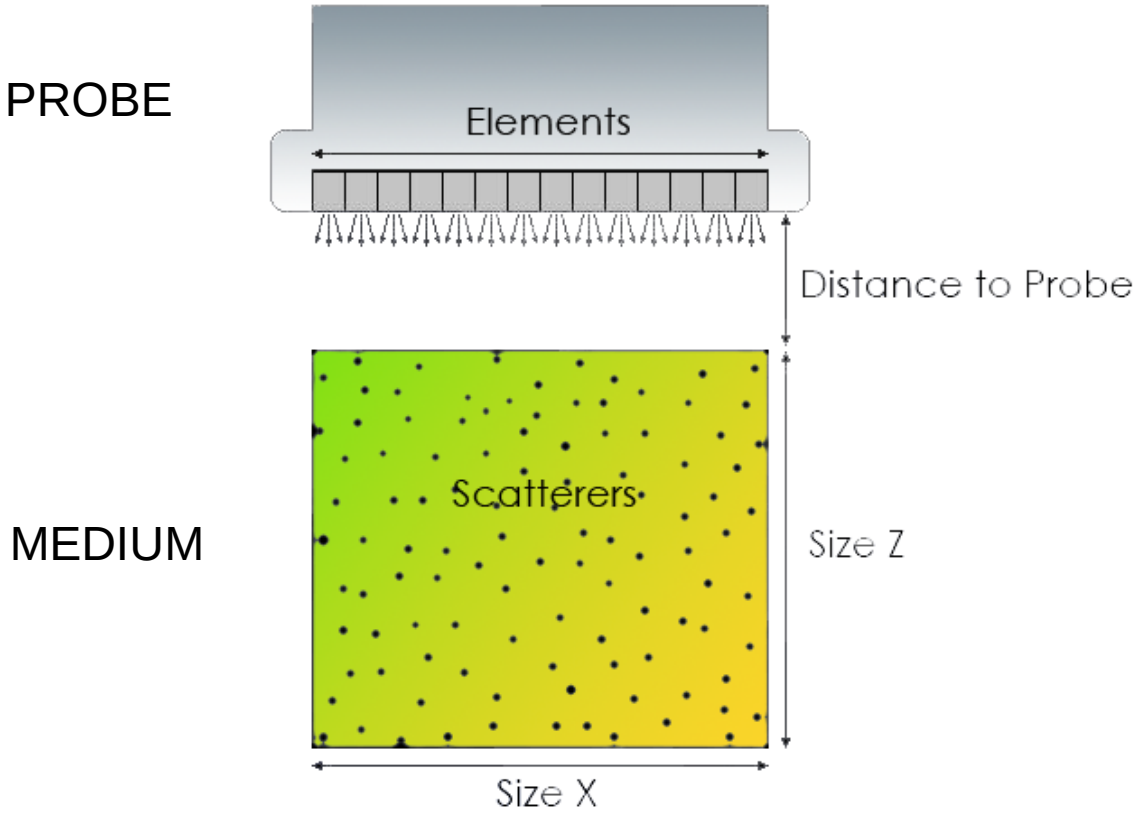


- **Workflow iterated on parameter/image sets**

FIELD US simulation

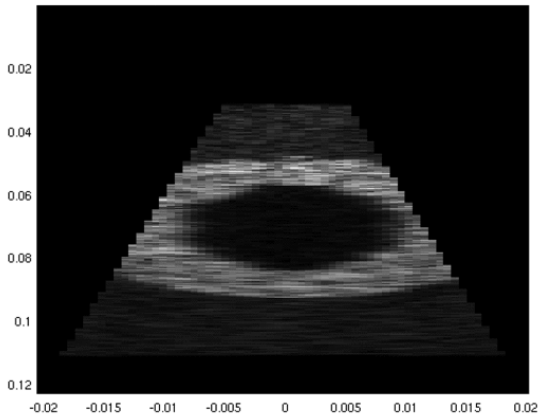
http://server.oersted.dtu.dk/personal/jaj/field/

