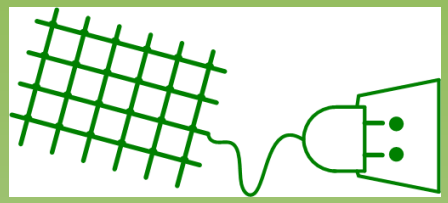


Green-Net

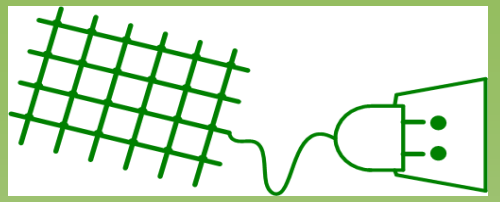
Towards a Green Grid'5000

Anne-Cécile Orgerie, Laurent Lefèvre
INRIA RESO – Université de Lyon
annececile.orgerie@ens-lyon.fr, laurent.lefevre@inria.fr

April 8, 2009, Grid'5000 Spring School, Nancy - France

ECOLE NORMALE SUPERIEURE DE LYON · RESO · INRIA RHÔNE-ALPES · LIP · UCBL · CNRS CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE · UNIVERSITÉ DE LYON · Grid'5000

Green-Net Project



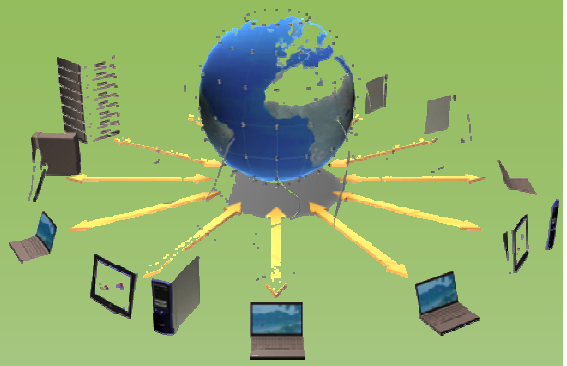
- **Power aware software frameworks for high performance data transport and computing in large scale distributed systems**
- ARC GREEN-NET : (Action de Recherche Coopérative supported by INRIA)
- Partners teams :
  - IRIT (Toulouse)
  - INRIA MESCAL (Grenoble)
  - INRIA RESO (Lyon)
  - Virginia Tech (USA)

http://www.ens-lyon.fr/LIP/RESO/Projects/GREEN-NET

Plan

1. Introduction
2. Two years in the life of Grid'5000
3. The consumption measurement infrastructure
4. EARI: Energy-Aware Reservation Infrastructure
5. Greening Grid'5000
6. Conclusion and future works

Adressed challenges



- **How to reduce energy usage without compromising QoE: Quality of Experiment ?**

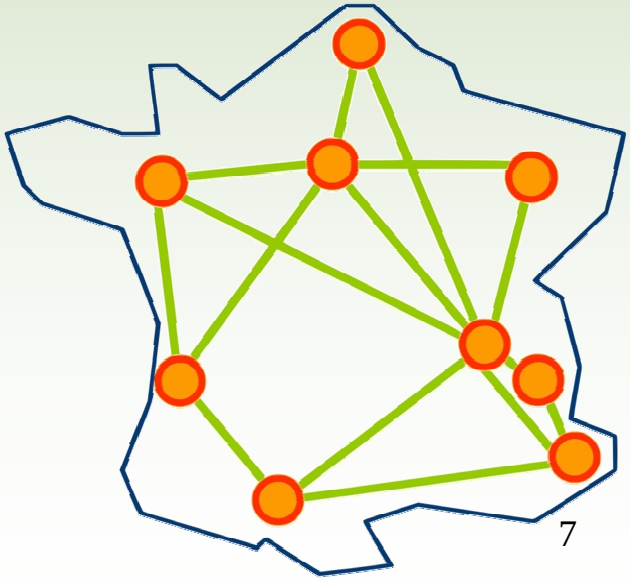
- How to **understand** and to analyze the **usage** of large scale platforms?
- How to **apply** energy usage **models** on this experimental usage ?
- How to **monitor** lively such usage (multiple views (Grids, datacenters, clusters, nodes, services, processes, threads)) ?
- How to **design** energy aware software **frameworks** ?

- **Our context:** Ressources always powered on / Reservation infrastructure / Large-scale distributed systems

Two years in the life of Grid'5000

Grid'5000 usage

- Two years of logs: 2007 & 2008
- Logs furnished by *oarstat*
- 8 sites
- Grid view, site view and node view
- 2 INRIA reports to appear



