A science-gateway workload archive application to the self-healing of workflow incidents

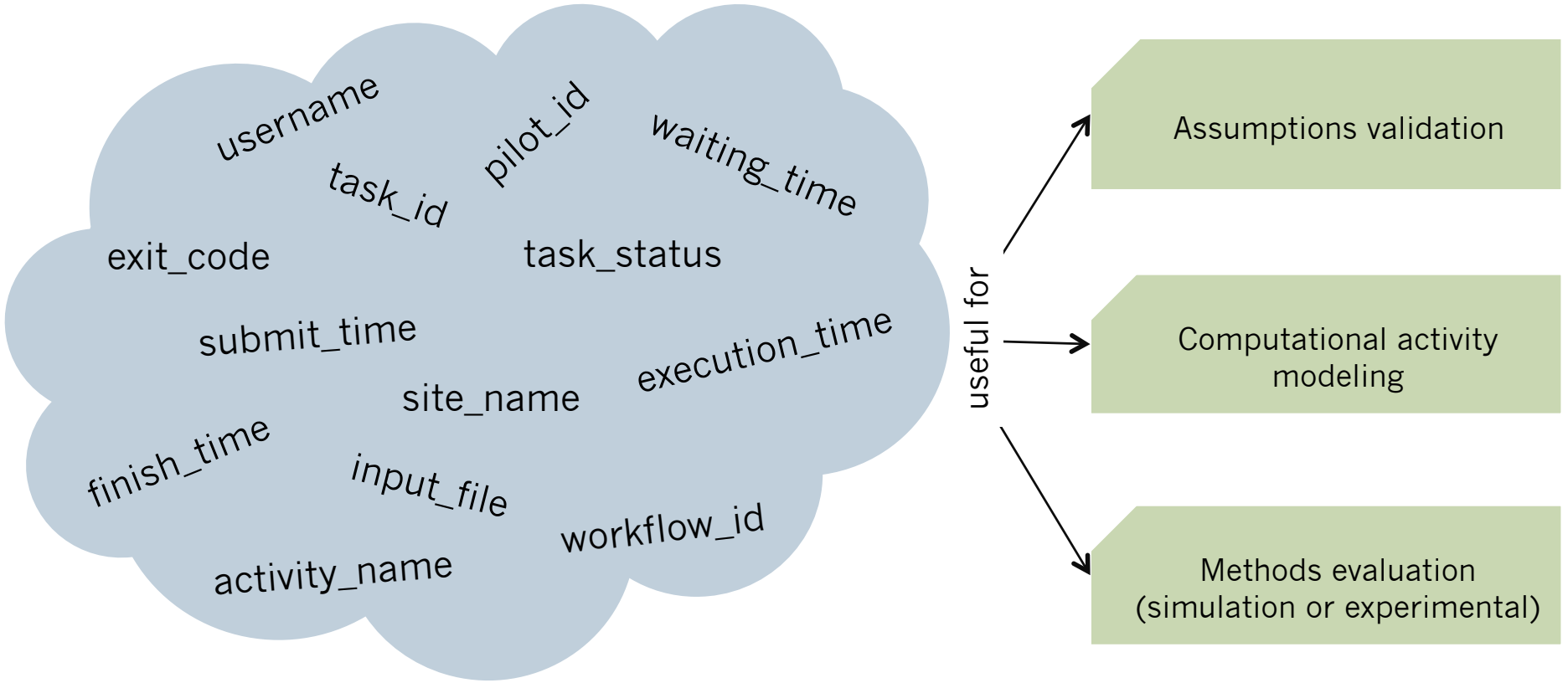
Rafael FERREIRA DA SILVA, **Tristan GLATARD**
University of Lyon, CNRS, INSERM, CREATIS
Villeurbanne, France

Frédéric DESPREZ
INRIA, University of Lyon, LIP, ENS Lyon
Lyon, France

Journées Scientifiques Mésocentres et France Grilles
October 1st-3rd 2012

Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

Context: Workload Archives



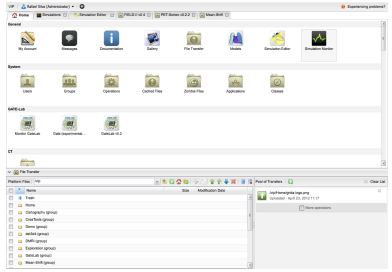
Information produced by grid workflow executions

Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

Science-gateway architecture

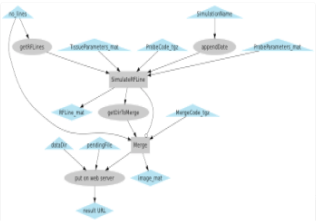
User

0. Login
1. Send input data



Web Portal

3. Launch workflow

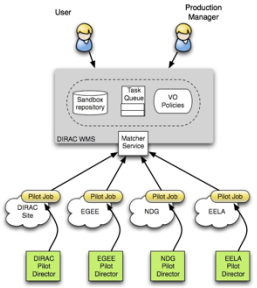


Workflow Engine

2. Transfer input files

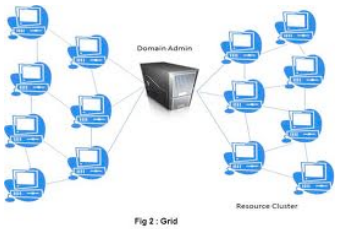
Storage Element

4. Generate and submit task



Pilot Manager

8. Get files
9. Execute
10. Upload results



Computing site

7. Get task

6. Schedule pilot jobs

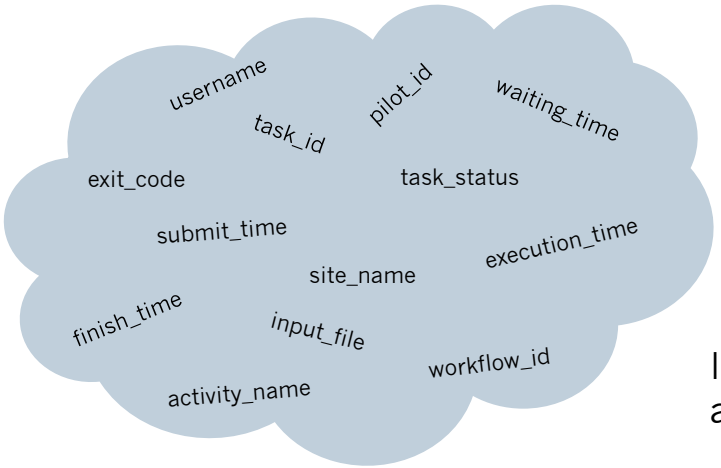
gLite

Meta-Scheduler

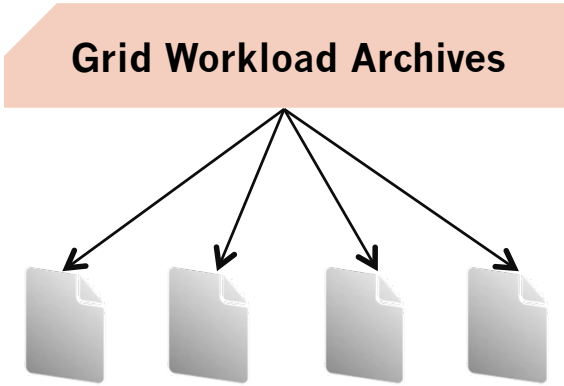
5. Submit pilot jobs

Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

State of the Art



Information gathered at **infrastructure-level**



tasks

Lack of critical information:
- Dependencies among tasks
- Task sub-steps
- Application-level scheduling artifacts
- User