- **Principle**



  - Parallelism on lines
  - Parallelism on mediums

- **Example on 2D beating heart**

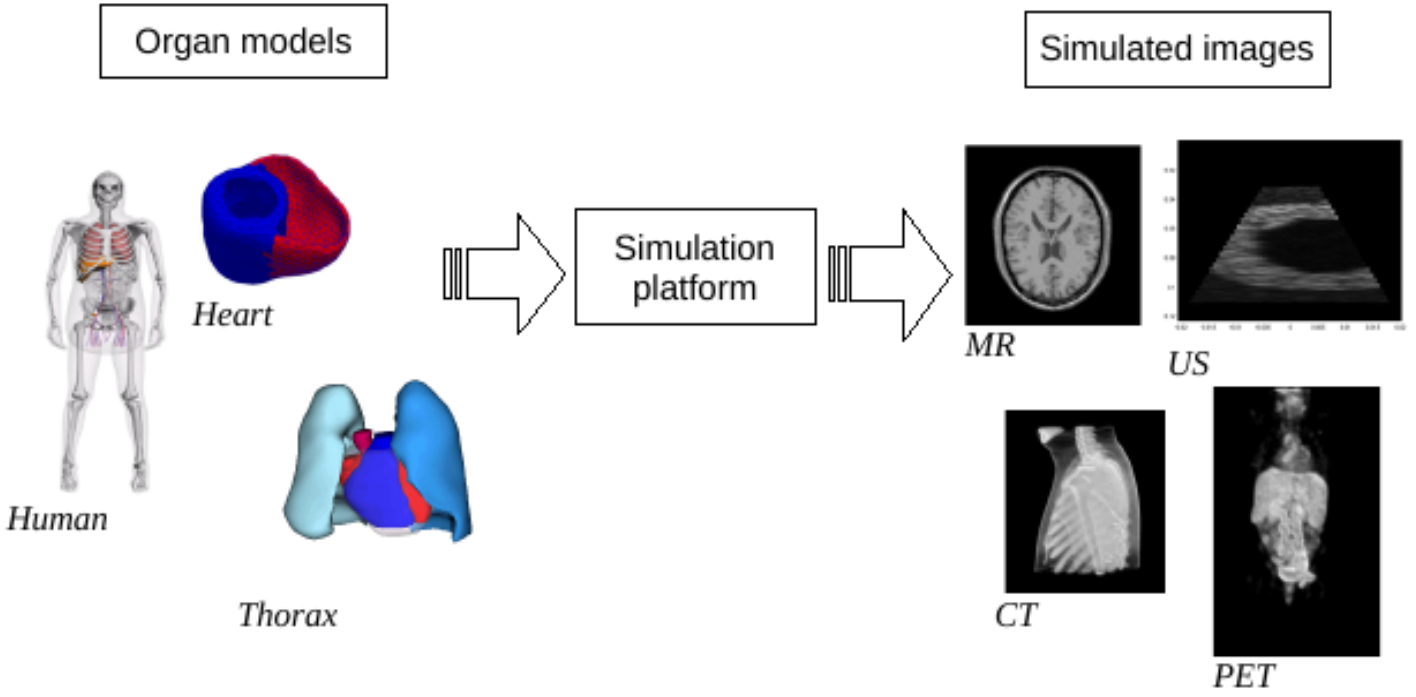


  - 1920 lines (30 mediums x 64 lines)
  - > 16h => < 3h
  - 12% error (first try) => 2% (fine-tuning)