Grid'5000

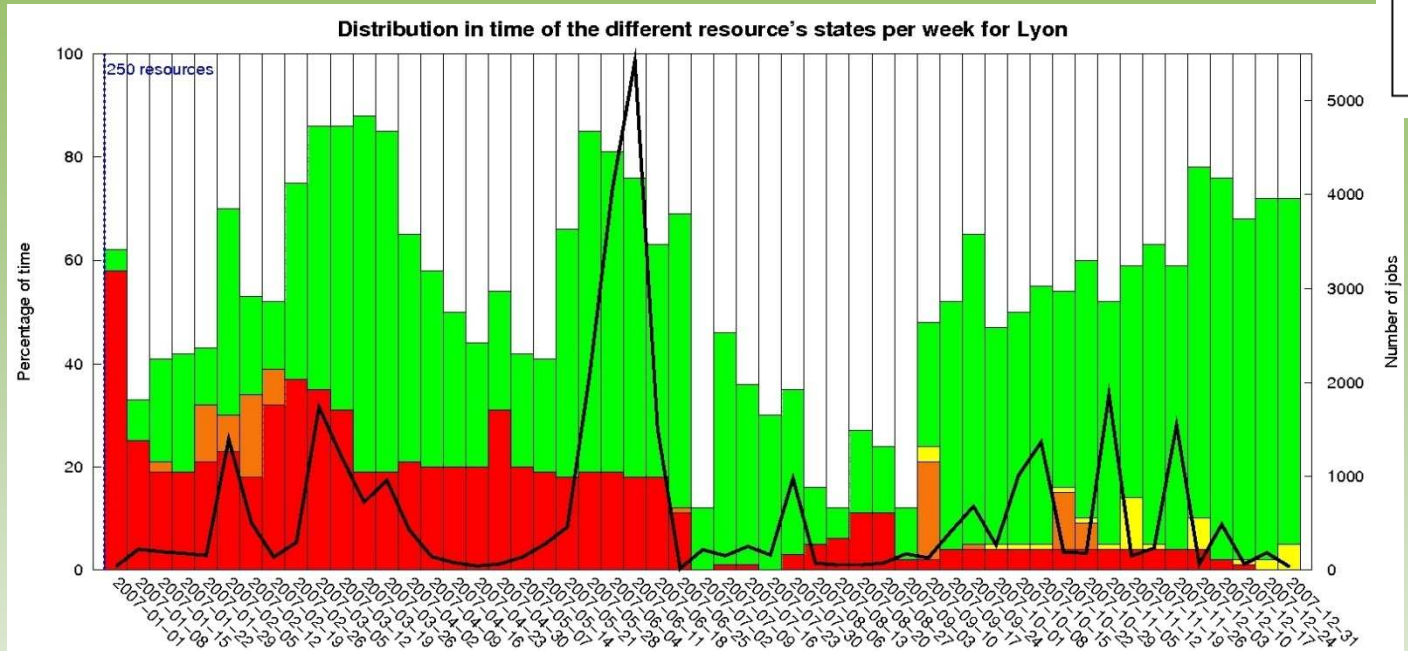
Grid statistics for 2007

job = reservation; resource = core

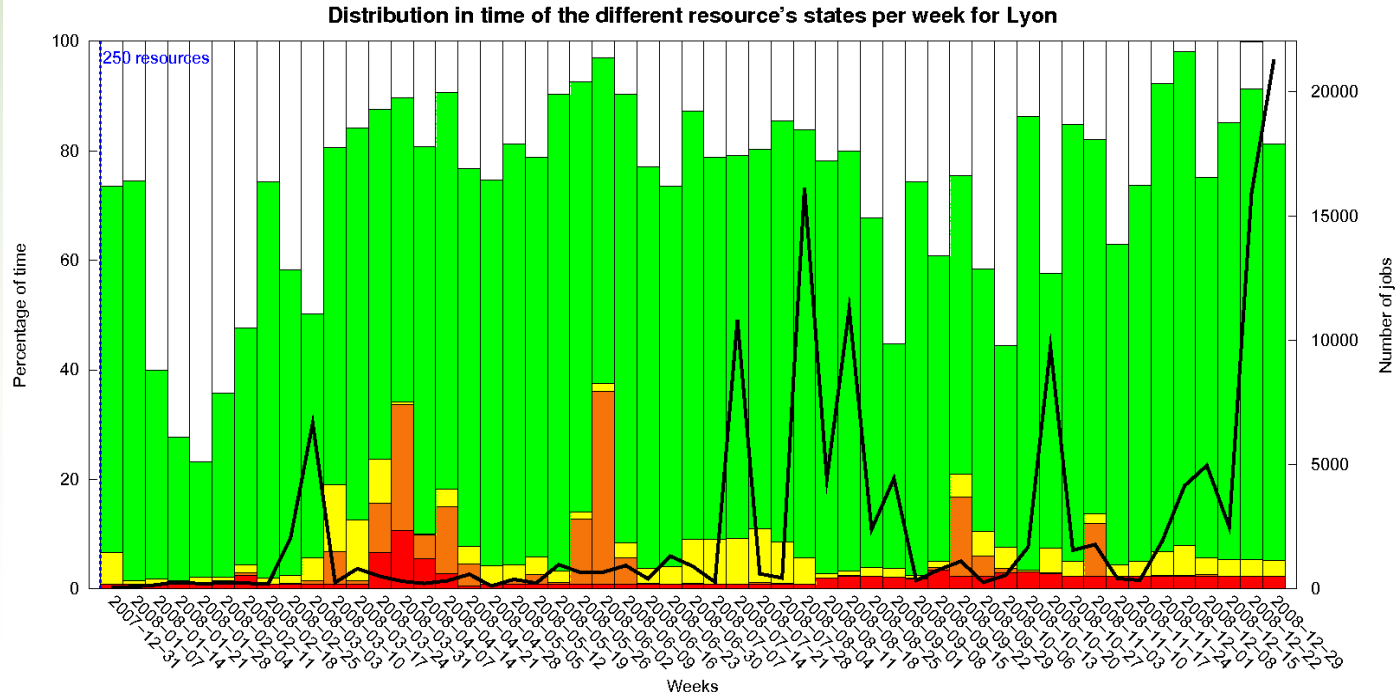
'real' activity = without dead and absent time

Site	Number of jobs	Number of resources	Mean number of resources per job	Mean duration of a job in seconds	Percentage of 'real' activity
Bordeaux	45775	650	55.50	5224.59	47.40%
Grenoble	19211	72	4.06	4473.76	14.25%
Lille	330694	250	4.81	1446.13	36.08%
Lyon	33315	322	41.64	3246.15	45.92%
Nancy	63435	574	22.46	19480.49	56.21%
Orsay	26448	684	47.45	4322.54	18.69%
Rennes	36433	714	54.85	7973.39	49.42%
Sophia	35179	568	57.93	4890.28	51.36%
Toulouse	20832	434	12.89	7420.07	49.99%