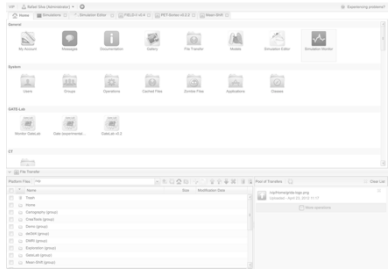
- Parallel Workloads Archive (http://www.cs.huji.ac.il/labs/parallel/workload/)
- Grid Workloads Archive (http://gwa.ewi.tudelft.nl/pmwiki/)

Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

At infrastructure-level

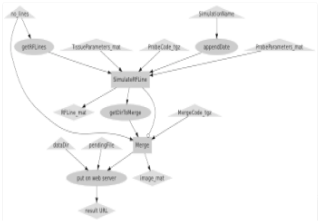
User

0. Login
1. Send input data



Web Portal

3. Launch workflow

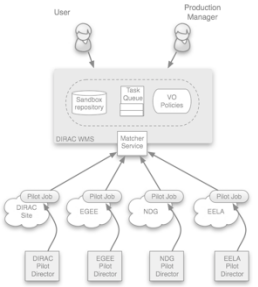


Workflow Engine

2. Transfer input files

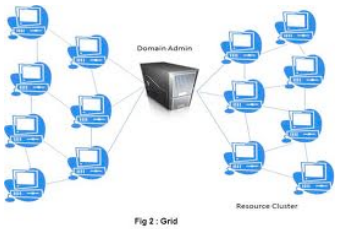
4. Generate and submit task

Storage Element



Pilot Manager

8. Get files
9. Execute
10. Upload results



Computing site

7. Get task

6. Schedule pilot jobs

Meta-Scheduler

5. Submit pilot jobs

Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

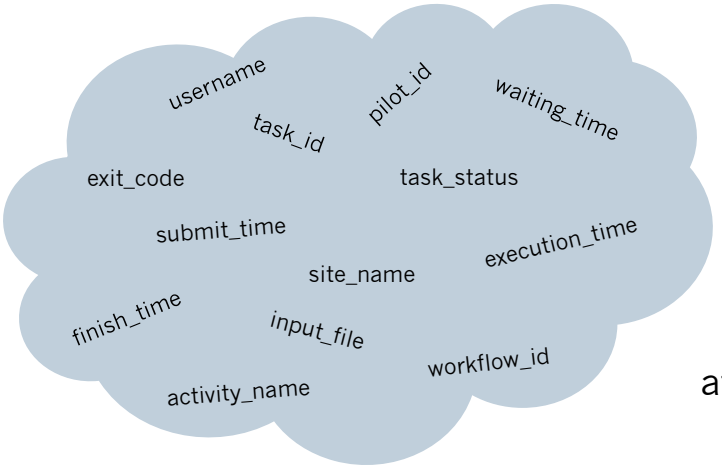
Creatis

Outline

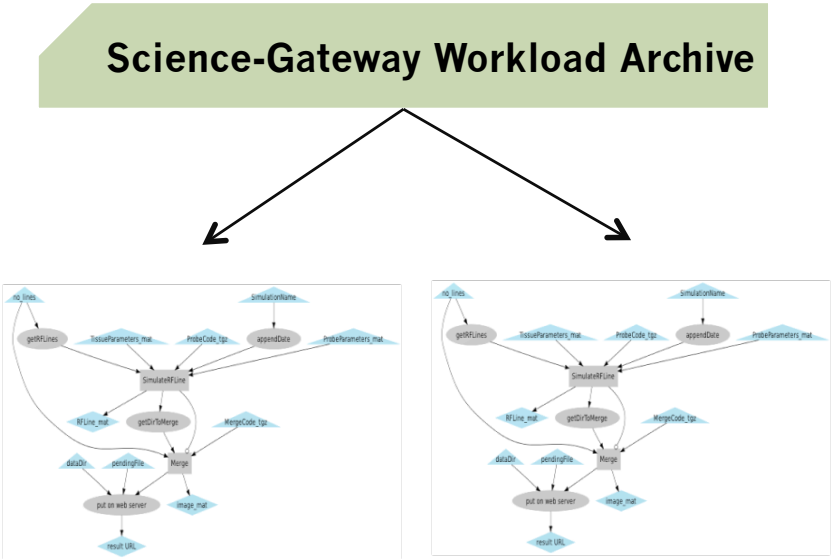
- A science-gateway workload archive
- Case studies
  - Pilot Jobs
  - Accounting
  - Task analysis
  - Bag of tasks
- Workflow Self-Healing
- Conclusions

Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

Our approach



Information gathered at **science-gateway level**



workflow executions

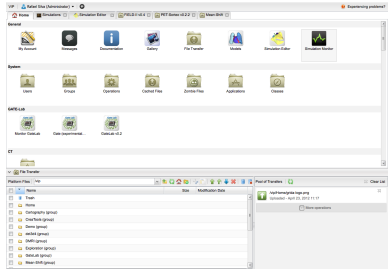
Advantages:
- Fine-grained information about tasks
- Dependencies among tasks
- Workflow characterization
- Accounting

Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

At science-gateway level

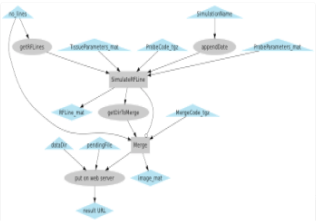
User

0. Login
1. Send input data



Web Portal

3. Launch workflow

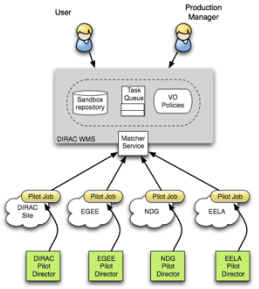


Workflow Engine

2. Transfer input files

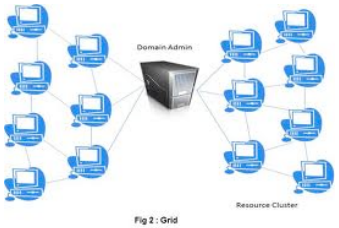
4. Generate and submit task

Storage Element



Pilot Manager

8. Get files
9. Execute
10. Upload results



Computing site

7. Get task

6. Schedule pilot jobs

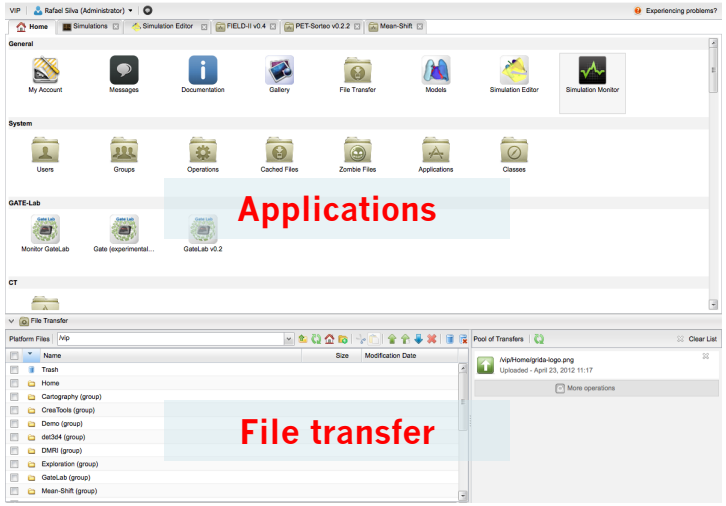
Meta-Scheduler

5. Submit pilot jobs

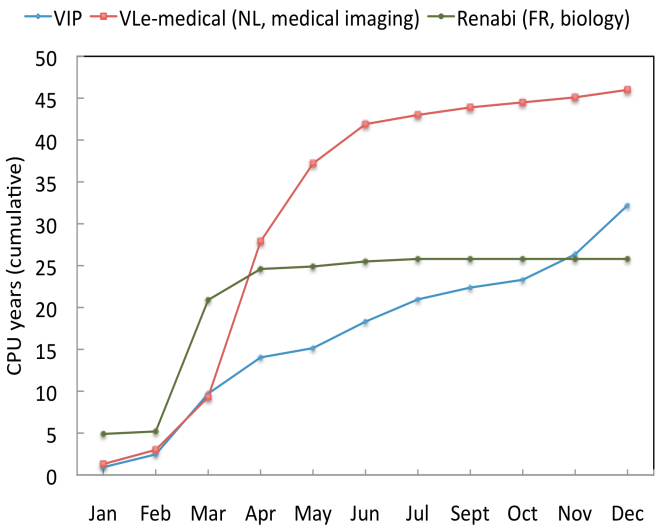
Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

Virtual Imaging Platform

- **Virtual Imaging Platform** (VIP)
  - Medical imaging science-gateway
  - Grid of 129 sites (EGI – http://www.egi.eu)
- **Significant usage**
  - Registered users: 244 from 26 countries
  - Applications: 18
  - Consumed 32 CPU years in 2011



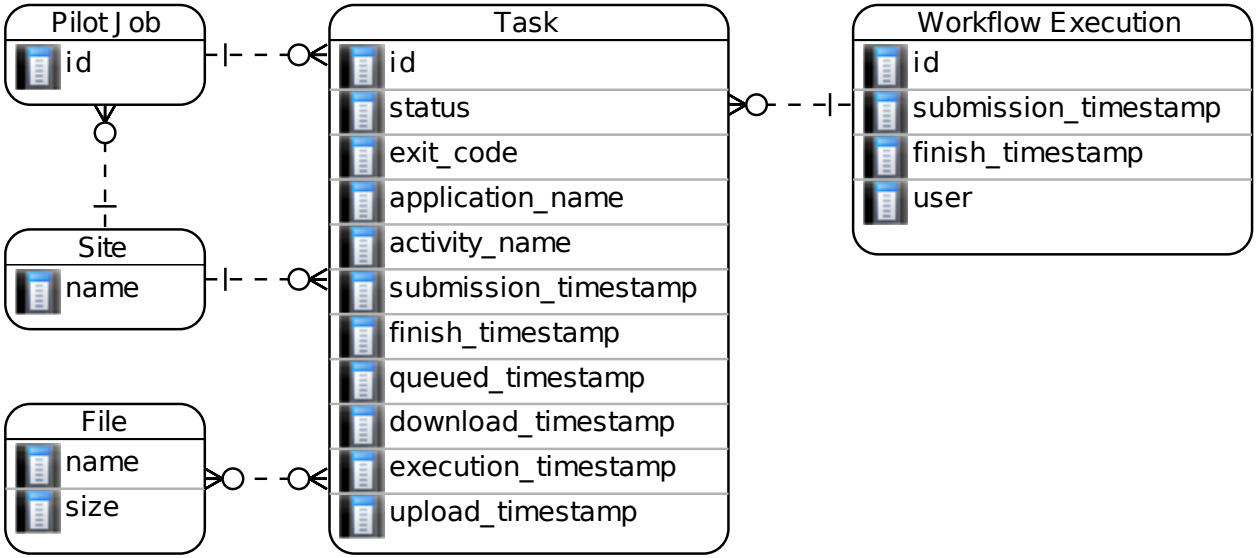
VIP – http://vip.creatis.insa-lyon.fr



VIP usage in 2011: CPU consumption of VIP and related platforms on EGI.

Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

- **Science Gateway Workload Archive (SGWA)**
  - Archive is extracted from VIP



Science-gateway archive model

Task, **Site** and **Workflow Execution** acquired from databases populated by the workflow engine at runtime

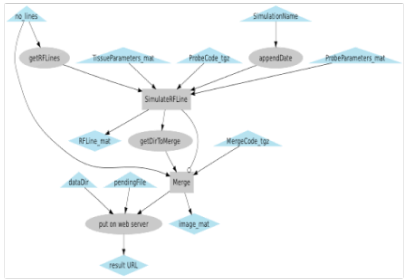
File and **Pilot Job** extracted from the parsing of task standard output and error files

Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

Workload for Case Studies

- **Based on the workload of VIP**
  - January 2011 to April 2012

112 users



2,941 workflow executions

680,988 tasks

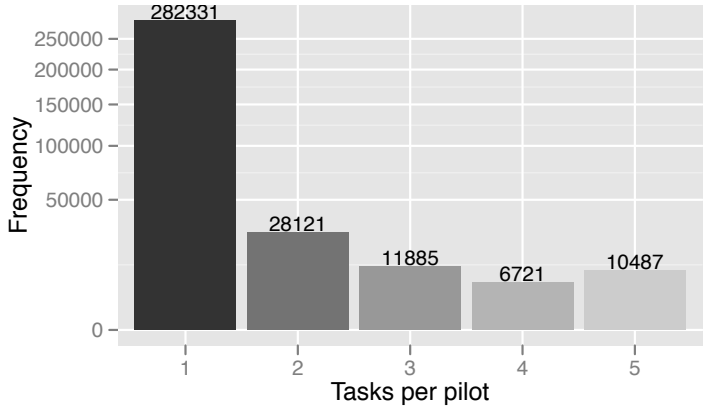
- 338,989 completed
- 138,480 error
- 105,488 aborted
  - 15,576 aborted replicas
- 48,293 stalled
- 34,162 queued