Virtual Imaging Platform

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

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

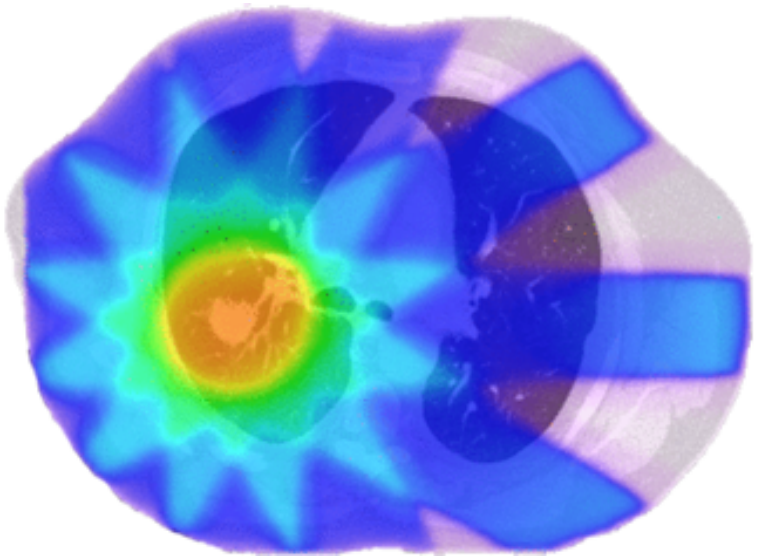


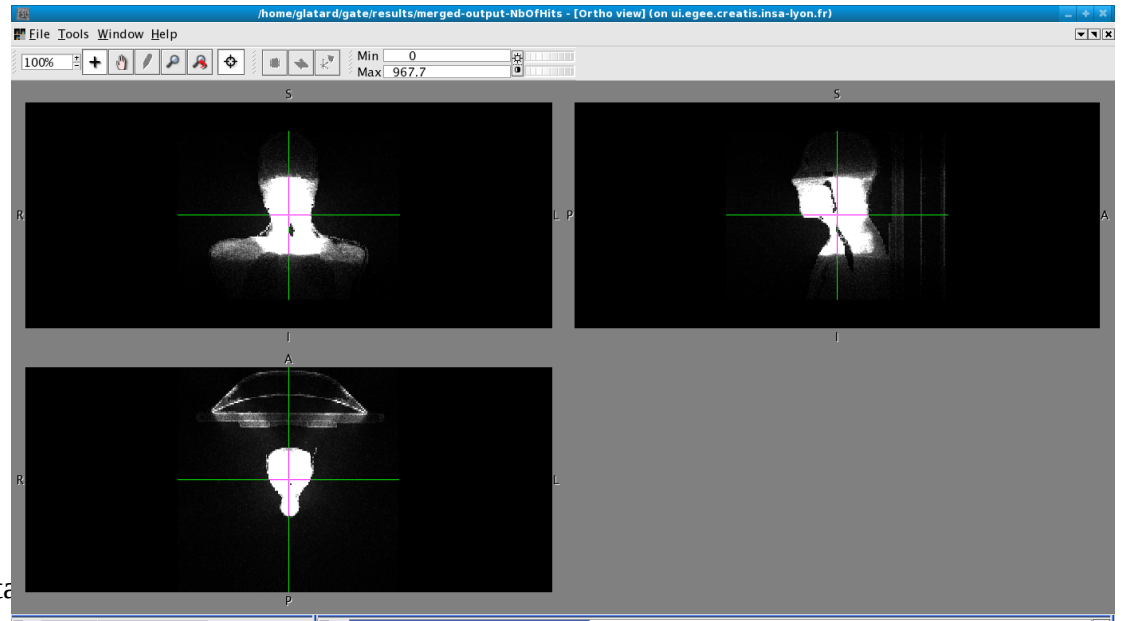
Hadrontherapy simulation with GATE

http://opengatecollaboration.healthgrid.org/

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

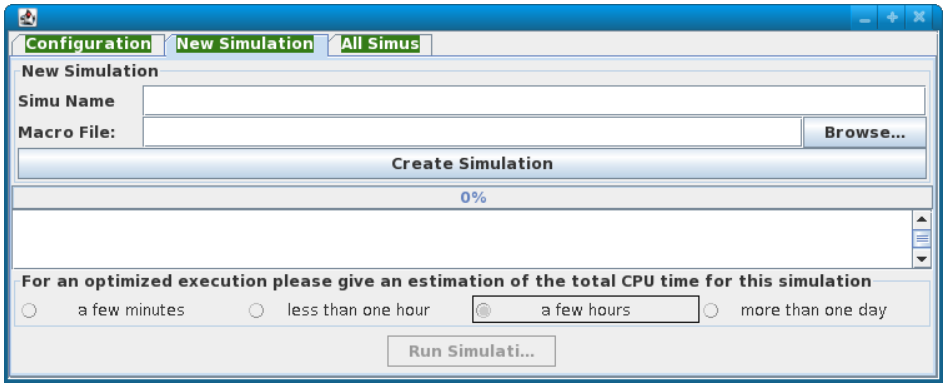
- **Involved effort**
  - High number of particles (~ 20,000,000)
  - Monte-Carlo simulation ►divisible load



