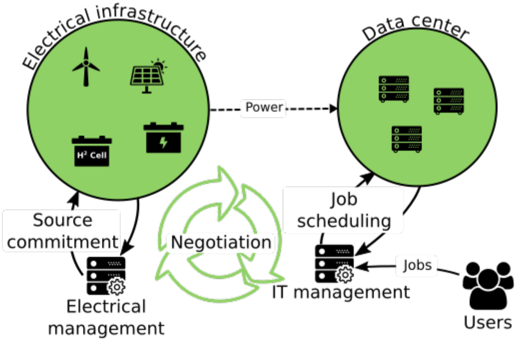
Impact of flexibility on QoS and power consumption in a green datacenter

Damien Landré [1, 2], Laurent Philippe [1], Jean-Marc Pierson [2]

[1] FEMTO-ST institute, Besançon
[2] IRIT, Toulouse

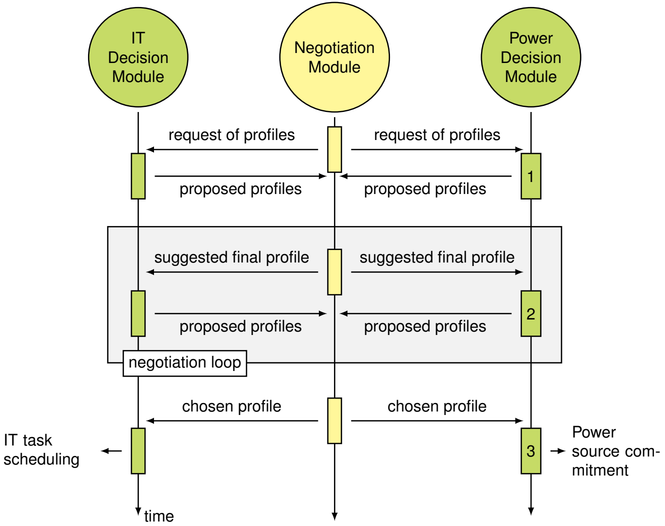
DATAZERO2: the big picture



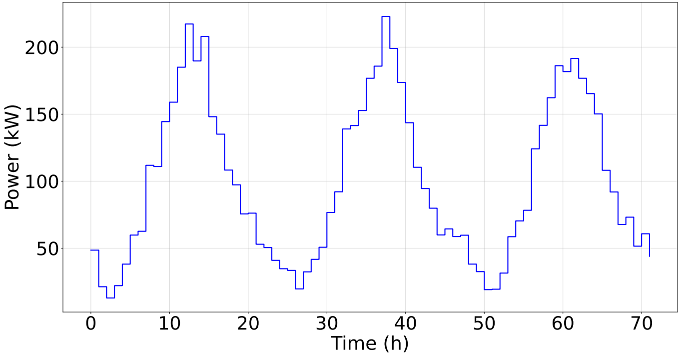
Partners: IRIT, femto-st SCIENCES & TECHNOLOGIES, Laplace, EATON Powering Business Worldwide

Funded by anr

DATAZERO2: the big picture



DATAZERO2: the big picture



DATAZERO2: the big picture

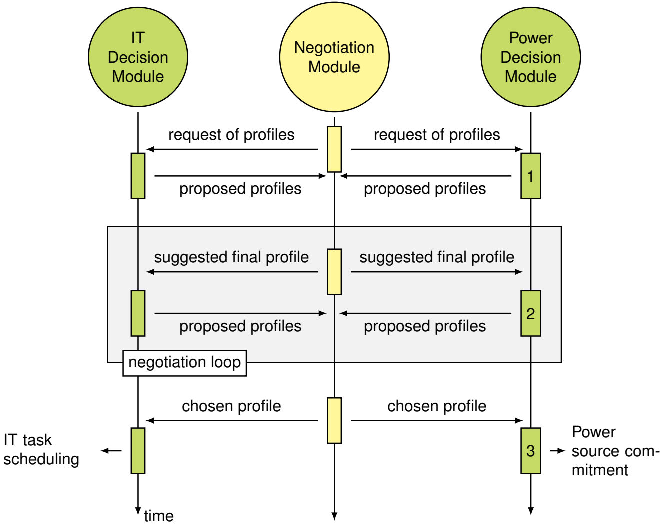


Table of contents

Table of contents

Summary results

- BS algorithm to generate the initial power profile in a reasonable computation time.
- MILP and 3 non-trivial heuristics that computes the maximum processing power in the heterogeneous case.
- BPP (Balance Power-Performance) heuristic seems the most suitable to solve this problem in a reasonable time.