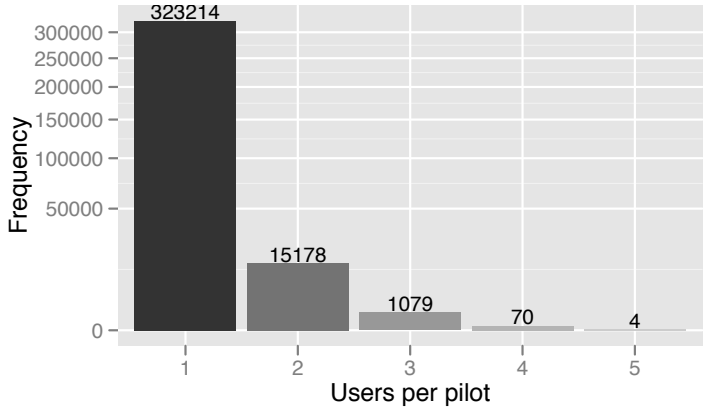
339,545 pilot jobs

Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

Pilot Jobs

- **A single pilot can wrap several tasks and users**
- At **infrastructure**-level
  - Assimilates pilot jobs to tasks and users
  - Valid for only 62% of the tasks
  - Valid for 95% of user-task associations
- At **science-gateway** level
  - Users and tasks are correctly associated to pilots

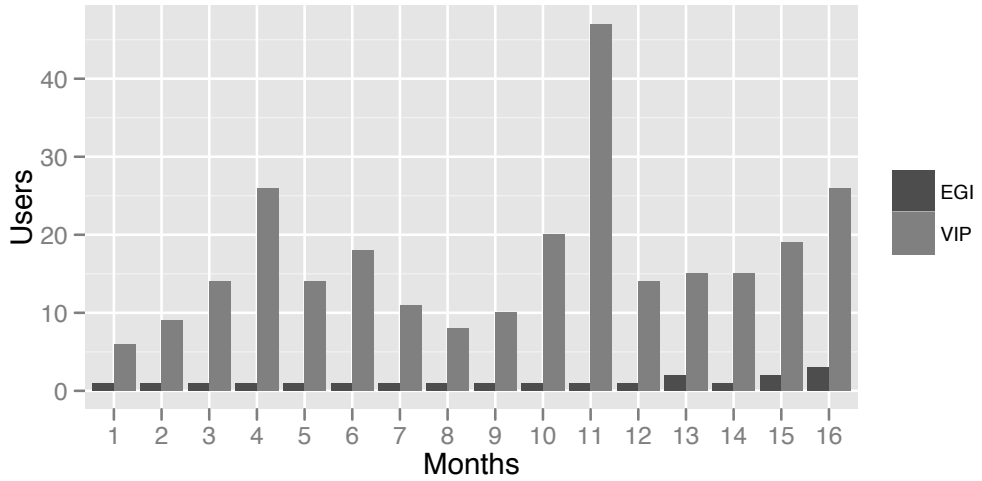




Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

Accounting: Users

- **Authentications based on login and password are mapped to X.509 robot certificates**
- At **infrastructure**-level
  - All VIP users are reported as a single user
- At **science-gateway** level
  - Maps task executions to VIP users



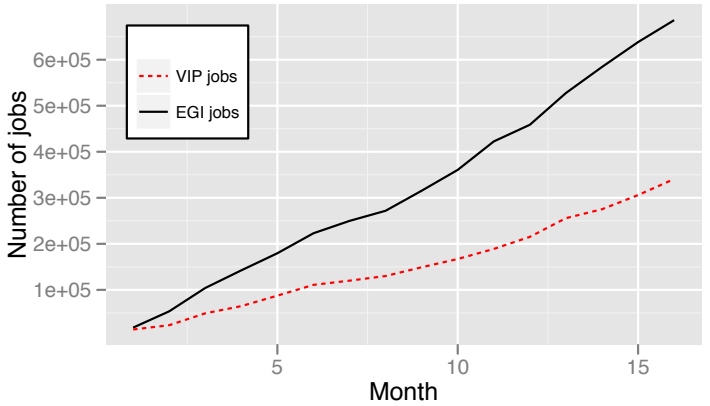
Number of reported EGI and VIP users

Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

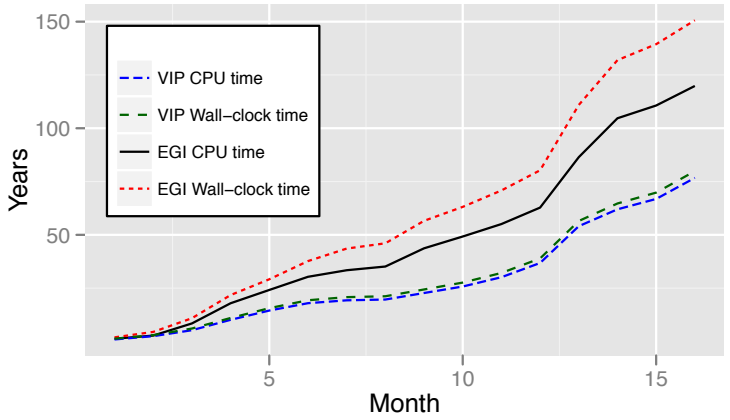
Accounting: CPU and Wall-clock Time

- **Huge discrepancy of values**
  - Pilot jobs do not register to the pilot system
  - Absence of workload
  - Outputs unretrievable
  - Pilot setup time
  - Lost tasks (a.k.a. stalled)