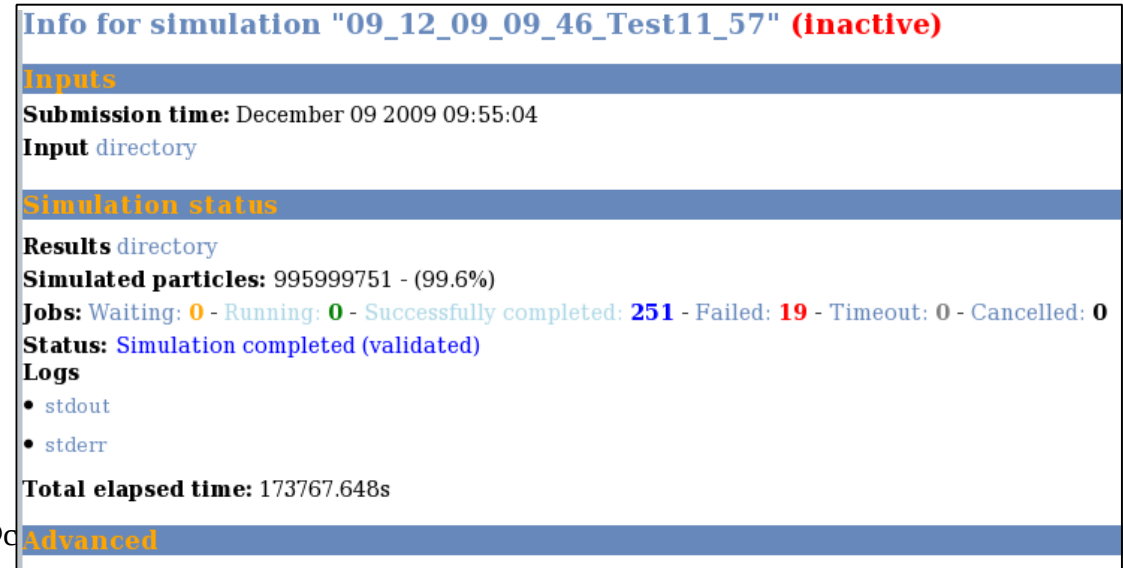


GATE-Lab interface

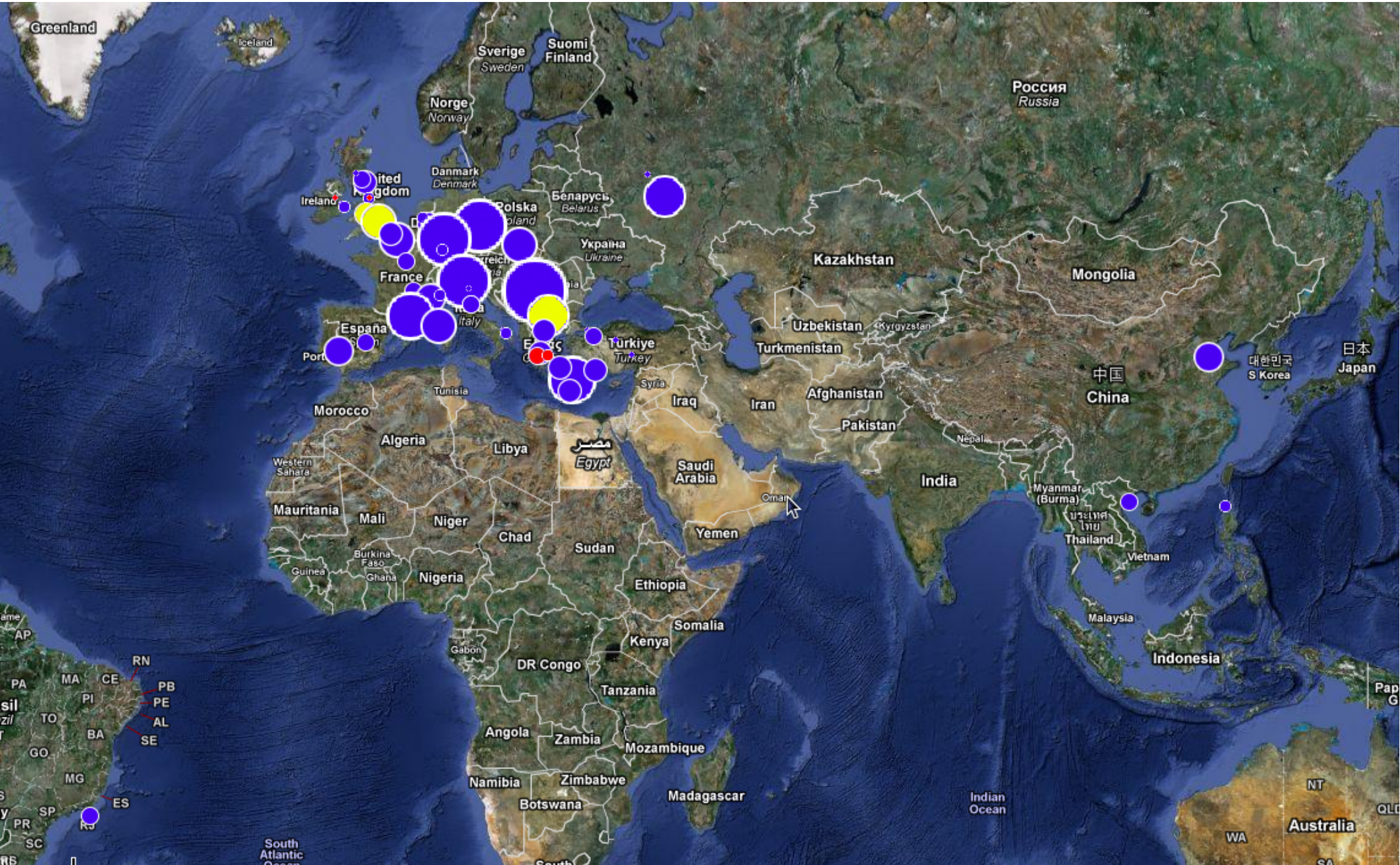
- **Dedicated VBrowser plugin**
  - Simulation file parsing
  - Parameter checking
  - Input files bundling and upload
  - Time estimation
  - History management (+cleanup)

