PhD: Energy- and Carbon-aware HPC scheduling

Jules EVANS[1]

[1]Team SEPIA @ IRIT, CNRS, Exa-Soft, WP6

Supervisors : Georges DA COSTA et Patricia STOLF

Start date : 1er avril 2025

September 25th 2025

PhD Objectives

Context

- Clusters use a lot of electricity, some of them require dedicated power plants to run.

Objectives:

- Minimize electric consumption & carbon emissions of clusters
- At the scheduler level

What are the available leverages?

Leverages

Leverages for Minimizing Electricity Consumption:

- **Speed Scaling:** Limiting CPU frequency, thus reducing power draw.
- **Sleep Modes:** Switching off inactive nodes.
- **Power Capping:** Limiting maximum power draw.
- **Moldability:** More flexible workloads.

Leverages for Minimizing Carbon Emissions:

- **Temporal Load Shifting** Shifting tasks in time when electricity generation emits less carbon. Will mimic scheduling algorithms that dynamically load the cluster according to electricity pricing.

The contributions will be on scheduling algorithms.

Validation methods

Cluster Emulation

- Deploys containers or vms to host the machines required to make the cluster and the RJMS.
- Using `OAR` [1] (the one on `grid5000`)
  - 1 frontend
  - 1 server
  - some nodes
- Using NixOS-Compose [2] for deploying the containers/vms that has reproducibility built-in.
- Using GROG [3] for workload generation.

