Green Days@Toulouse v2.0 (Edition 2018)

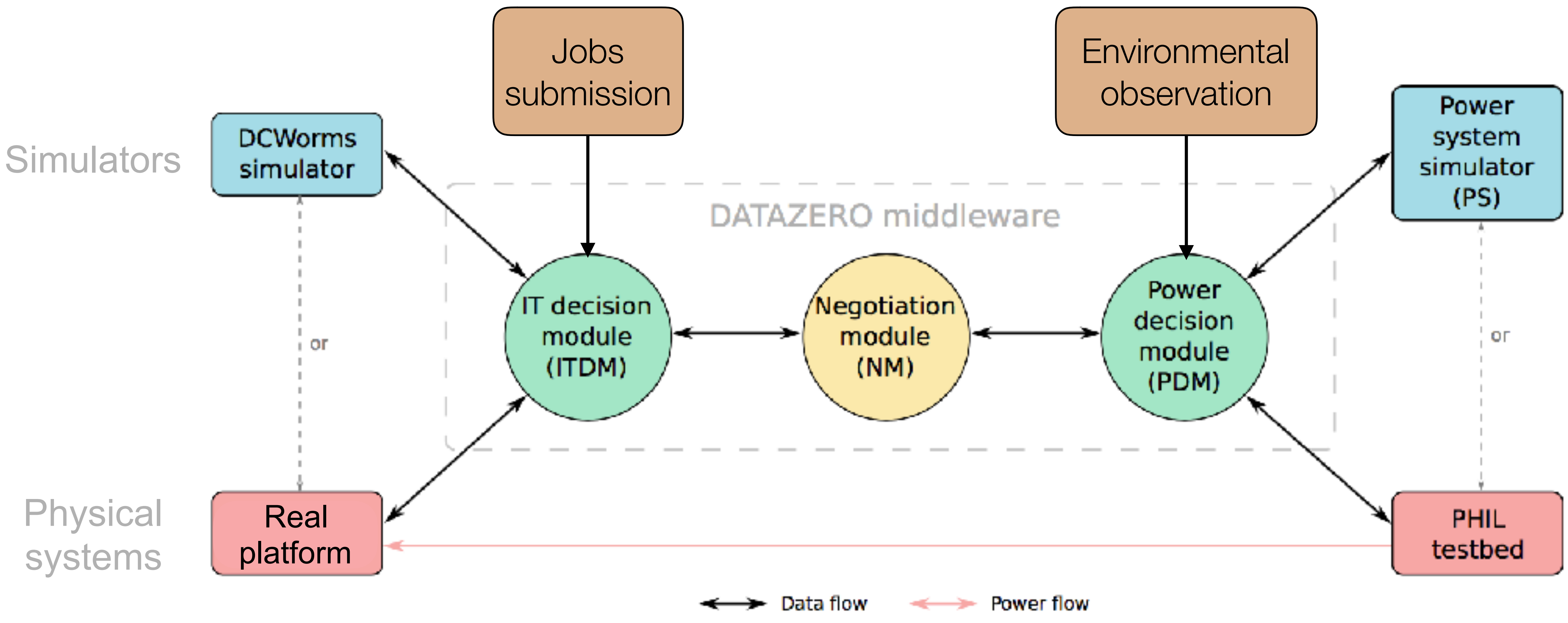
IT Optimization Under Renewable Energy Constraint

Gustavo Rostirolla

gustavo.rostirolla@irit.fr

Stephane Caux, Paul Renaud-Goud, Gustavo Rostirolla, Patricia Stolf. *IT Optimization for Datacenters Under Renewable Power Constraint (regular paper). Euro-Par, Turin, 27/08/2018 - 31/08/2018*.

DATAZERO Project



Credits: Robin Roche

The work presented is supported by the French ANR DATAZERO project ANR-15-CE25-0012.