Usage evolution



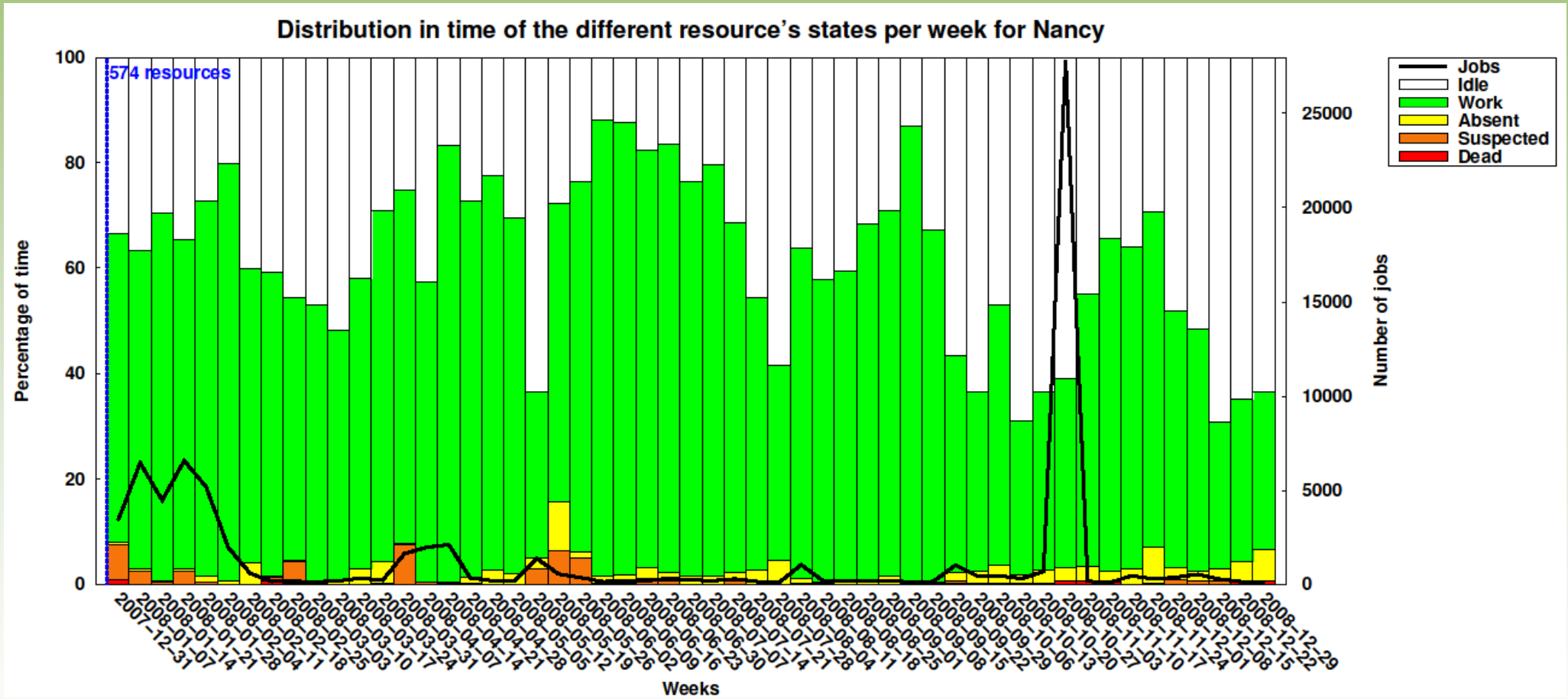
46% of activity



69% of activity

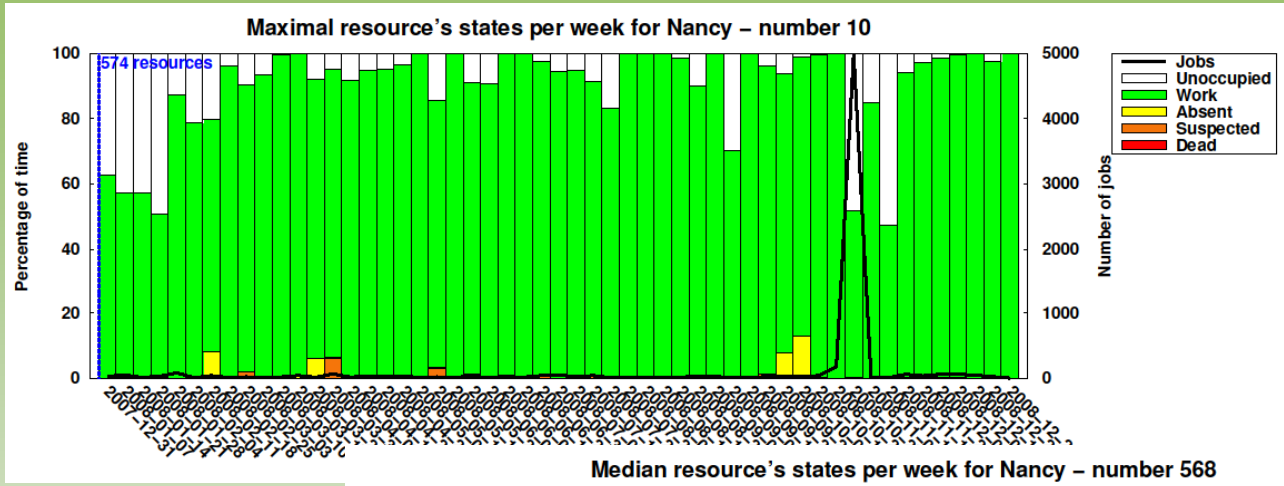
Grid'5000

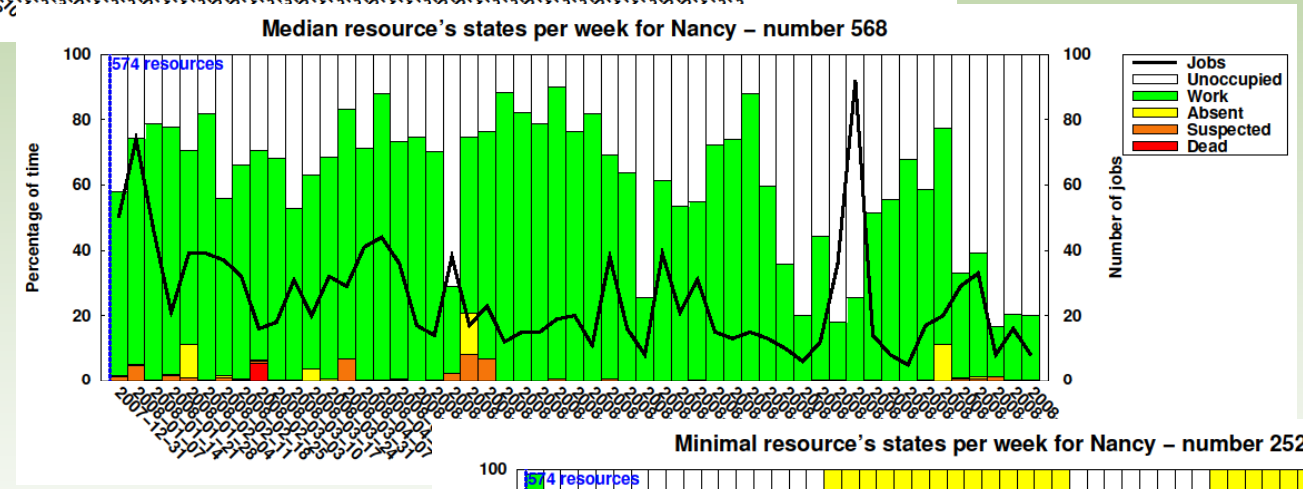
Nancy in 2008: big burst / impact of best effort jobs