- **<u>Undetectable</u> at infrastructure-level**



Number of submitted pilot jobs by EGI and VIP

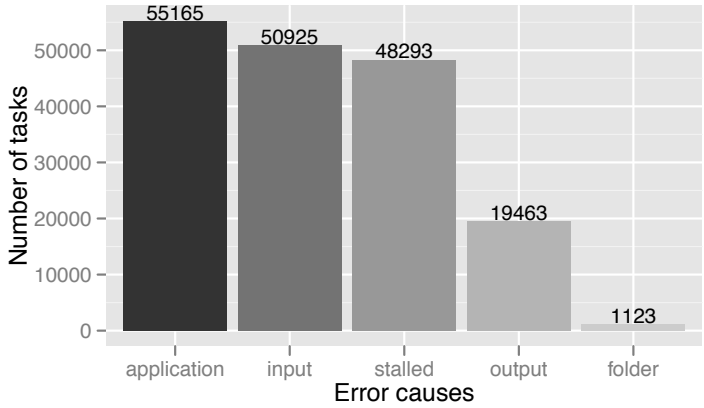


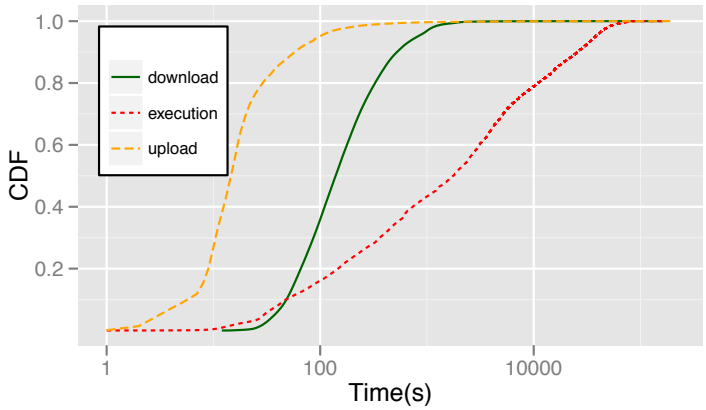
Consumed CPU and wall-clock time by EGI and VIP

Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

Task Analysis

- At **infrastructure**-level
  - Limited to task exit codes
- At **science-gateway** level
  - Fine-grained information
  - Steps in task life
  - Error causes
  - Replicas per task





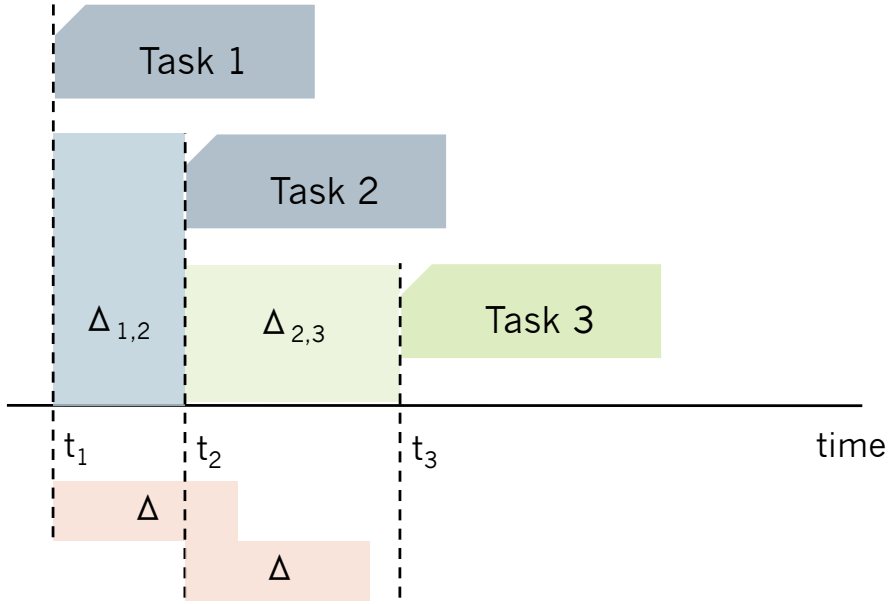
Different steps in task life

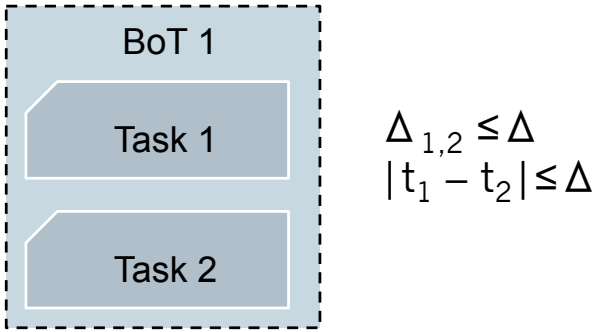
Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

Bag of Tasks: at Infrastructure level

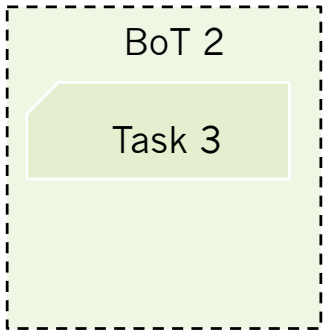
- Evaluation of the accuracy of Iosup et al.[8] method to detect bag of tasks (BoT)

- Two successively submitted tasks are in the same BoT if the time interval between submission times is lower or equal to Δ.





$\Delta_{1,2} \leq \Delta$
$|t_1 - t_2| \leq \Delta$

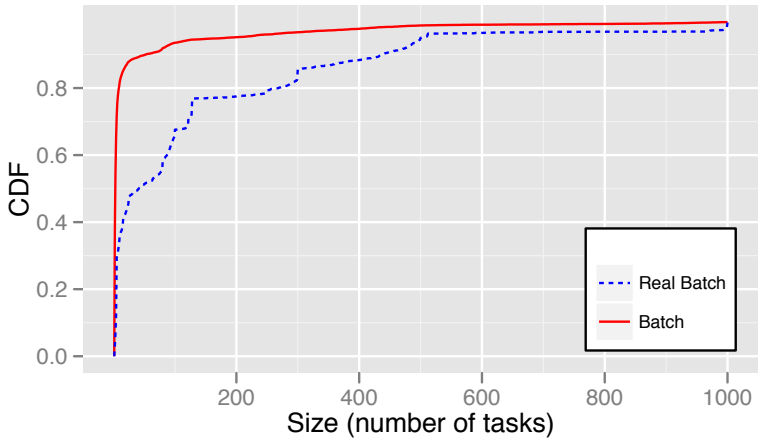


$\Delta_{2,3} > \Delta$
$|t_2 - t_3| > \Delta$

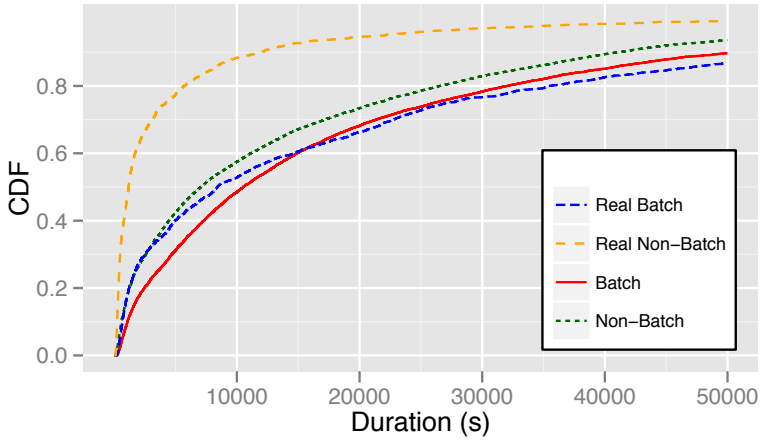
[8] Iosup, A., Jan, M., Sonmez, O., Epema, D.: The Characteristics and performance of groups of jobs in grids. In: Euro-Par. (2007) 382-393

