

Data caching in the Virtual Imaging Platform

Tristan Glatard¹

¹Creatis, CNRS, INSERM, Université de Lyon, France

on behalf of the VIP project consortium



EGI technical forum, September 2011

Virtual Imaging Platform

Enabling distributed computing for medical image simulation

- **Computed Tomography**

Sindbad

{CEA-Leti}

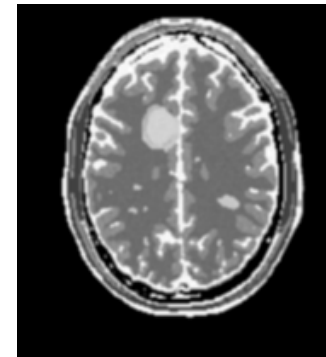


XCAT

- **Magnetic Resonance Imaging**

SIMRI

{CREATIS}



Brainweb + MS lesions + USPIO

- **Ultrasound imaging**

Field-II

{Technical Univ. Denmark}



Echocardiography

- **Positrons Emission Tomography**

PET-Sorteo

{CERMEP}



Zubal + tumors

Integration of simulators

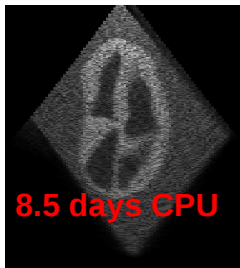
- **Workflow descriptions**

- Task graphs
- Data dependencies

- **Motivations**

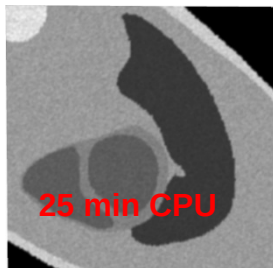
- Express data parallelism
- Automated processing of applications
- No modification of the simulator codes