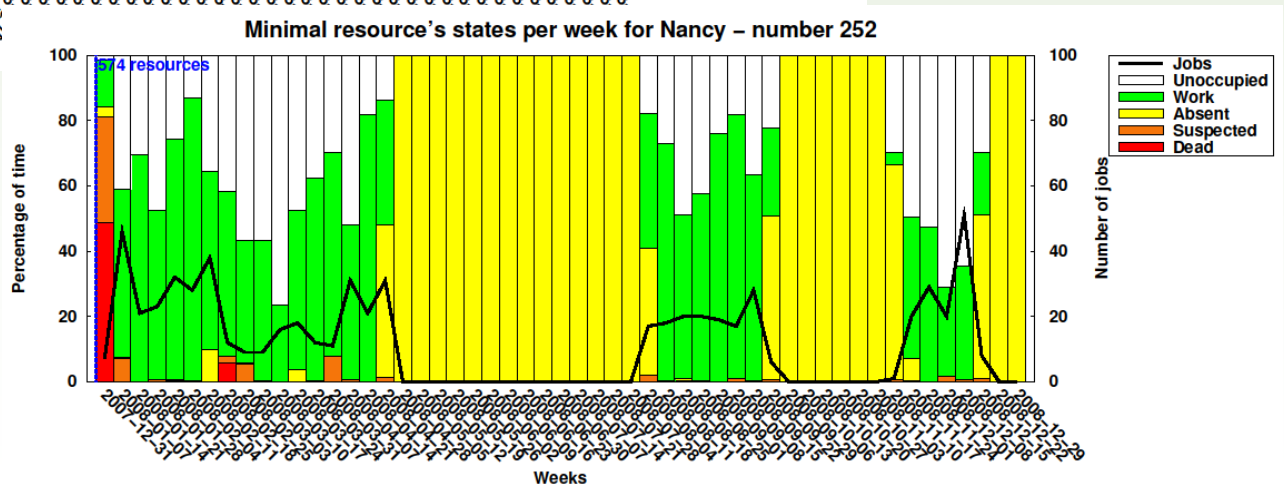


Grid'5000

Nancy: max, min and median resources

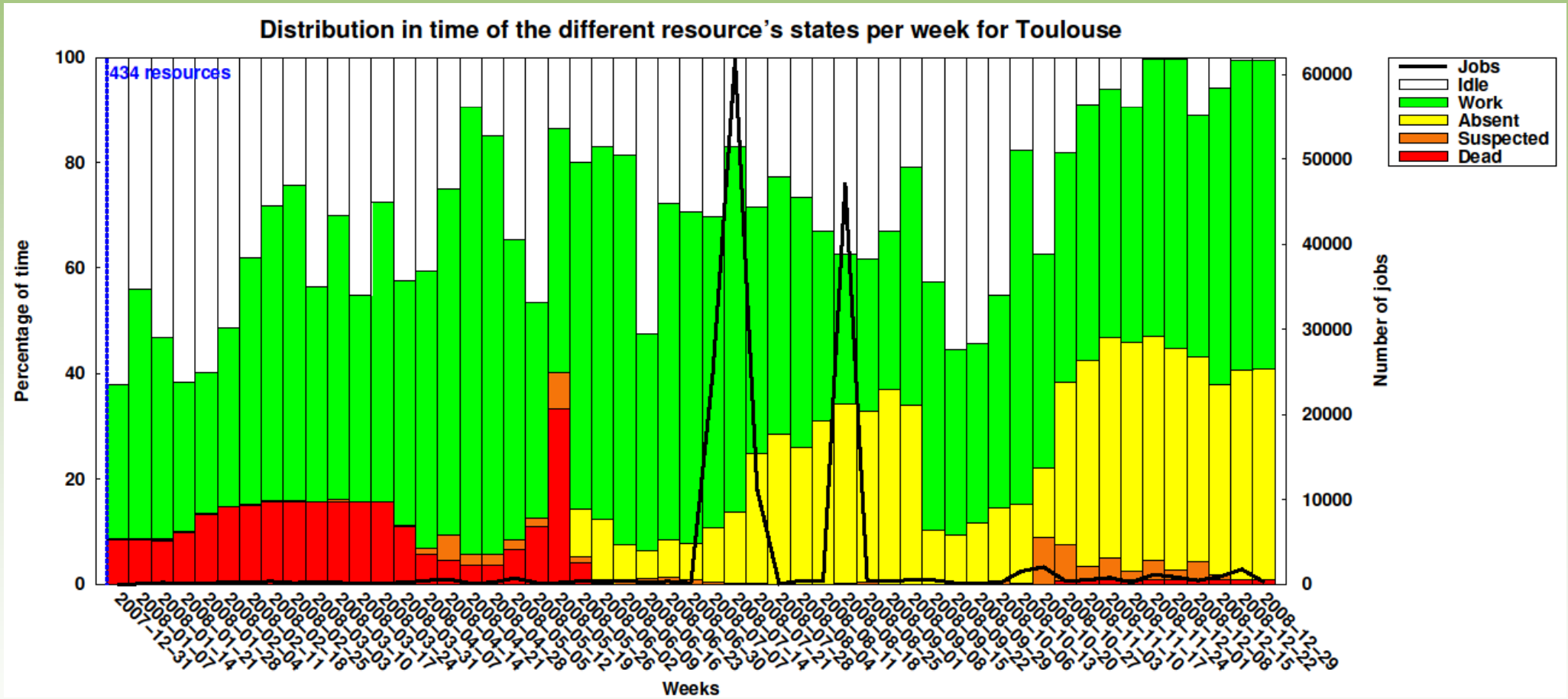






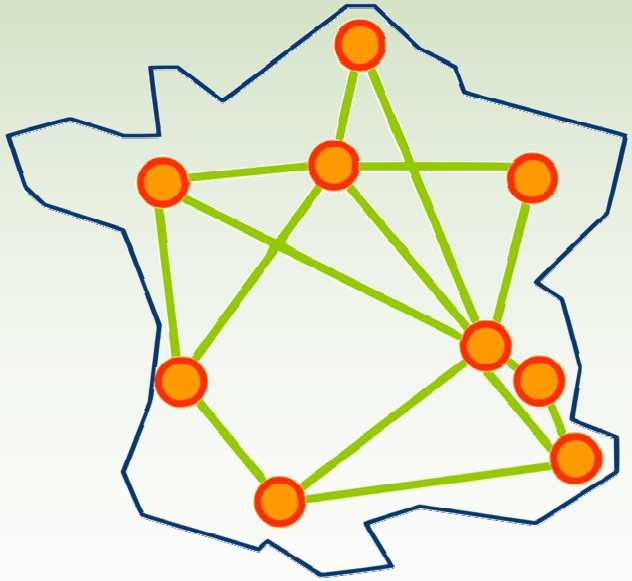
Grid'5000

Toulouse in 2008: 15% of absent



Conclusions about the usage

- Specific usage of an experimental Grid
- Great differences between 2007 and 2008
- Significative bursts → significative gaps
- A lot of small reservations

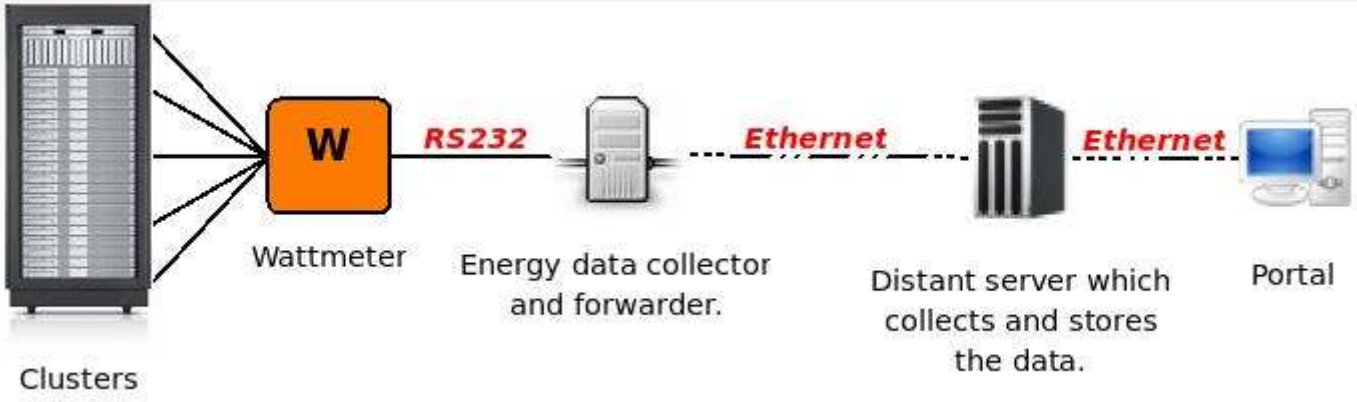


Grid'5000

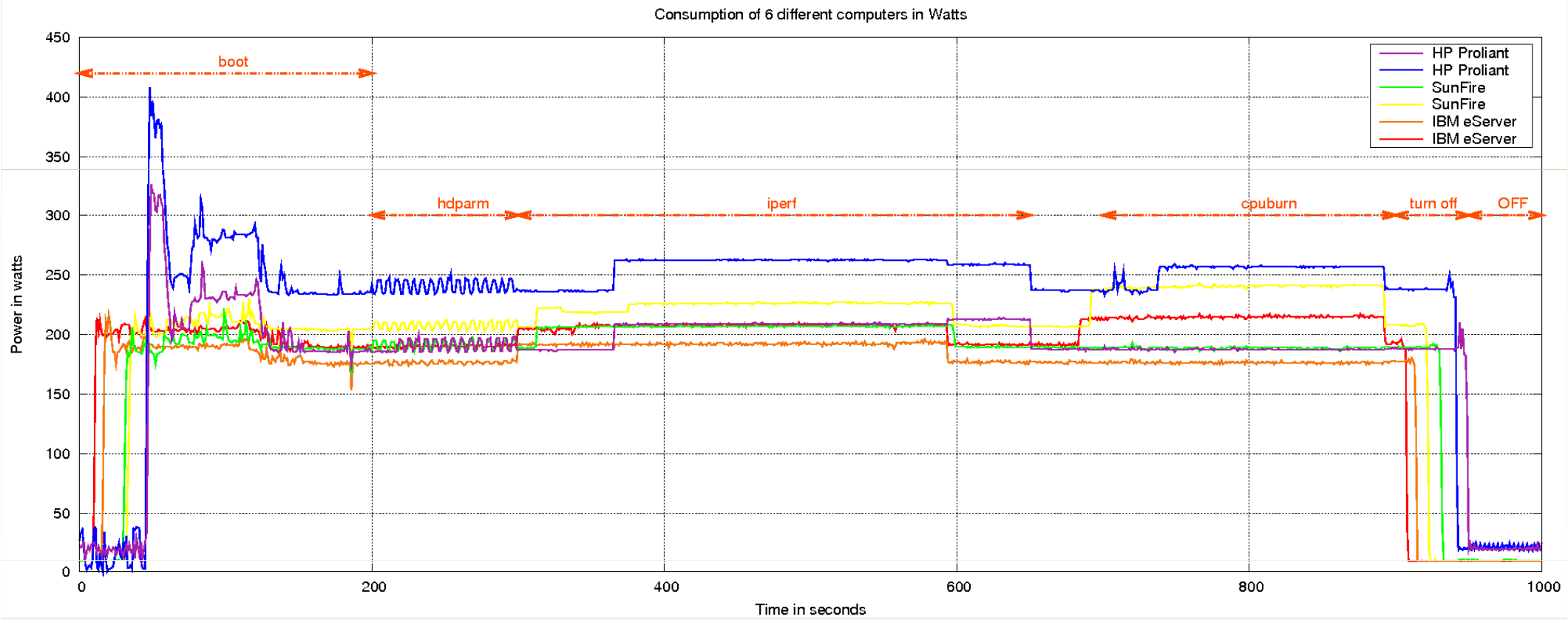
The consumption measurement infrastructure deployed on Grid'5000