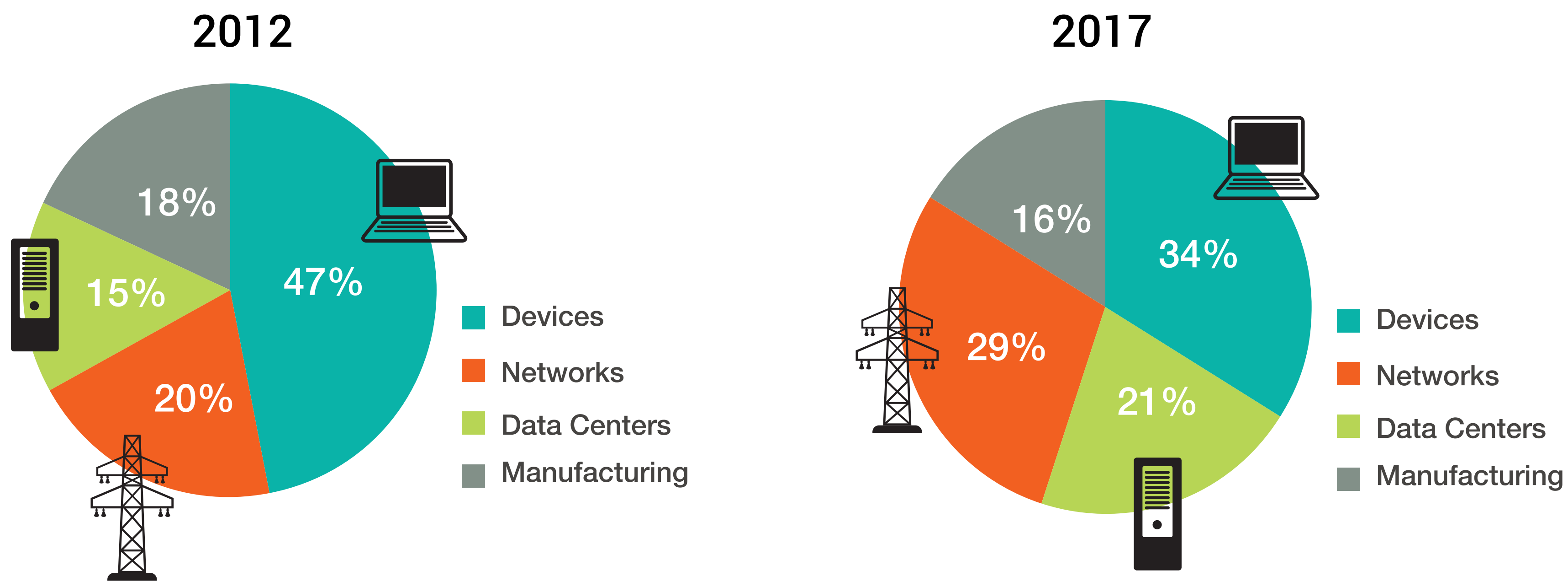
Agenda

- Introduction
- IT Optimization Module
- Methodology
- Results
- Conclusion
- Future Works