Validation pipeline

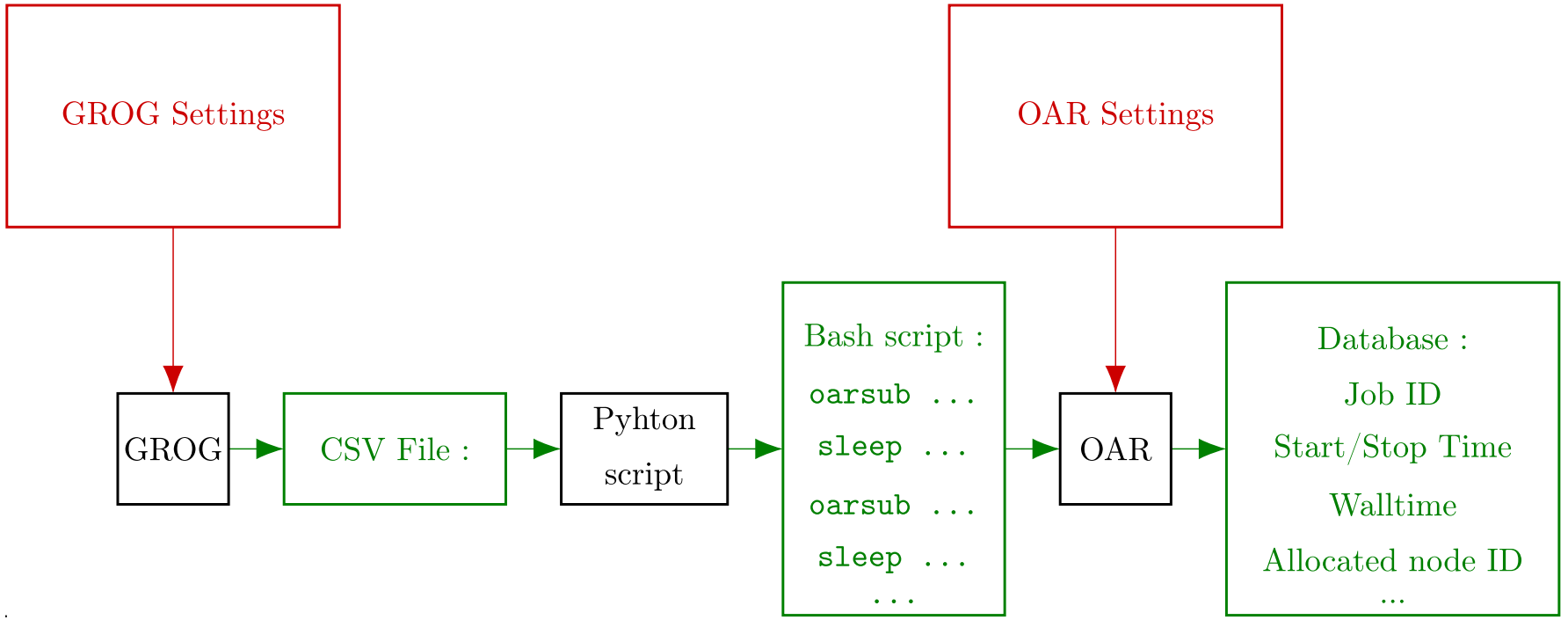

Figure: Validation pipeline

Results

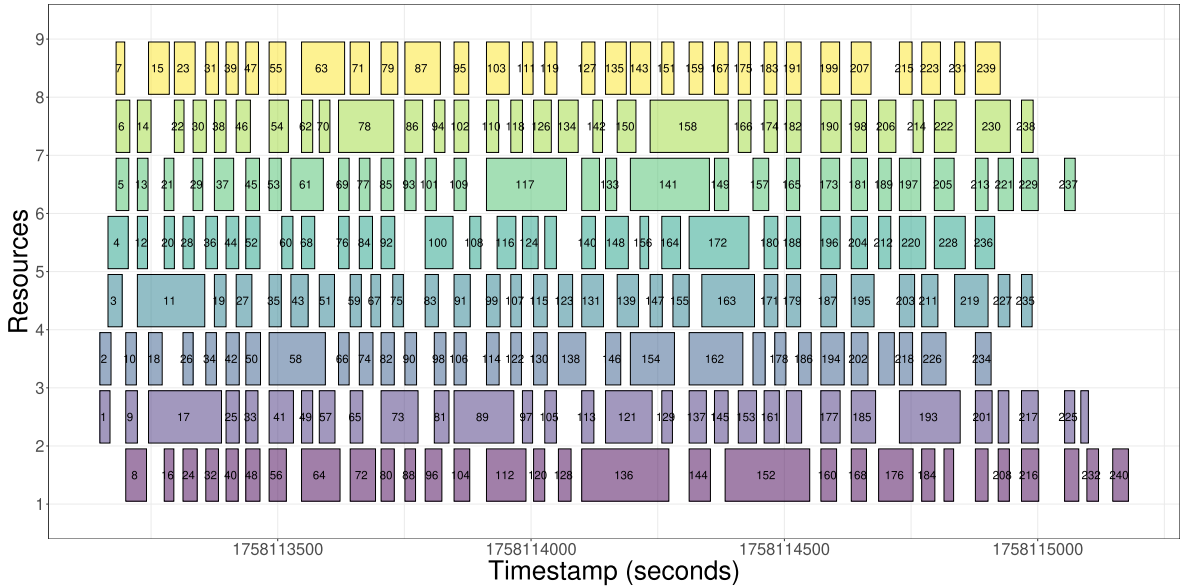

Figure: Gantt chart of the round-robin scheduler on 8 virtual nodes.

- **Duration:** 24 hours
- **Density:** 10 jobs per hour
- **Speedup:** 100
- **# Nodes:** 8
- **Algorithm:** round-robin

Thanks for listening!

Acknowledgements

As part of the “France 2030” initiative, this work has benefited from a national grant managed by the French National Research Agency (Agence Nationale de la Recherche) attributed to the Exa-Soft project of the NumPEx PEPR program, under the reference ANR-22-EXNU-0003.

Sources

[1] N. Capit *et al.*, “A batch scheduler with high level components.” [Online]. Available: https://arxiv.org/abs/cs/0506006

[2] Q. Guilloteau, J. Bleuzen, M. Poquet, and O. Richard, “Painless Transposition of Reproducible Distributed Environments with NixOS Compose,” in *CLUSTER 2022 - IEEE International Conference on Cluster Computing*, Heidelberg, Germany, Sep. 2022, pp. 1–12. [Online]. Available: https://hal.science/hal-03723771

[3] G. Da Costa, L. Grange, and I. de Courchelle, “Modeling, classifying and generating large-scale Google-like workload,” *Sustainable Computing: Informatics and Systems*, vol. 19, pp. 305–314, 2018, doi: https://doi.org/10.1016/j.suscom.2017.12.004.