Consumption measurement

- Autonomic wattmeter
- Furnished by Omegawatt
- One measure per second and per node
- 6 nodes at a time
- 3 different sites: Lyon, Grenoble and Toulouse



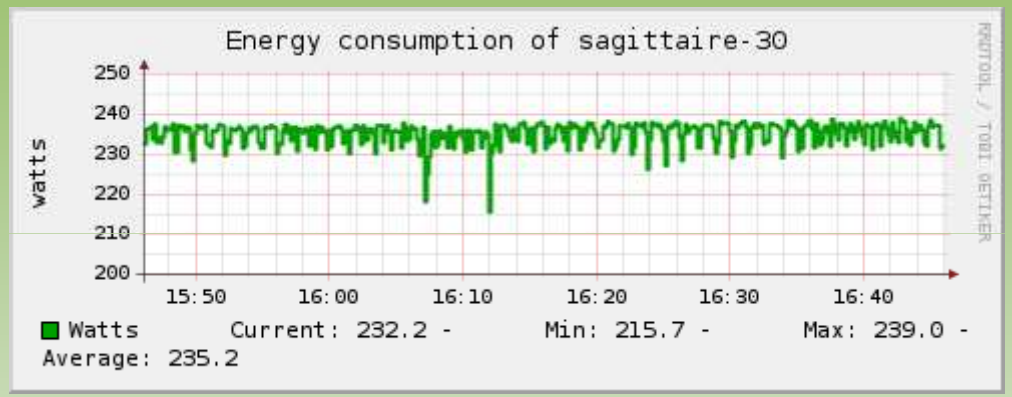
Example : Monitoring of 6 nodes in Lyon

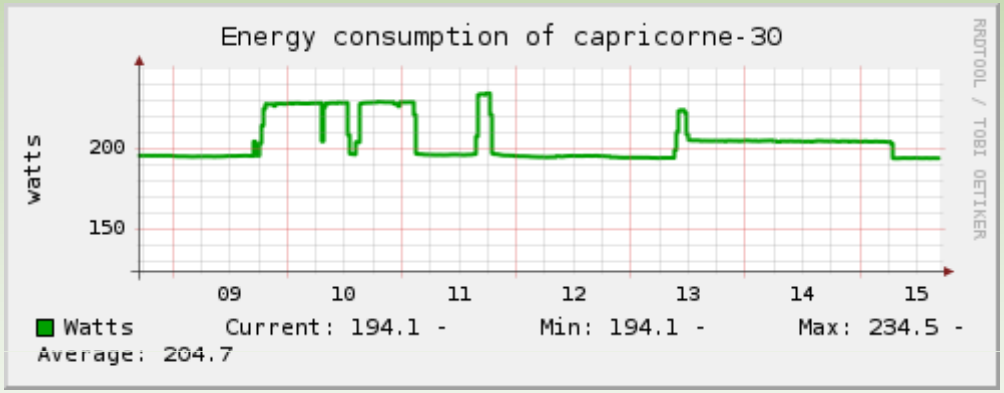


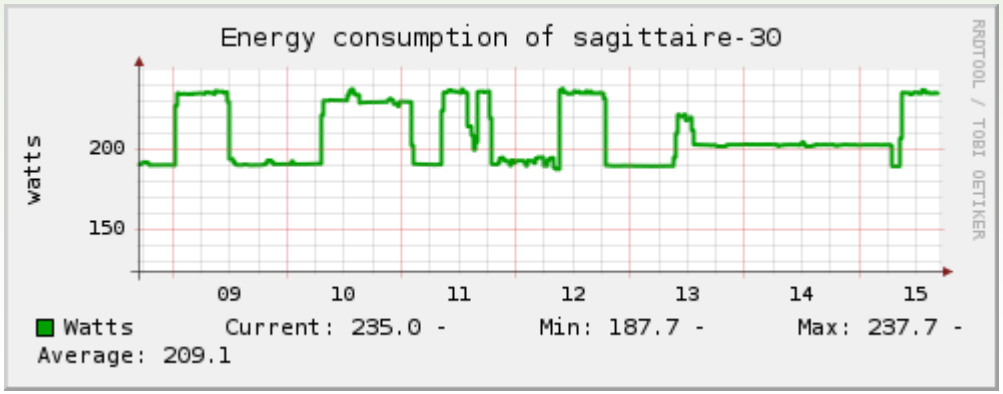
- Idle consumption really high: around 190 Watts
- Off consumption around 10 Watts