- **Integrated 4 simulators**



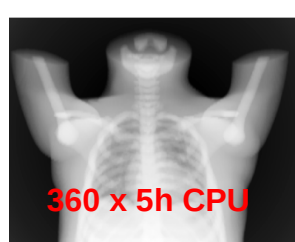
8.5 days CPU

US



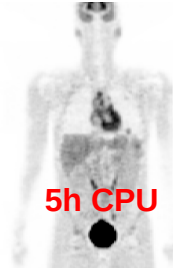
25 min CPU

MRI



360 x 5h CPU

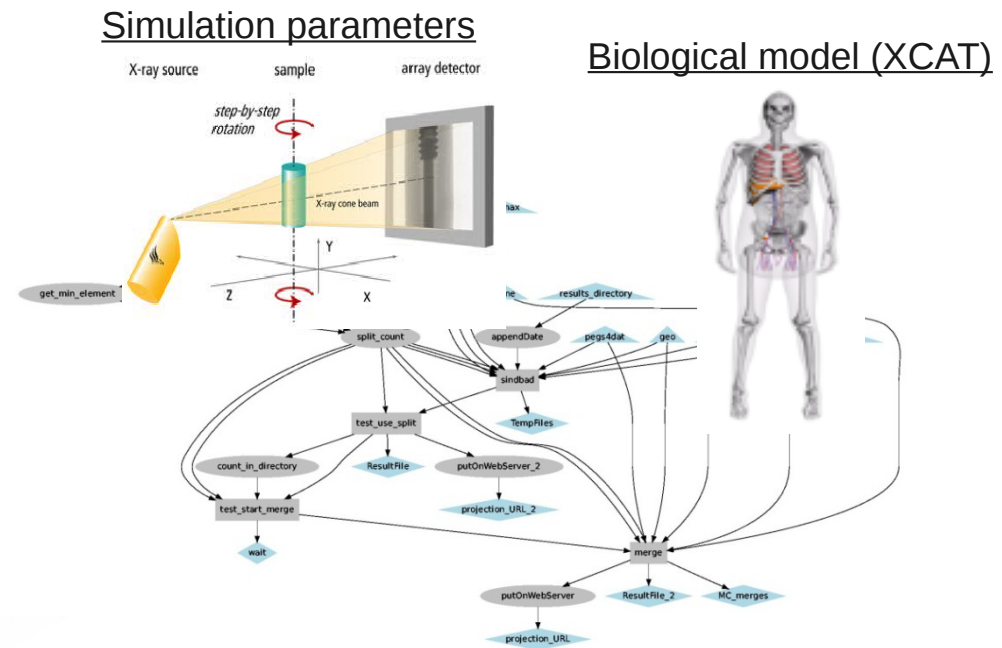
CT



5h CPU

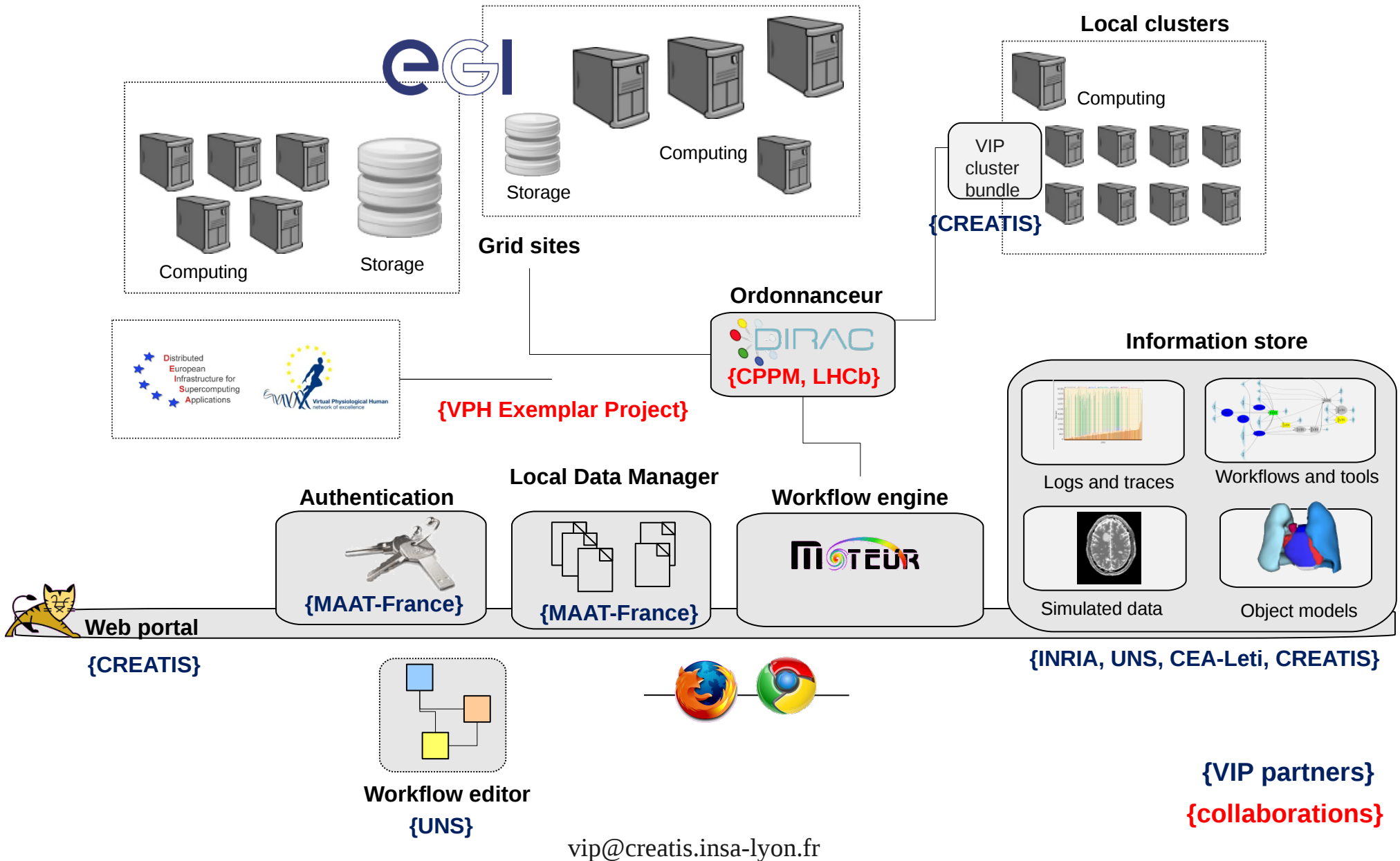
PET

Example for Sindbad - CT



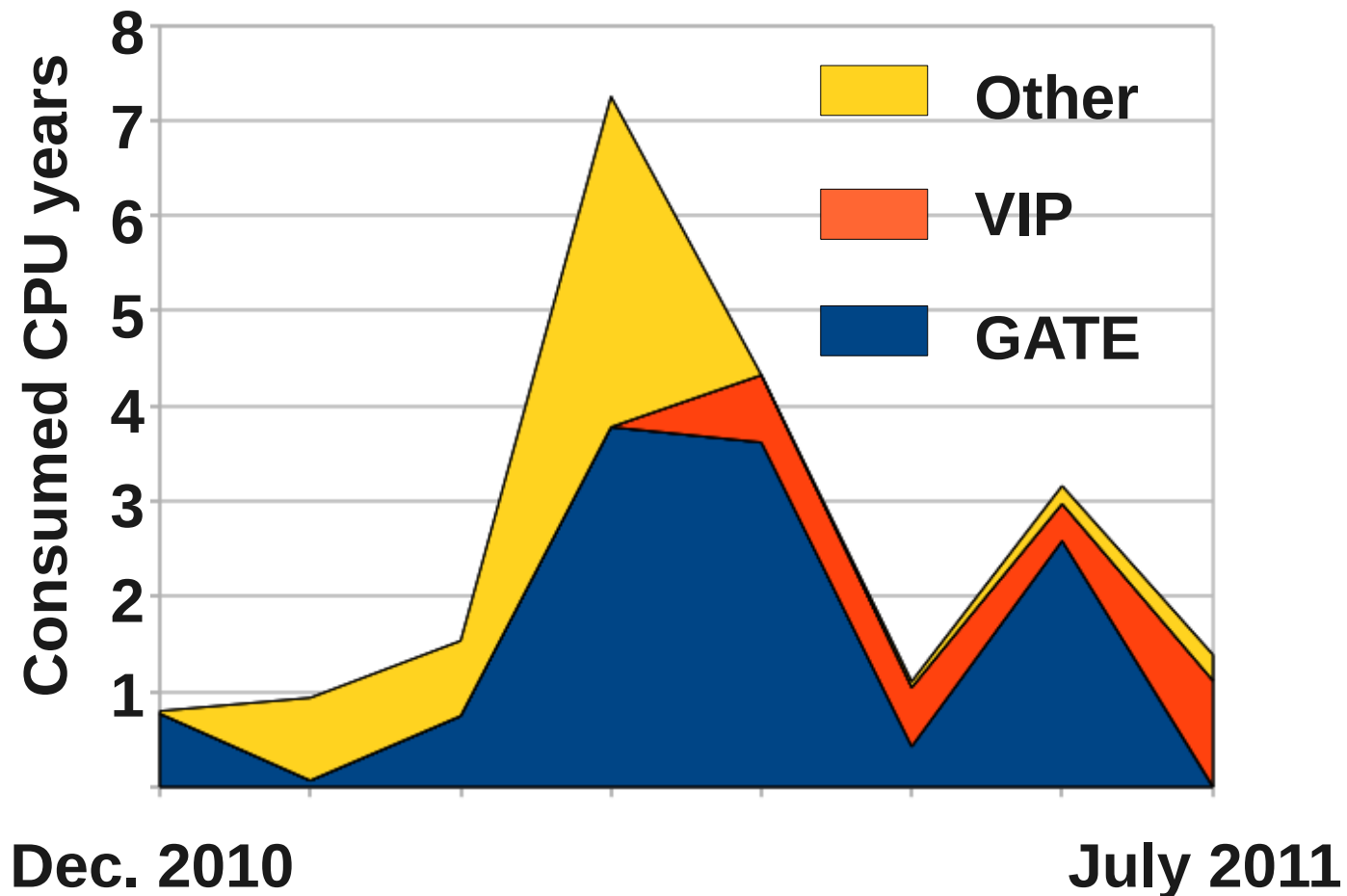
Simulated data

Platform outline



Platform usage

- Subset of LSGC, biomed VO



Data management problem

- **EGI three-tier data management**
 - Logical File Catalog (LFC) – single index space
 - Storage Elements (SE) – DPM, dCache, STORM, Castor
 - **Challenge: data availability between 80-95%**
- **Data management in VIP**
 - Users upload input files to process on LFC (web interface)
 - Platform replicates these files 3 times
 - Files are cached by (pilot) jobs
 - Output files are stored on site SE; central SEs as failovers
 - Job error rates due to data transfer issues: 5-10%