Bag of Tasks: Size and Duration Infrastructure *vs* science-gateway

- 90% of **Batch** BoTs size ranges from 2 to 10 while it represents 50% of **Real Batch**



- **Non-Batch** duration is overestimated up to 400%

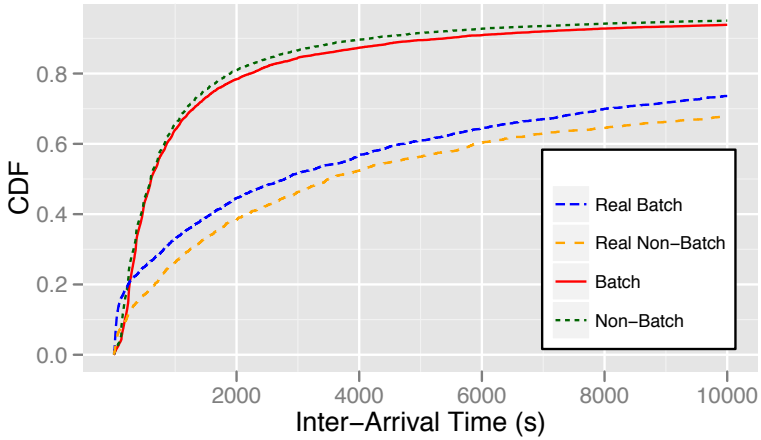


Real Batch = ground-truth BoT
Real Non-Batch = ground-truth non-BoT
Batch = Iosup et al. BoT
Non-Batch = Iosup et al. non-BoT

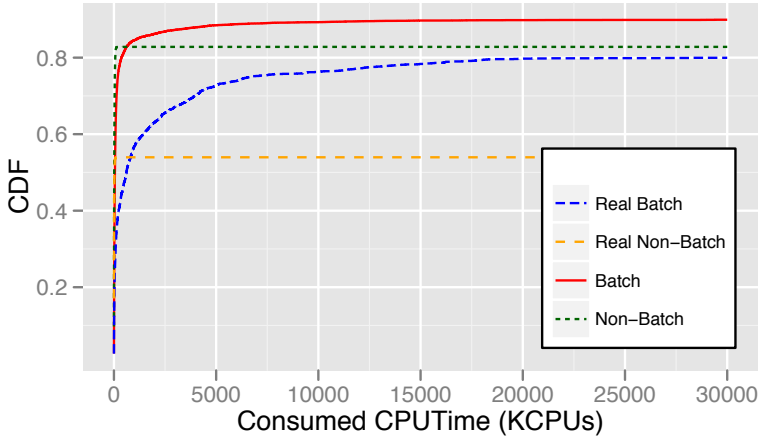
Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

Bag of Tasks: Inter-arrival Time and Consumed CPU Time

- **Batch** and **Non-Batch** inter-arrival times are underestimated by about 30%



- CPU times are underestimated of 25% for **Non-Batch** and of about 20% for **Batch**



Real Batch = ground-truth BoT
Real Non-Batch = ground-truth non-BoT
Batch = Iosup et al. BoT
Non-Batch = Iosup et al. non-BoT

Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

Creatis

Outline

- A science-gateway workload archive
- Case studies
  - Pilot Jobs
  - Accounting
  - Task analysis
  - Bag of tasks
- Workflow Self-Healing
- Conclusions

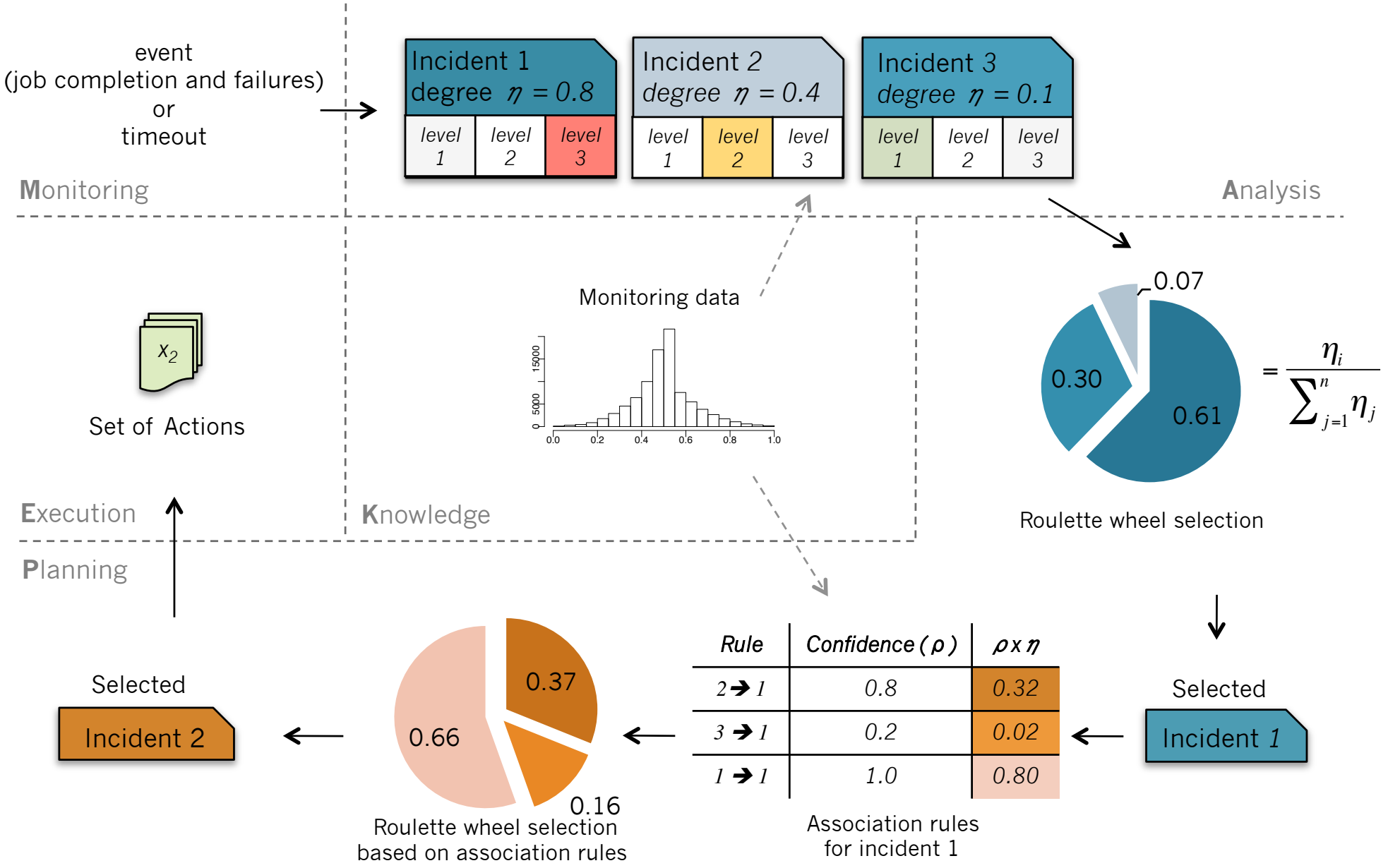
Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