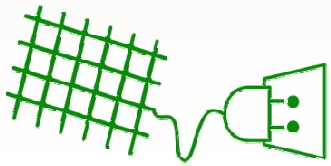
Idea: to switch off the unused nodes

Live energy monitoring



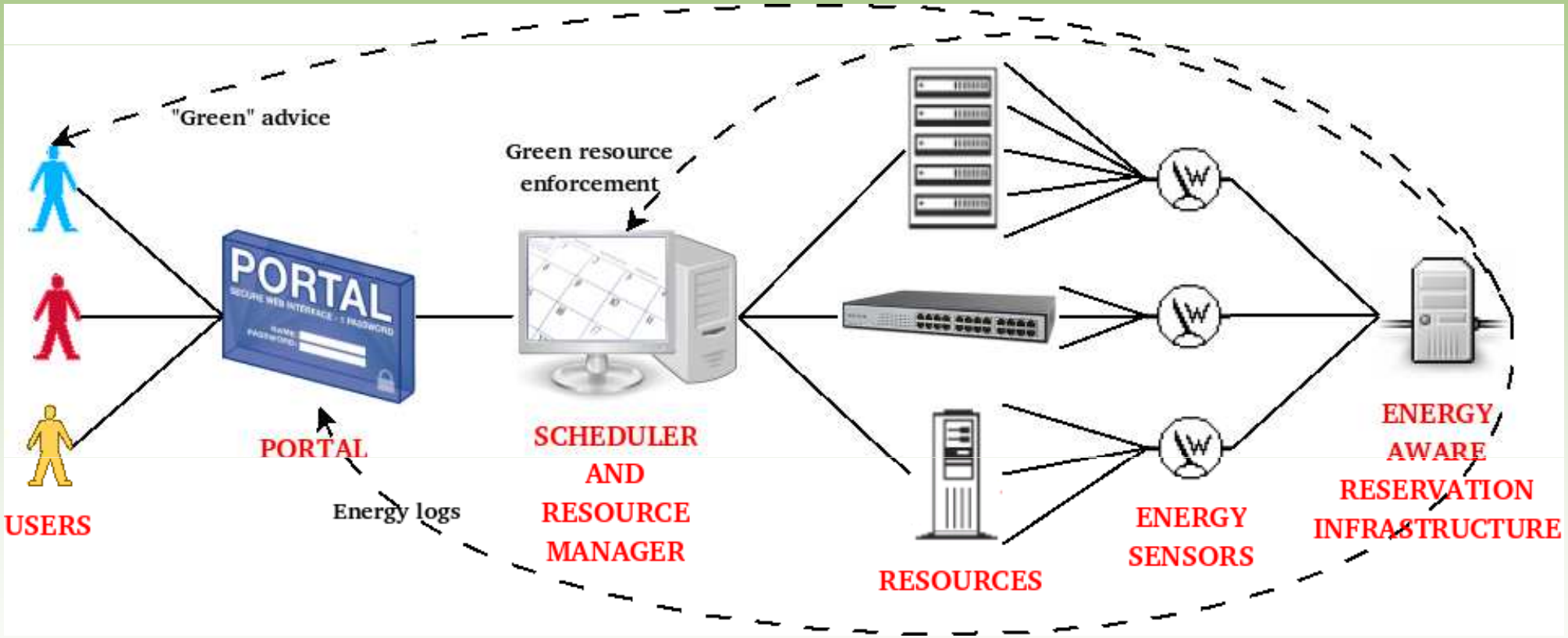




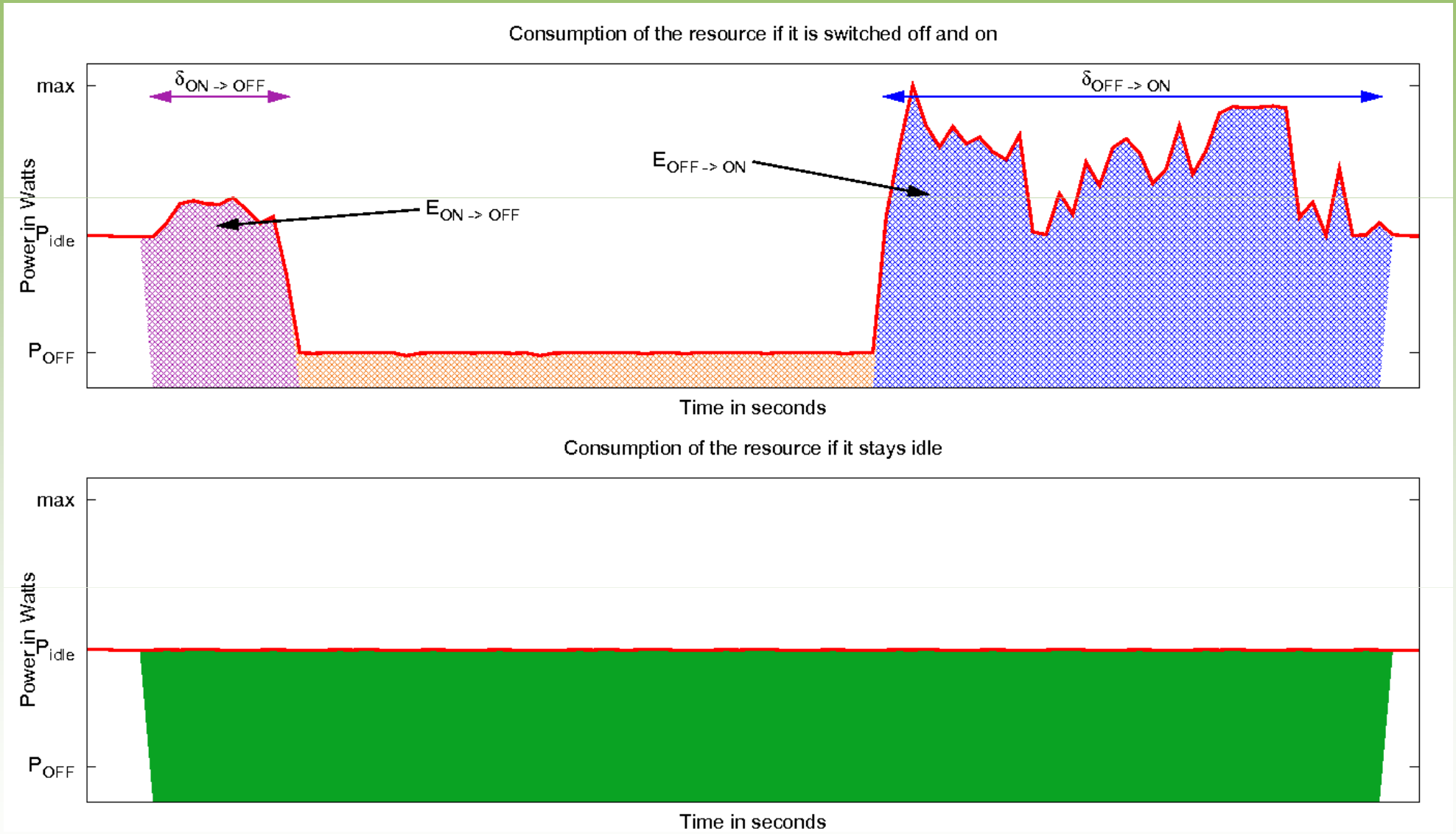


EARI: Energy-Aware Reservation Infrastructure

Architecture for energy efficient system

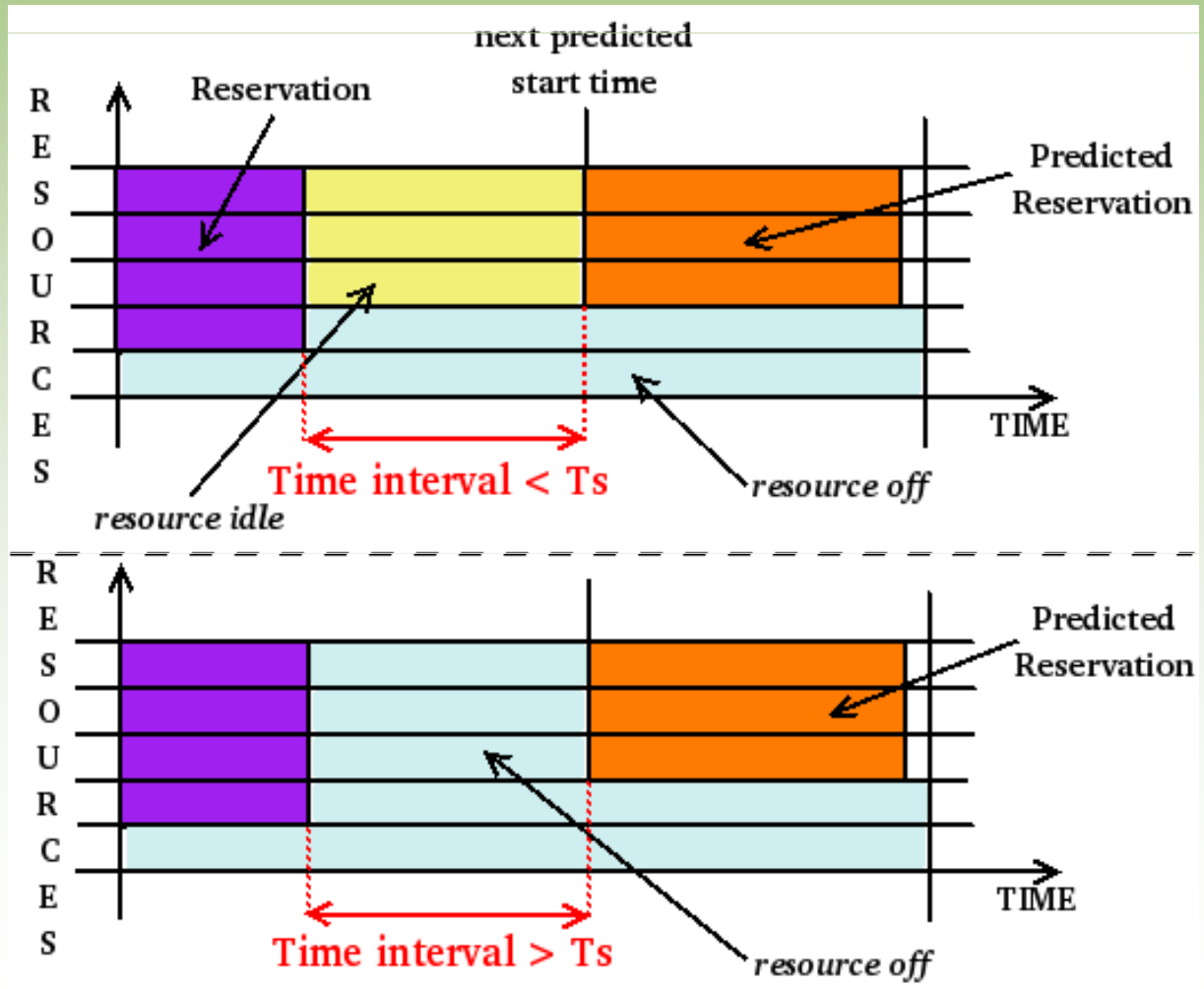


When can we switch off a node?