Local Data Manager

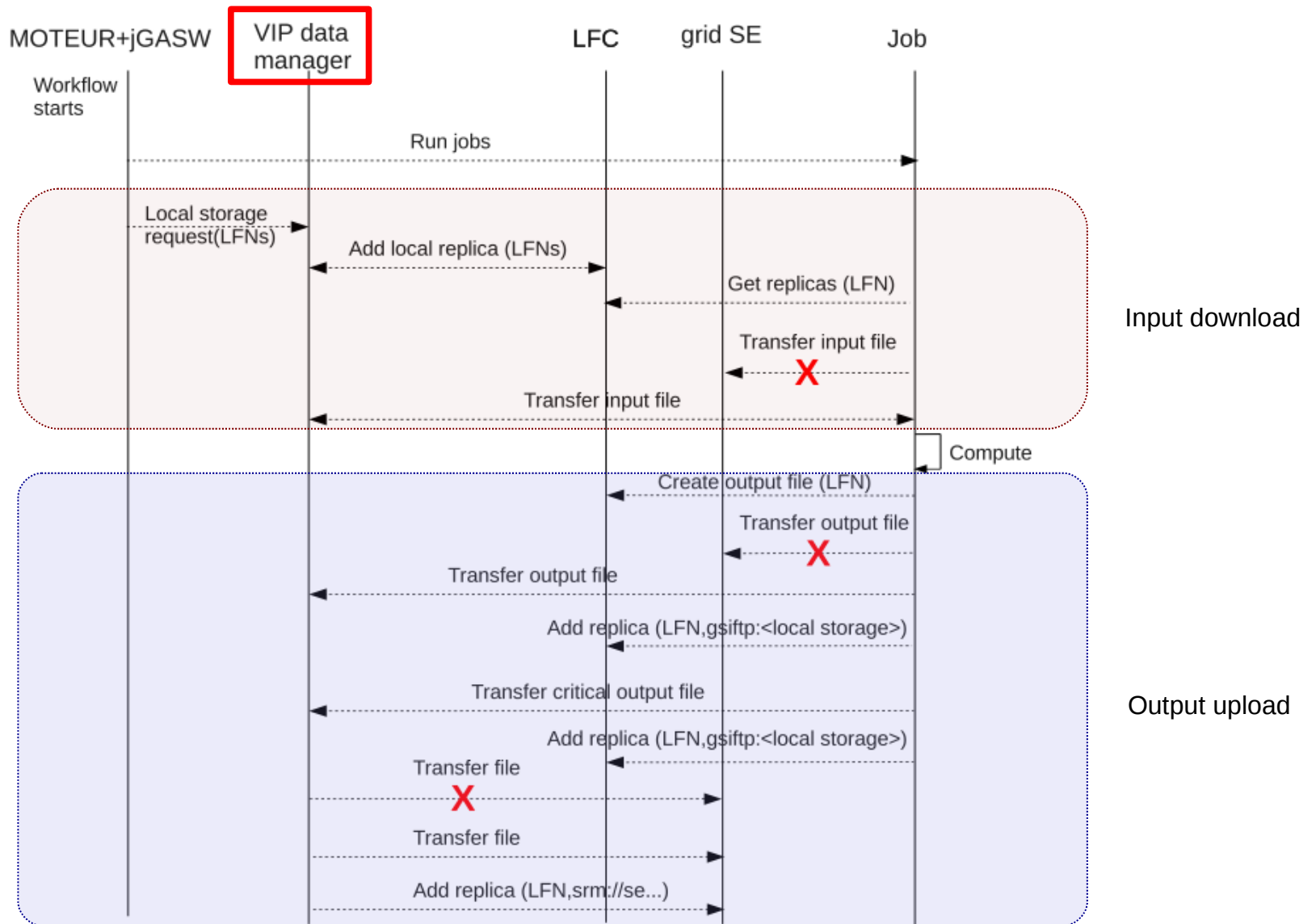
- **Principle**

- Dedicated cache SE, used at failover storage
- Periodically tries to replicate its files to grid SEs
- Available for users and grid jobs