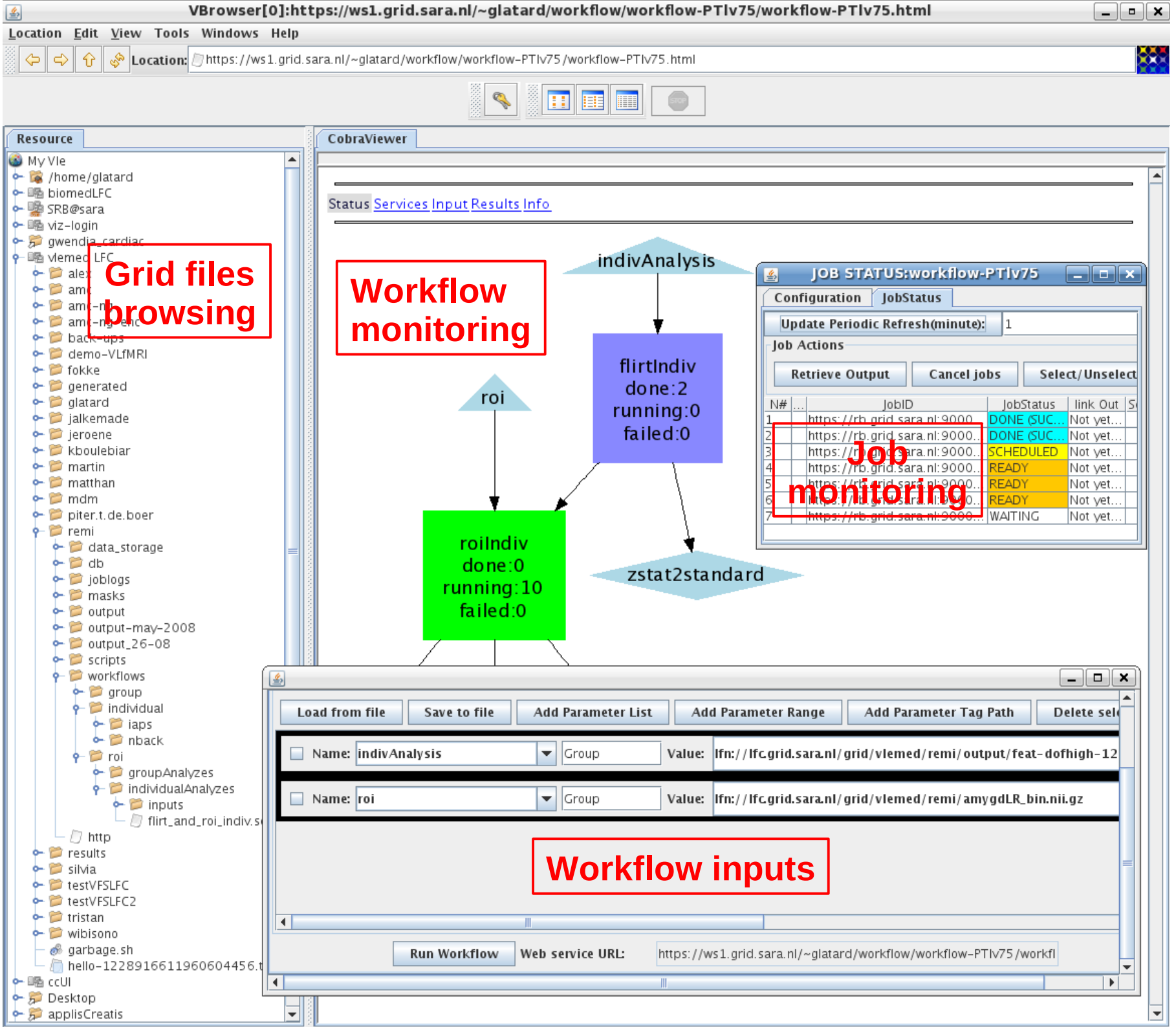


- **Customized server interface**
  - #simulated particles
  - Current status
  - Link to results directory
  - Confirm/retry