Articles

- Canon et al., “Assessing Power Needs to Run a Workload with Quality of Service on Green Datacenters”, Euro-Par 2023
- Canon et al., “Évaluation de la consommation d’énergie nécessaire à l’exécution d’un workload dans un datacenter vert”, COMPAS 2023

Table of contents

Objectives

- Challenge the BS algorithm, the MILP and the heuristics with a real workload to assess them in a practical case.
- Analysis of the power required to process the workload.
- Comparison of heuristics in terms of total energy consumed, including energy to switch-on and switch-off the machines, by the machine configurations required to process the workload.
- Impact of deadlines and computing power on the QoS.

Table of contents

Assumptions and settings

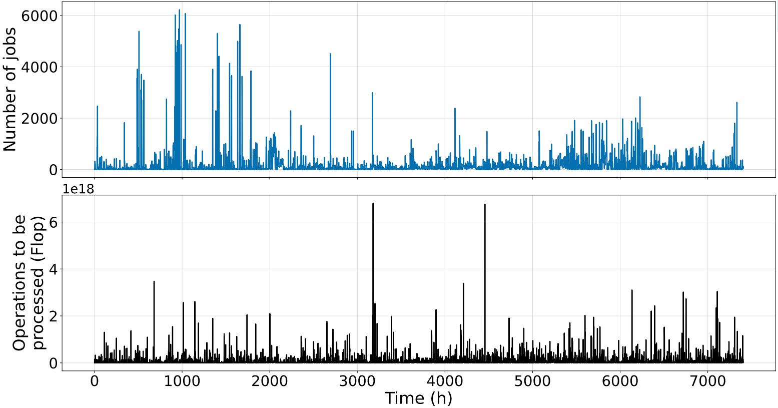
- Only Metacentrum[1] non-exotic nodes[2] are selected (partitions 7, 9, 11, 14, 15, 18 and 19 are kept for a total of 305 machines).
  - 12 to 16 CPU cores.
  - 12 to 134 GB of memory.
  - No GPU.
- Period from July 30, 2013 to September 19, 2014, totalling 7406 hours. During this period, the set of selected partitions remains the same[3].
- Load parts are derived from Metacentrum jobs.
  - Submit time: Same as for jobs.
  - Deadline: Sum of the wait time and the runtime of jobs.
  - Number of operations: Product of the number of cores allocated to the job, its runtime and a factor determining the number of Flops per core per second (9.72 GFlops).
- Target rate of operations processed: 1.

[1] Klusáček, Tóth, and Podolníková, "Real-life experience with major reconfiguration of job scheduling system".

[2] *HPC and HTC tutorial*.

[3] *The Metacentrum 2 log*.

Simulation data



Simulation data

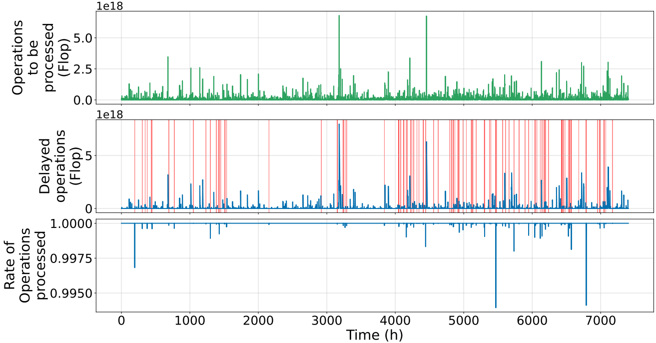
Table: Machine types[4] and their characteristics[5]

Machine name	No.	No. of states	Static power (W)	$\max(g_{max_j})$ (GFlops)
Taurus	113	13	93.0	220.80
Parasilo	112	12	94.1	614.40
Graoully	7	14	98.2	614.40
Grimoire	47	14	121.2	614.40
Grisou	26	14	90.5	614.40

[4] *Clusters Grid5000*.
[5] *Energy saving in large scale distributed platforms – Energumen*.

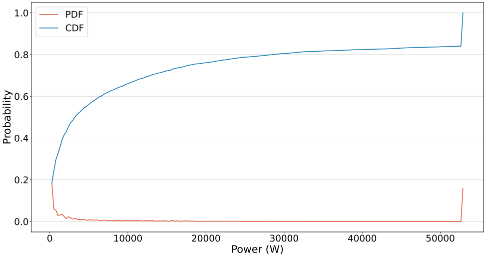
Table of contents

Operations Processed, Delayed, Killed



- Rate of operations killed: $\sim 2.5 \times 10^{-9}$.
- Rate of operations processed: > 0.99.
- QoS is met for more than 98% of time intervals.