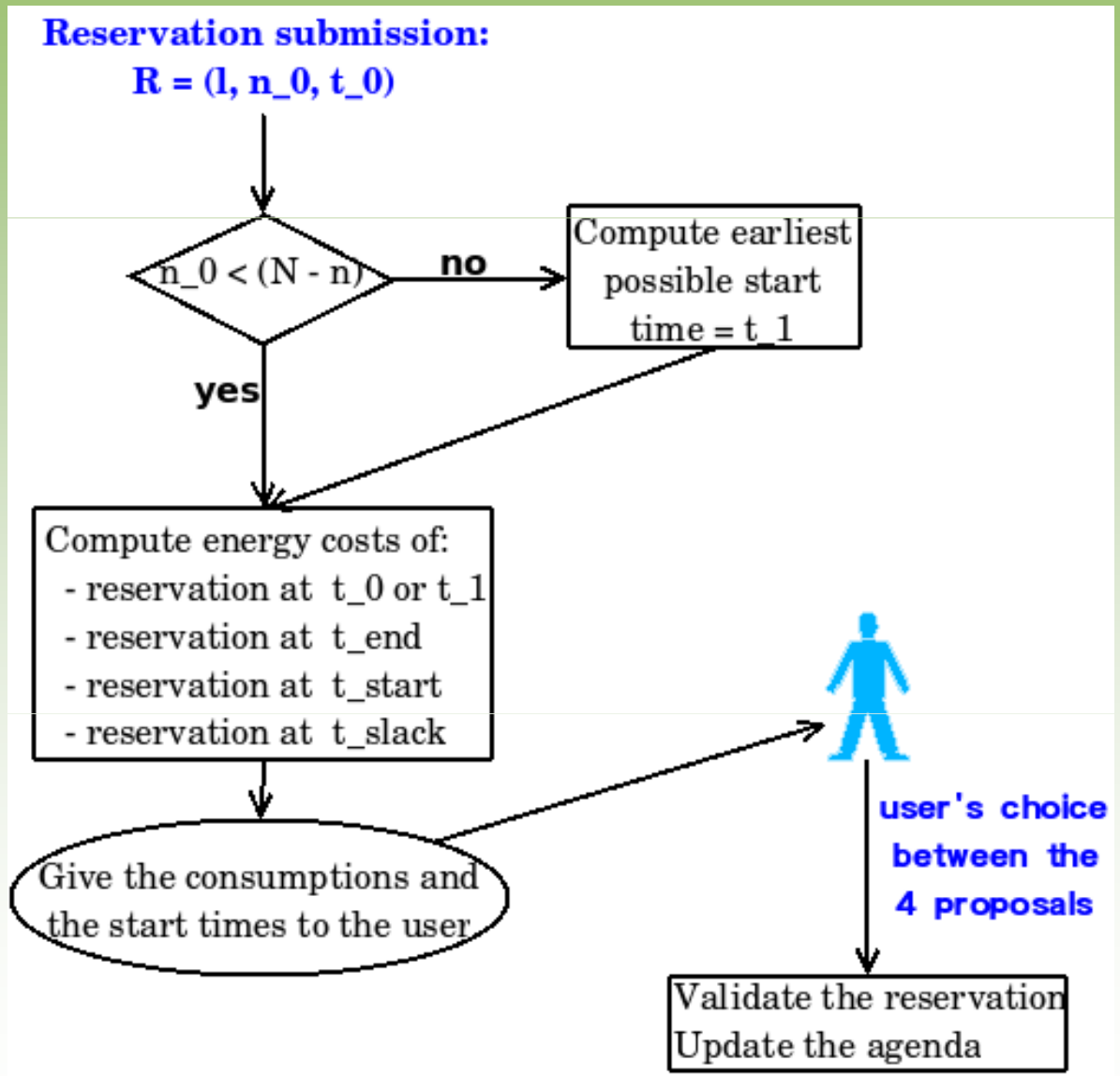


Role of T_s

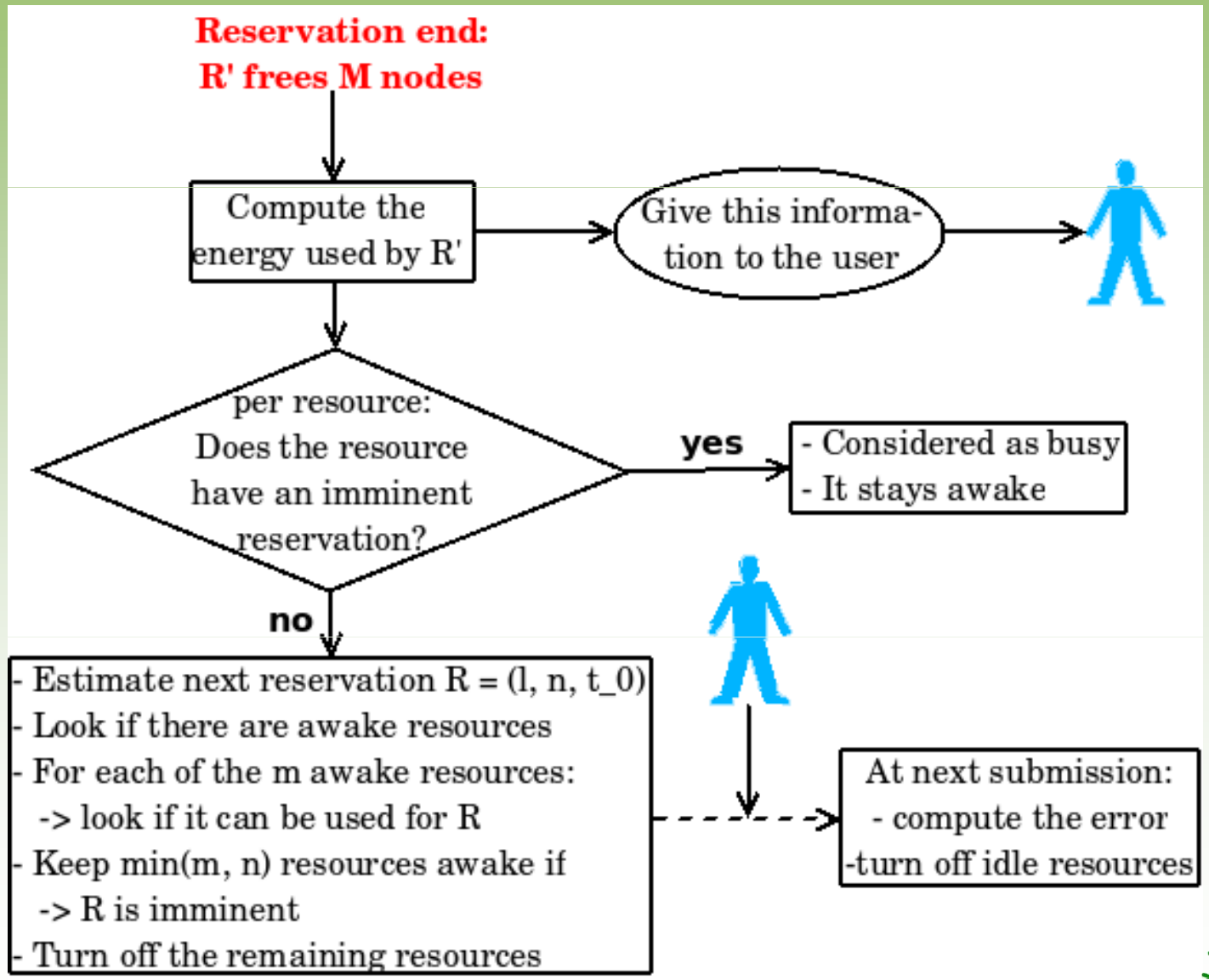
$$T_s = \frac{E_s - P_{OFF}(\delta_{ON\rightarrow OFF} + \delta_{OFF\rightarrow ON}) + E_{ON\rightarrow OFF} + E_{OFF\rightarrow ON}}{P_I - P_{OFF}} + T_r$$



Submission of a reservation



End of a reservation



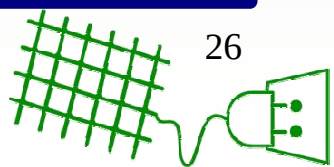
Predictions for greater energy savings

We should predict:

- The next reservation (length, size, start time)
- The next slack period
- The consumption of a given reservation

Solutions:

- Recent history (last reservations + feedback)
- History of the days before + feedback
- History of the user + type of ressource + number of ressources to switch on/off.



User behavior modelling

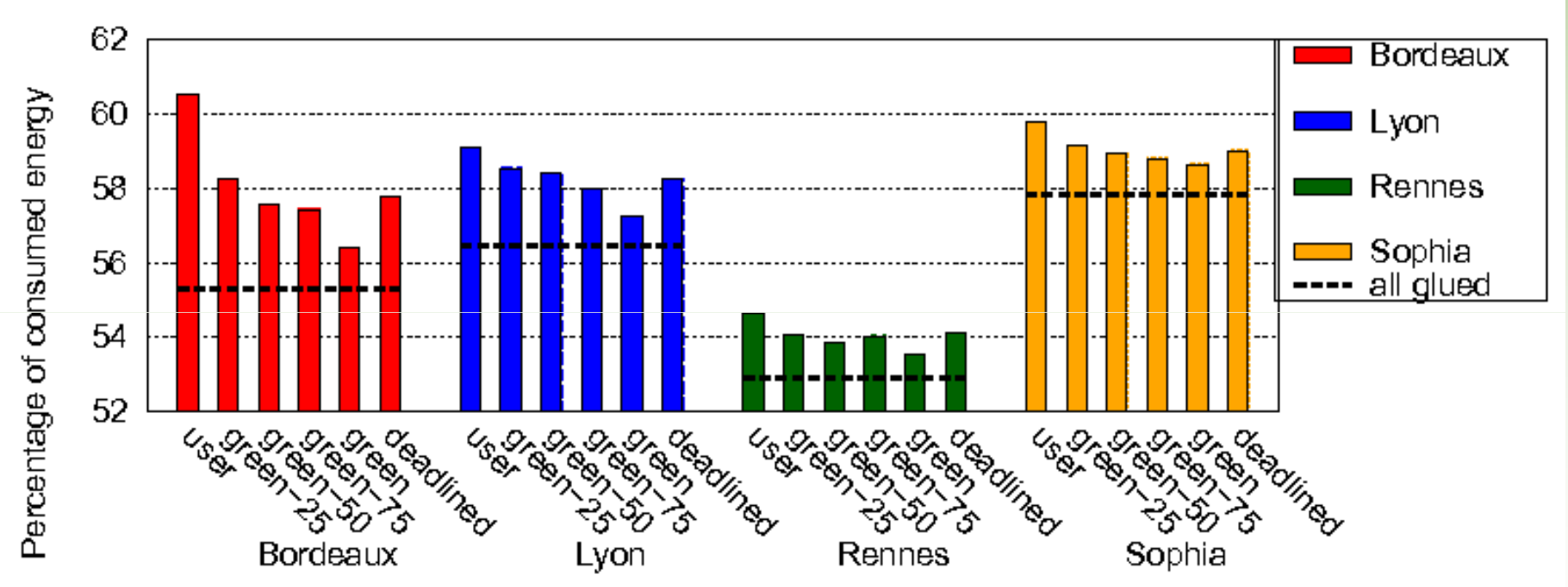
- **user**: wished date (or the nearest which is possible);
- **fully green**: solution that costs the less in terms of energy (the one where we switch on/off the smaller number of resources);
- **green-percentage-25**: 25% of fully green taken at random and user for the other ones;