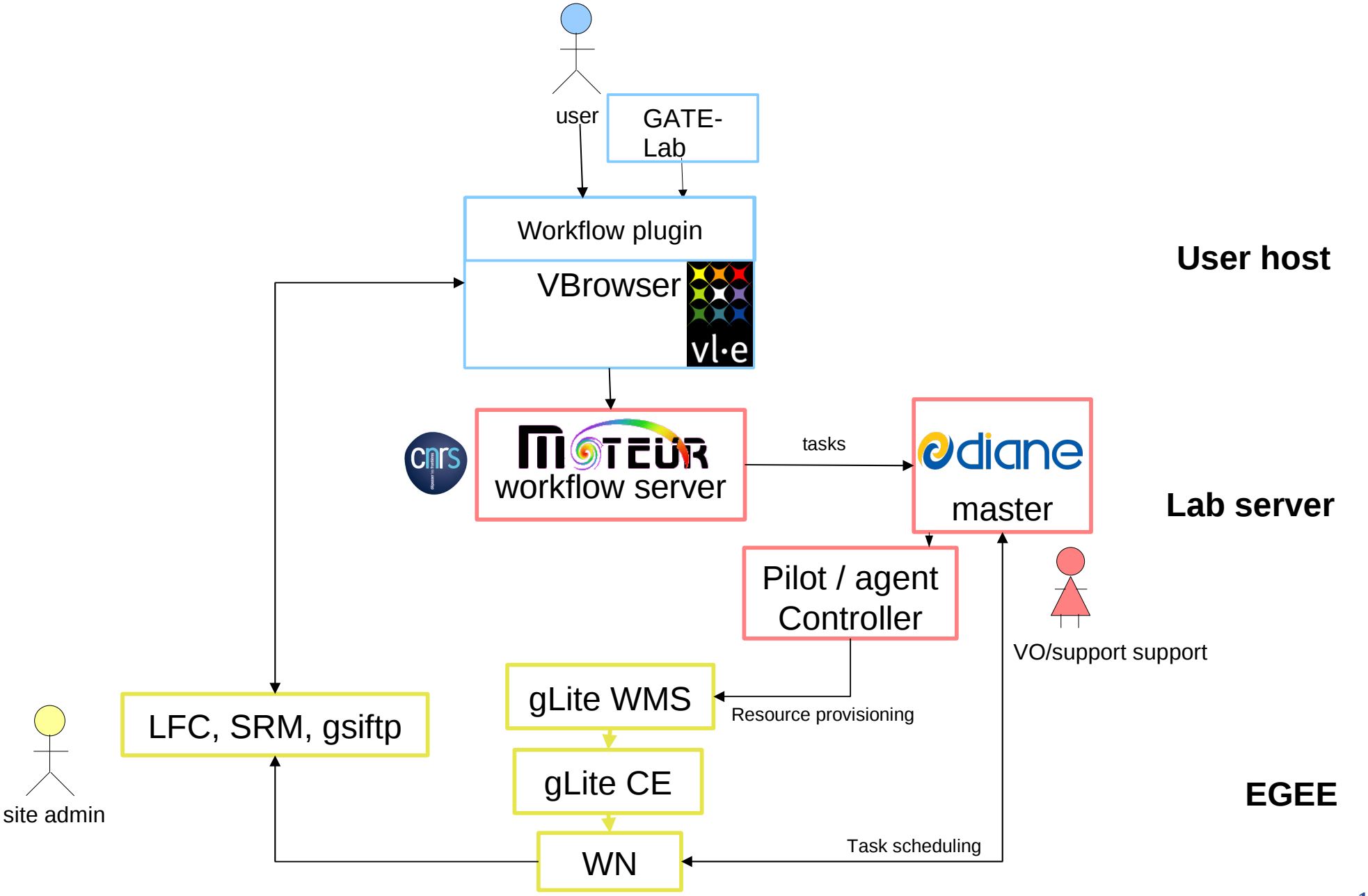


A (hadrontherapy) experiment




VBrowser[0]:https://ws1.grid.sara.nl/~glatard/workflow/workflow-PTlv75/workflow-PTlv75.html
Location Edit View Tools Windows Help
Location: https://ws1.grid.sara.nl/~glatard/workflow/workflow-PTlv75/workflow-PTlv75.html
Resource
CobraViewer
Status Services Input Results Info
Grid files browsing
Workflow monitoring
indivAnalysis
flirtIndiv done:2 running:0 failed:0
roi
roiIndiv done:0 running:10 failed:0
zstat2standard
JOB STATUS:workflow-PTlv75
Configuration JobStatus
Update Periodic Refresh(minute): 1
Job Actions
Retrieve Output Cancel jobs Select/Unselect
Job monitoring
Load from file Save to file Add Parameter List Add Parameter Range Add Parameter Tag Path
Name: indivAnalysis Group Value: lfn://lfc.grid.sara.nl/grid/vlemed/remi/output/feat-dofhigh-12
Name: roi Group Value: lfn://lfc.grid.sara.nl/grid/vlemed/remi/amygdLR_bin.nii.gz
Workflow inputs
Run Workflow Web service URL: https://ws1.grid.sara.nl/~glatard/workflow/workflow-PTlv75/workfl

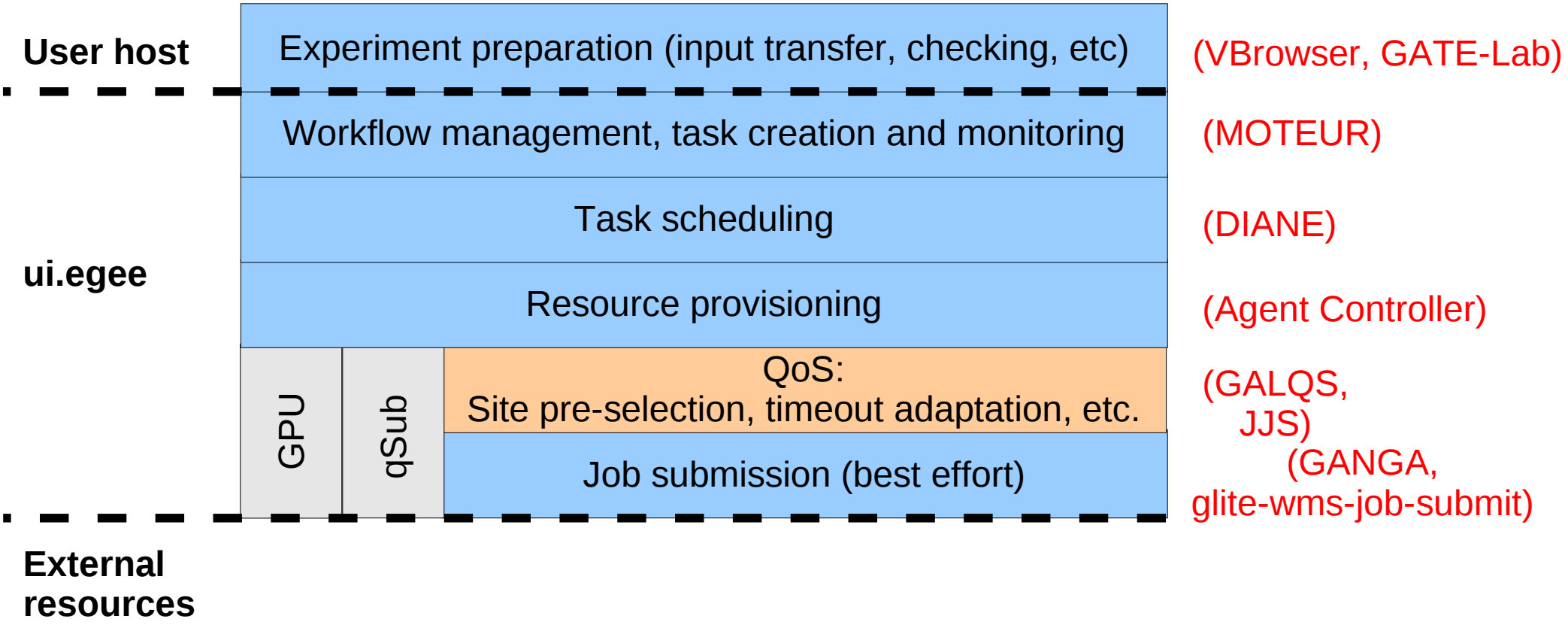
Grid set up used at Creatis



Application porting

- **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)