Workflow Self-Healing

- **Problem**: costly manual operations
  - Rescheduling tasks, restarting services, killing misbehaving experiments or replicating data files
- **Objective**: automated platform administration
  - Autonomous detection of operational incidents
  - Perform appropriate set of actions
- **Assumptions**: online and non-clairvoyant
  - Only partial information available
  - Decisions must be fast
  - Production conditions, no user activity and workloads prediction

Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

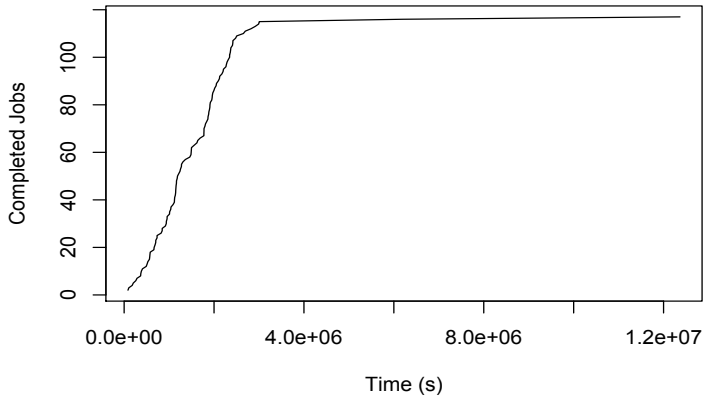
General MAPE-K loop



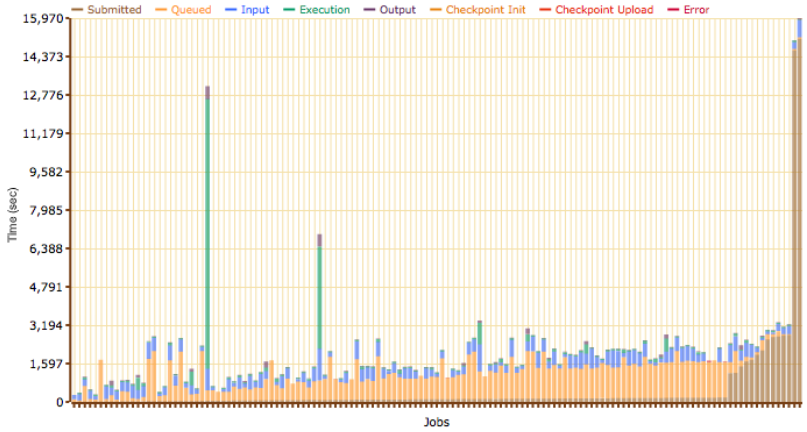
Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

Incident: Activity Blocked

- **An invocation is late compared to the others**



Invocations completion rate for a simulation



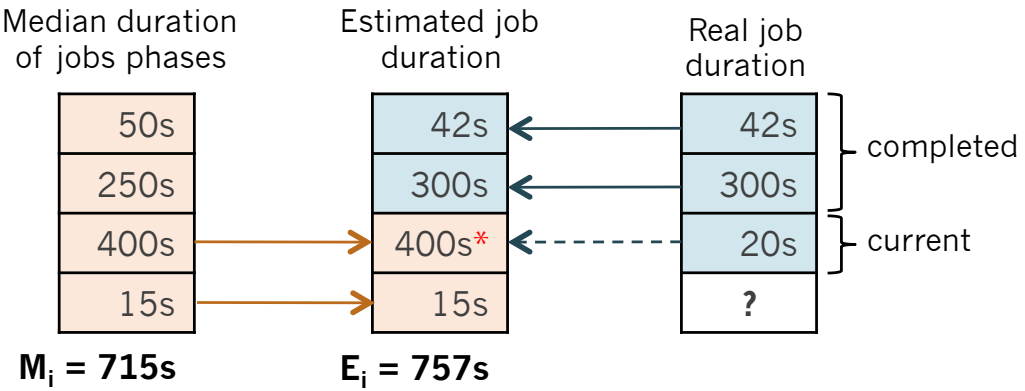
Job flow for a simulation

- **Possible causes**
  - Longer waiting times
  - Lost tasks (e.g. killed by site due to quota violation)
  - Resources with poor performance

Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

Activity blocked: degree

- **Degree computed from all completed jobs of the activity**
  - Job phases: setup ➔ inputs download ➔ execution ➔ outputs upload
  - Assumption: *bag-of-tasks* (all jobs have equal durations)
  - Median-based estimation:



*: max(400s, 20s) = 400s

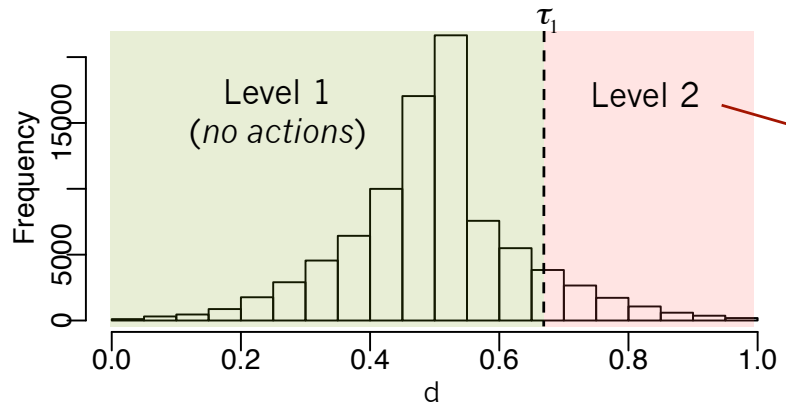
- **Incident degree:** job performance w.r.t median

$$d = \frac{E_i}{M_i + E_i} \in [0,1]$$

Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

Activity blocked: levels and actions

- **Levels:** identified from the platform logs



action: replicate jobs

- **Actions**
  - Job replication
    - Cancel replicas with bad performance
    - Replicate only if all active replicas are running