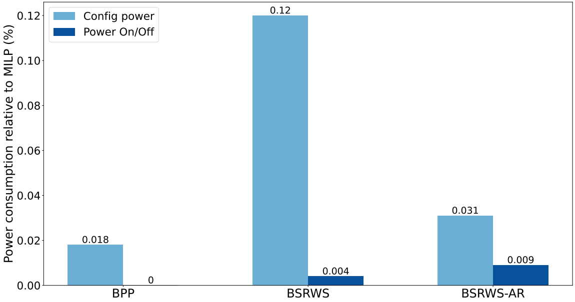
Power demand



- power consumption is 0 for more than 18% of the time intervals.
- Power consumption is $< 6\%$ of the maximum power that can be requested by the machines for 50% of the time intervals.
- The maximum power required by the machines is reached for 16% of the time intervals.
- Metacentrum HPC workload is more sporadic than cloud workloads[6].

[6] Di, Kondo, and Cirne, "Characterization and comparison of cloud versus grid workloads".

Heuristics power consumption



- Total energy consumed with the BS algorithm using MILP: $\sim$ 102 MWh.
- $\sim 2MWh$ to switch-on and -off machines ($\sim$ 2% of all energy consumed).
- BPP consumes the least excess power with $\sim$ 0.02% (18 kW).

Flexibility

$$d_i = sub_i + run_i \quad => \quad d_i = sec$$

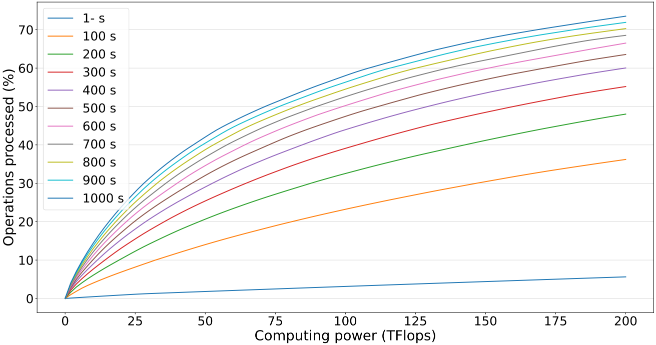


Table of contents

Conclusion & Perspectives

Conclusion

- Present results consolidate the choice of the BPP heuristic[7].
- Sporadic nature of the HPC workload was observed.
- Power consumed for switching-on/off machines: $\sim$ 2% of total consumption.
- Minor variation in the workload flexibility results in a major variation in the computing power required.

Perspectives

- Implement and test the BS algorithm with the BPP heuristic in a real environment (DATAZERO2[8] platform).
- Assess its resilience against uncertainty.

[7] Canon et al., "Assessing Power Needs to Run a Workload with Quality of Service on Green Datacenters".

[8] Pierson et al., "DATAZERO: Datacenter With Zero Emission and Robust Management Using Renewable Energy".

Thank you for your attention

Canon, L.-C. et al. “Assessing Power Needs to Run a Workload with Quality of Service on Green Datacenters”. In: *European Conference on Parallel Processing*. Springer. 2023, pp. 229–242.

— .“Évaluation de la consommation d’énergie nécessaire à l’exécution d’un workload dans un datacenter vert”. In: *Conférence d’informatique en Parallélisme, Architecture et Système, COMPAS 2023*. 2023.

Clusters Grid5000. https://www.grid5000.fr/. [Online; accessed 21-October-2022]. 2022.

Di, S., D. Kondo, and W. Cirne. “Characterization and comparison of cloud versus grid workloads”. In: *2012 IEEE International Conference on Cluster Computing*. IEEE. 2012, pp. 230–238.

Energy saving in large scale distributed platforms – Energumen. https://anr.fr/Project-ANR-18-CE25-0008. [Online; accessed 20-November-2022]. 2018.

HPC and HTC tutorial. https://www.grid5000.fr/w/HPC_and_HTC_tutorial. Accessed: 2023-11-23.

- Klusáček, D., Š. Tóth, and G. Podolníková. “Real-life experience with major reconfiguration of job scheduling system”. In: *Job Scheduling Strategies for Parallel Processing: 19th and 20th International Workshops, JSSPP 2015, Hyderabad, India, May 26, 2015 and JSSPP 2016, Chicago, IL, USA, May 27, 2016, Revised Selected Papers 19*. Springer. 2017, pp. 83–101.
- Pierson, J.M. et al. “DATAZERO: Datacenter With Zero Emission and Robust Management Using Renewable Energy”. In: *IEEE Access* 7 (2019). DOI: 10.1109/ACCESS.2019.2930368.
- *The Metacentrum 2 log*. https://www.cs.huji.ac.il/labs/parallel/workload/l_metacentrum2/index.html. Accessed: 2023-10-26.