Execution stack



Medical Data Retrieval

Worker Node

User Interface

1. get GUID from metadata

AMGA Metadata

Metadata ACL control

2. lcg client

LCG File Catalog

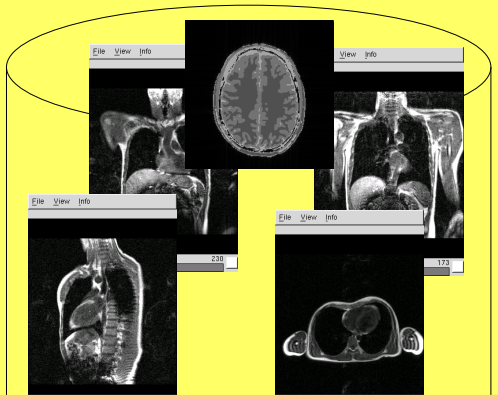
File ACL control

GFAL API

3. get SURL from GUID

4. request file

SRM-DICOM interface



Anonimization & encryption

DICOM server

6. on-the-fly encryption and anonimyzation return encrypted file

7. get file key and decrypt file locally

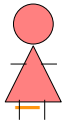
Key ACL control

Hydra Key store

5. get file key

Data management issues

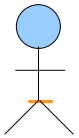
- **experiment level**



"Jobs are failing due to file transfer issues"
=> Data placement and replication of volatile data ?

VO/user support

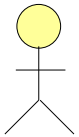
- **user level**



"The file I have uploaded to the grid in December is not available"
=> Data placement and replication of permanent data ?

user

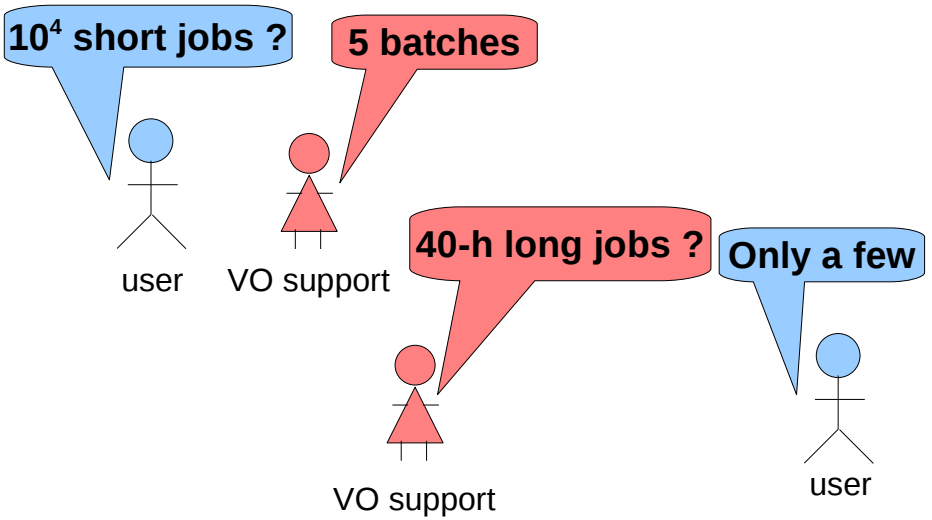
- **VO level**

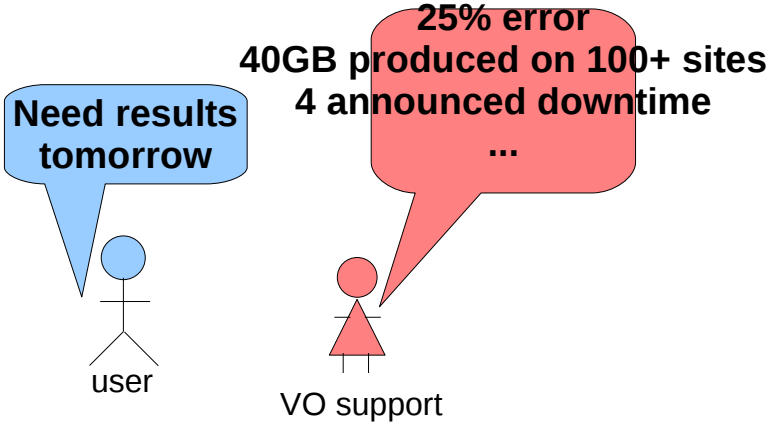


"Storage Element is full / being decommissioned ; please organize migration"
=> Data placement and replication of VO data ?