- **Implementation**

- Overlay of DPM SE
- Not published in BDII
- Accessed using **-no-bdii** options of lcg-utils

Use-cases



Evaluation

- **Conditions**

- EGI, biomed VO (production infrastructure)
- Ultrasonic simulation of 128 jobs
- Each job has 5 input files + 1 output file
- Failure rate: 1%

- **Results**

Data manager	Total submitted jobs	Data transfer errors	Total failed jobs
ON	130	0	2
OFF	142	11	14

- **Job data transfers failure rate is almost zeroed when failover dedicated storage is used**

Conclusion

- **Developments related to data management in VIP**
 - Web-interface to access LFC
 - Local cache used as failover SE, periodically tries to replicate its files to grid Ses
- **Substantial improvement of job reliability**
- **Web interface**

The screenshot displays a web-based file management interface. The left pane, titled 'File Transfer', shows a directory tree for the path '/vip'. The right pane, titled 'Pool of Transfers', lists various file transfer operations with their names and dates.

Name	Size	Modification Date
Trash	0 B	
Home	0 B	
VIP (group)	0 B	
CreaTools (group)	0 B	
DMRI (group)	0 B	
Demo (group)	0 B	
Nadia (group)	0 B	
Mean-Shift (group)	0 B	
Object Preparation (group)	0 B	
Test (group)	0 B	
Physical Properties (group)	0 B	
SHIWA pilots (group)	0 B	
det3d4 (group)	0 B	
Cartography (group)	0 B	
GateLab (group)	0 B	
Biomed	0 B	

Name	Date
/vip/Home/MergeField.m	09/21/2011 12:52
/vip/Home/simu_Field/simu_adam/probes/simulation_settings_with	09/15/2011 17:34
/vip/VIP (group)/TestData/Simri/heart-dynamic/slices/slice01.zip	09/15/2011 17:28
hello.gwendia	09/15/2011 14:50
hello.gwendia	09/15/2011 14:48
85312-merged-result-1316007307-kingkong.grid.creatis.insa-lyon.	09/14/2011 17:49
/vip/Home/Test_V5/14-09-2011_14:06:13/partial-results/85312-me	09/14/2011 17:39
Done	
11_23:06:19/mergers/kingkong-1315984705-1	09/14/2011 10:12
/vip/Home/13-09-2011_23:06:19/mergers/clrwn159-1315948670-g	09/14/2011 10:12
/vip/Home/13-09-2011_22:04:03/mergers/wn126-1315944485-140	09/13/2011 23:02
/vip/Home/13-09-2011_09:36:42/merge-history/121805-partial-resi	09/13/2011 21:30
/vip/Home/13-09-2011_09:36:42/partial-results/55622-partial-resul	09/13/2011 21:28
/vip/Home/12-09-2011_20:26:50/partial-results/24246-merged-resi	09/12/2011 21:01
/vip/Home/05-09-2011_18:02:26/partial-results/70688522-merged-	09/07/2011 08:24
/vip/Home/06-09-2011_15:53:15/data/result-instant0-0-simri-outSir	09/06/2011 19:48
/vip/Home/06-09-2011_15:53:15/data/result-instant0-1-simri-outSir	09/06/2011 19:48
/vip/Home/05-09-2011_18:02:26/partial-results/159719-partial-resi	09/06/2011 09:46
/vip/Home/05-09-2011_18:02:26/partial-results/52861416-merged-	09/06/2011 09:42
/vip/Home/dwi-small-test-vendredi/22-07-2011_11:40:14	09/05/2011 17:18
/vip/Home/02-09-2011_15:21:43/partial-results/14363-partial-resul	09/02/2011 16:57

Authenticated as: Tristan Glatard / CREATIS v0.4.4

Thank you !

- **More information**

- VIP project website <http://www.creatis.insa-lyon.fr/vip>
- VIP platform: <http://vip.creatis.insa-lyon.fr>
- Development roadmap: <http://vip.creatis.insa-lyon.fr:9002>

- **Demo and poster at France-Grilles booth (15:30-16:00)**