Replication process for one task

Input: Set of replicas R of a task i

01. rep = true
02. **for** $r \in R$ **do**
03. **for** $j \in R$, $j \neq r$ **do**
04. **if** $p(t_r, t_j) > \tau$ **and** j is a step further than r **then**
05. abort r
06. **done**
07. **if** r is started **and** $p(t_r, \tilde{t}) \leq \tau$ **then**
08. rep = false
09. **else if** r is queued **then**
10. rep = false
11. **done**
12. **if** rep == true **then**
13. replicate r

Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

Creatis

Experiments

- **Goal: Self-Healing vs No-Healing**
  - Cope with recoverable errors
- **Metrics**
  - Makespan of the activity execution
  - Resource waste

$$w = \frac{(CPU + data)_{self-healing}}{(CPU + data)_{no-healing}} - 1$$

  - For $w < 0$: self-healing consumed less resources
  - For $w > 0$: self-healing wasted resources

Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

Experiment Conditions

- **Software**
  - Virtual Imaging Platform
  - MOTEUR workflow engine
  - DIRAC pilot job system
- **Infrastructure**
  - European Grid Infrastructure (EGI): production, shared
  - Self-Healing and No-Healing launched simultaneously
- **Experiment parameters**
  - Task and file replication limited to 5
  - Failed task resubmission limited to 5

Creatis

Applications

FIELD-II/pasa

- Ultrasound imaging simulation
- 122 invocations
- CPU Time: 15 min
- ~210 MB
- Data-intensive

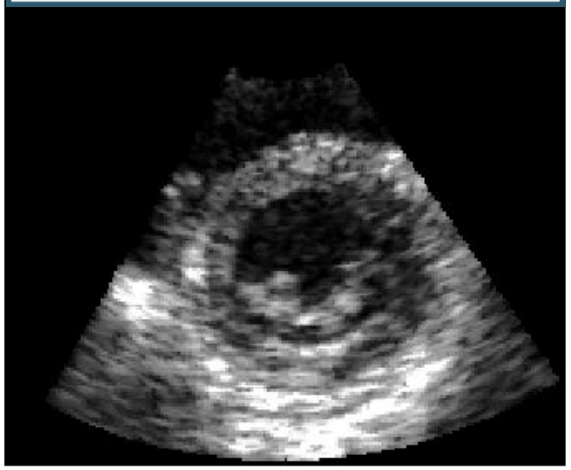


Image courtesy of ANR project US-Tagging
http://www.creatis.insa-lyon.fr/us-tagging/news
O. Bernard, M. Alessandrini

Mean-Shift/hs3

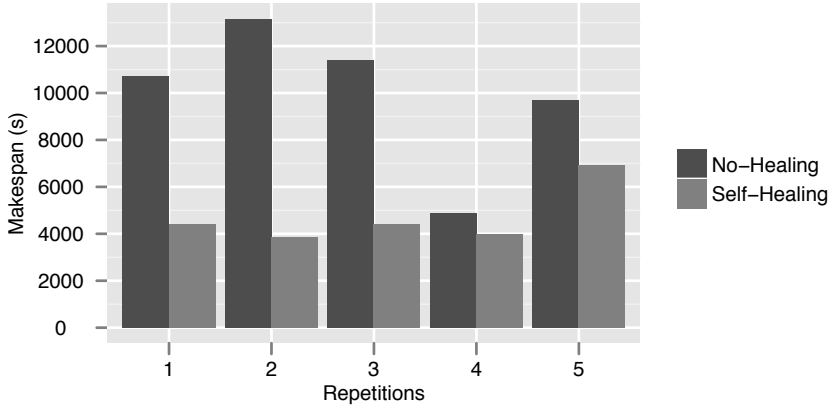
- Image denoising
- 250 invocations
- CPU Time: 1 hour
- ~182 MB
- CPU-intensive

Image courtesy of Ting Li
http://www.creatis.insa-lyon.fr

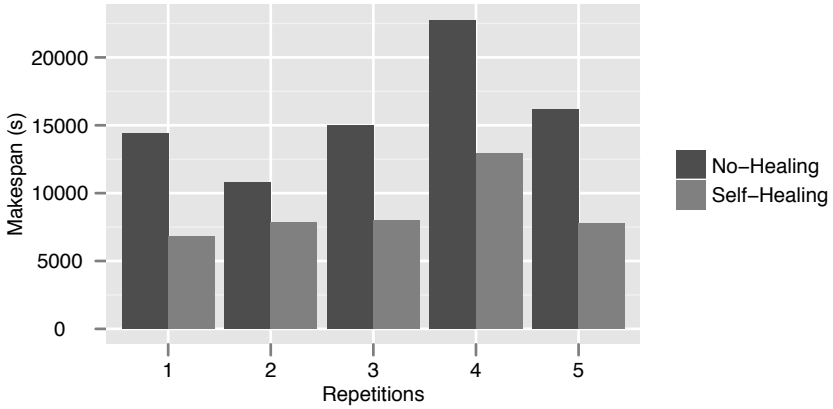
Results

- **Experiment**: tests if recoverable errors are detected

FIELD-II/pasa



Mean-Shift/hs3



speeds up execution up to 4

speeds up execution up to 2.6

Repetition	*w*
1	–0.10
2	–0.15
3	–0.09
4	0.05
5	–0.26

Self-Healing process reduced resource consumption up to 26% when compared to the *No-Healing* execution

Repetition	*w*
1	–0.02
2	–0.20
3	–0.02
4	–0.02
5	–0.01

Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

Creatis

Conclusions

- **Science-gateway model of workload archive**
  - Illustration by using traces of the VIP from 2011/2012
- **Added value when compared to infrastructure-level traces**
  - Exactly identify tasks and users
  - Distinguishes additional workload artifacts from real workload
  - Fine-grained information about tasks
  - Ground-truth of bag of tasks
- **Self-healing of worklfow incidents**
  - Implements a generic MAPE-K loop
  - Incident degrees computed online
  - Speeds up execution up to a factor of 4
  - Reduced resource consumption up to 26%
  - Successfull example of self-healing loop deployed in production
- **VIP is openly available at** http://vip.creatis.insa-lyon.fr
- **Traces are available to the community in the Grid Observatory:** http://www.grid-observatory.org

Rafael Ferreira da Silva – rafael.silva@creatis.insa-lyon.fr

A science-gateway workload archive application to the self-healing of workflow incidents

Thank you for your attention.
Questions?

ACKNOWLEDGMENTS
VIP users and project members
French National Agency for Research (ANR-09-COSI-03)
European Grid Initiative (EGI)
France-Grilles

Rafael FERREIRA DA SILVA, Tristan GLATARD
University of Lyon, CNRS, INSERM, CREATIS
Villeurbanne, France

Frédéric DESPREZ
INRIA, University of Lyon, LIP, ENS Lyon
Lyon, France