site admin

Experiment planning

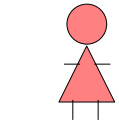




=> experiment planning effort

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

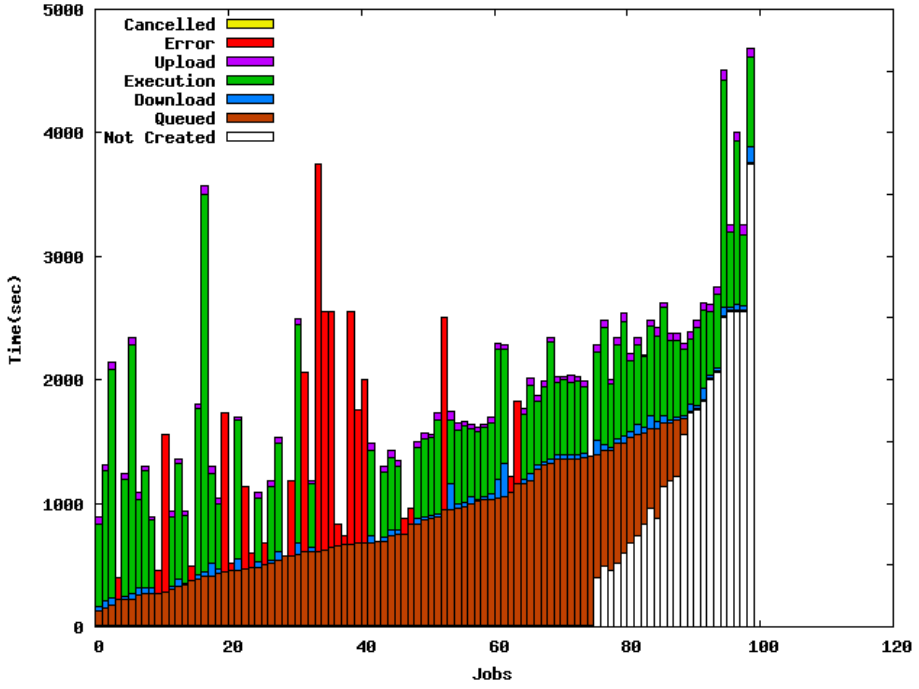
Robust task scheduling



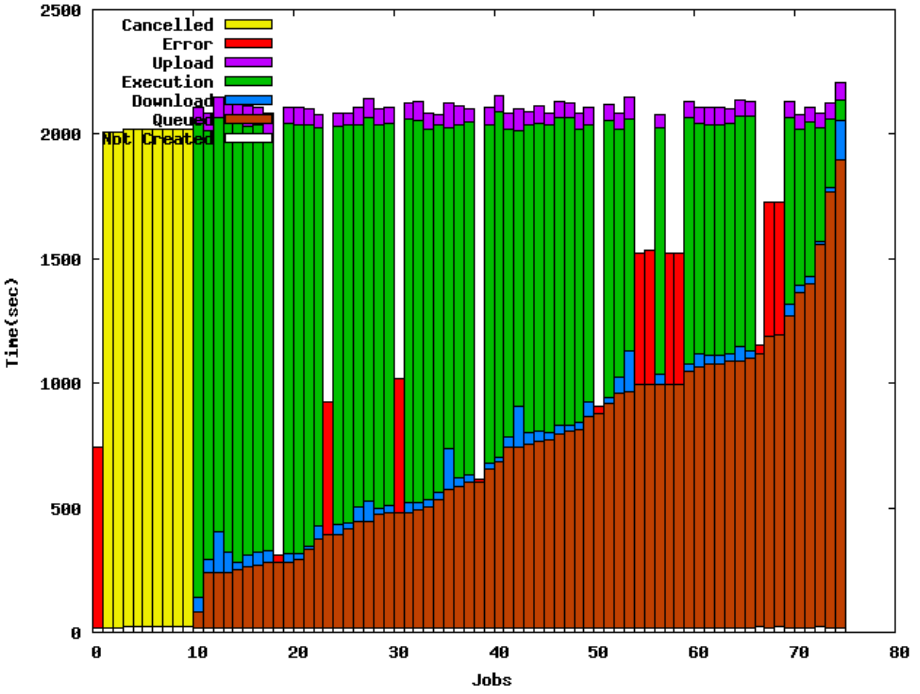
"99% of your experiment has completed ; the last 3 jobs will be available in 12 hours"

- **adjustements for GATE simulations**

Static partitioning

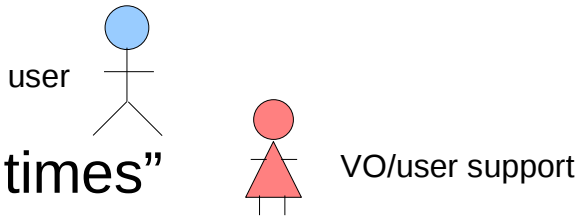


Dynamic partitioning



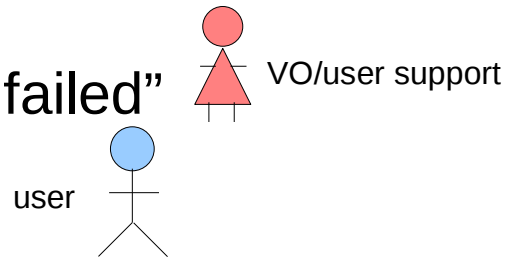
Reliability

- **Job error handling**
  - “I made a typo my file name”
  - “200 jobs were resubmitted 10 times”

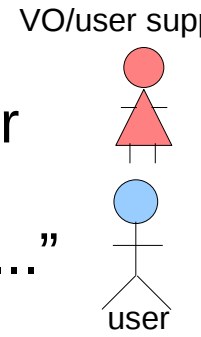


  - => detect permanent VS temporary errors ; local VS grid-wide errors ; user VS system errors

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



- **(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...”



More information and contact points

- **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

- **Other Life-Science Vos**
  - embrace, enmr.eu, moldyngrid.org, vo.neugrid.vo, bio, gene, libi, tps.infn.it, vlemed, vo.renabi.fr, lsgrid, fkppl.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: Gwendia (2007-2010)
Radiotherapy simulation: hGATE (2010-2012)
Image simulation: VIP (2010-2012)

EGEE-III Life-Science cluster

Creatis
Hugues Benoit-Cattin ; MRI simulation
Sorina Camarasu-Pop
Patrick Clarysse ; C*ardiac 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
https://gus.fzk.de