- **green-percentage-50**: 50% of fully green and user for the other ones;
- **green-percentage-75**: 75% of fully green and user for the other ones;
- **deadlined**: fully green if it does not delay the reservation for more than 24 hours, otherwise user.

Evaluation of EARI

Replay on 4 different traces of Grid'5000 in 2007

100% = present consumption (all nodes always on)

all glued = theoretical lower bound



Greening Grid'5000

Demo

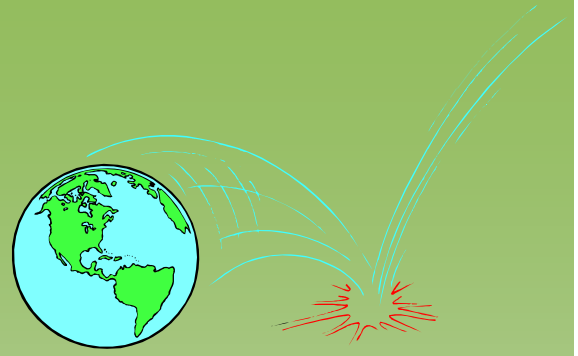
Conclusion and future works

Conclusions

Our Grid'5000 experience:

- Logs collecting with *oarstat*
- Usage over 2007 and 2008
- Consumption monitoring of 18 nodes
- Replay of the logs
- Modelling user's behavior

Soon on your G5K screens



- 'Green' advices and reservations (with Dynamic Voltage Scaling)
- Availability of the energy logs
- The energy profile per job available on the web
- Full Energy Monitoring of the complete Lyon site (135 nodes)

Thank you for your attention!

Questions?