Grid5000

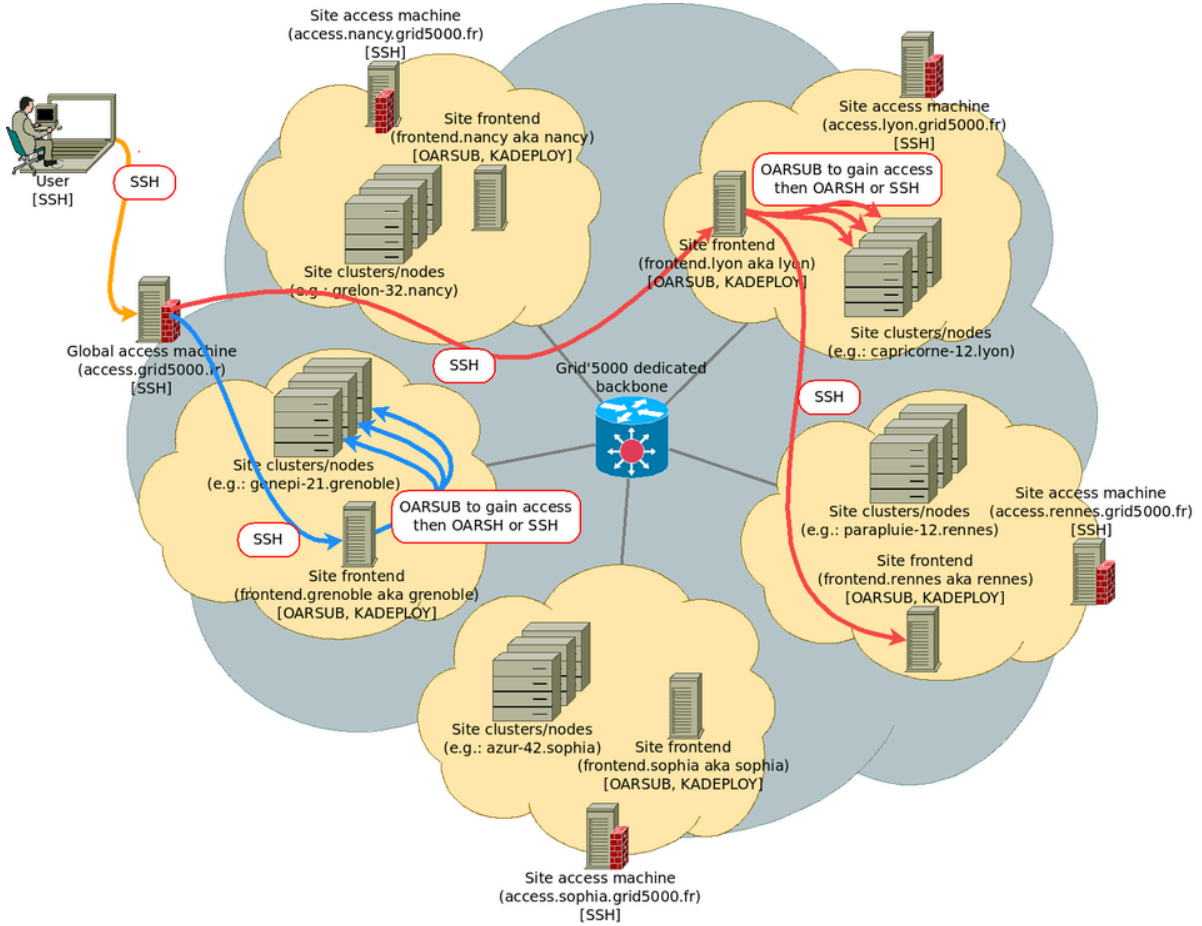

Figure: Grid5000's structure