Source: www.toulouse-tourisme.com

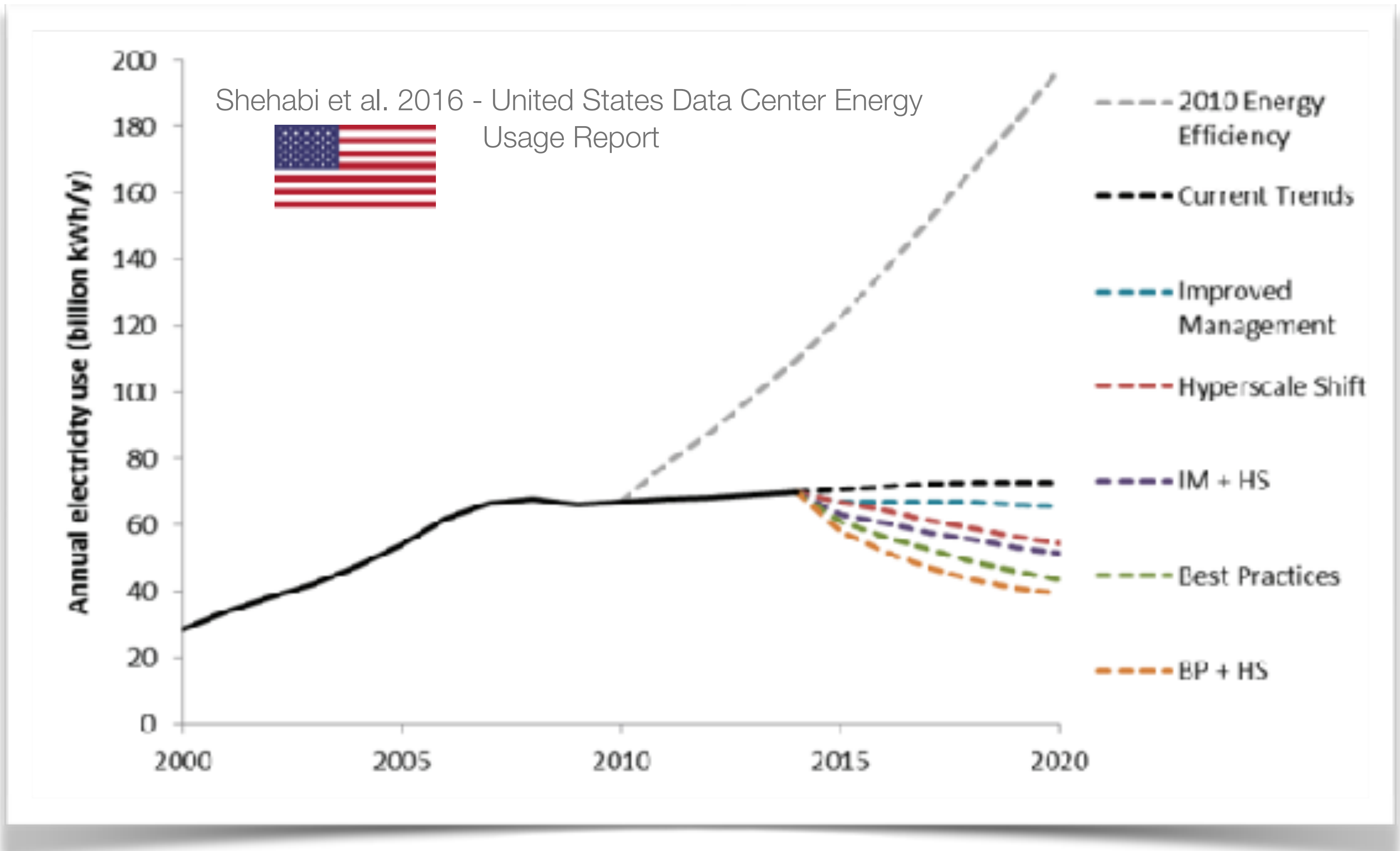
Introduction

- Data centers are known as one of the big players when talking about energy consumption;
- In 2006, were responsible for 61.4 billion kWh in the United States;
- In 2010 about 1.3% of world's electricity;



Cook, Gary, et al. "Clicking Clean: Who is winning the race to build a Green Internet?." Greenpeace Inc., Washington, DC (2017).

Introduction

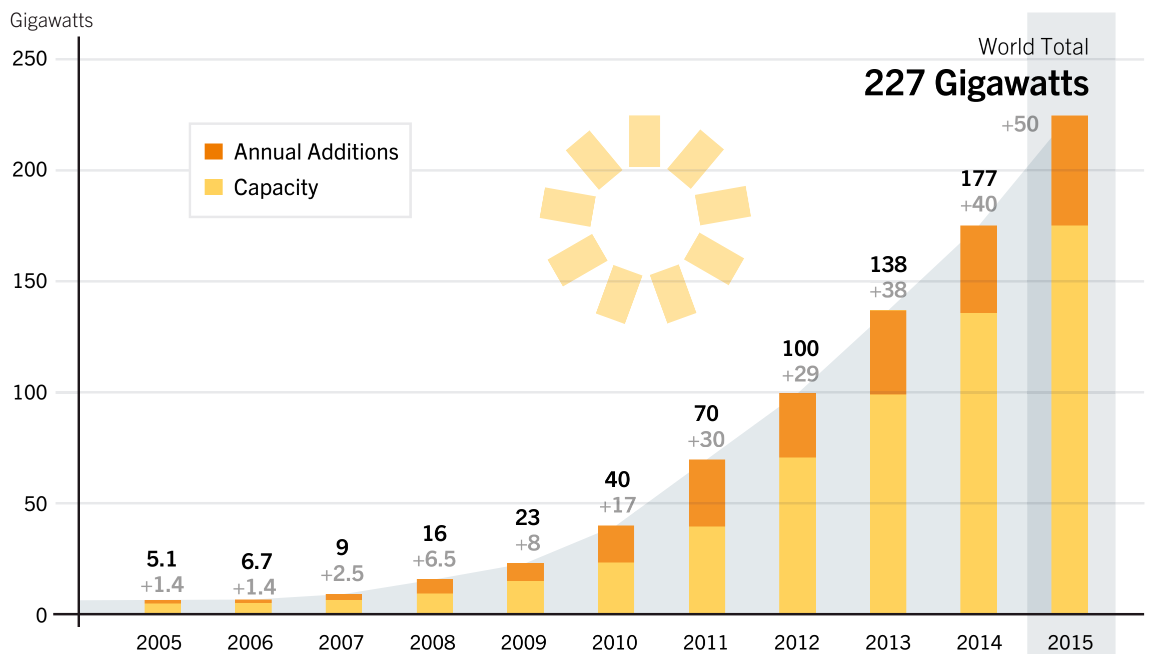


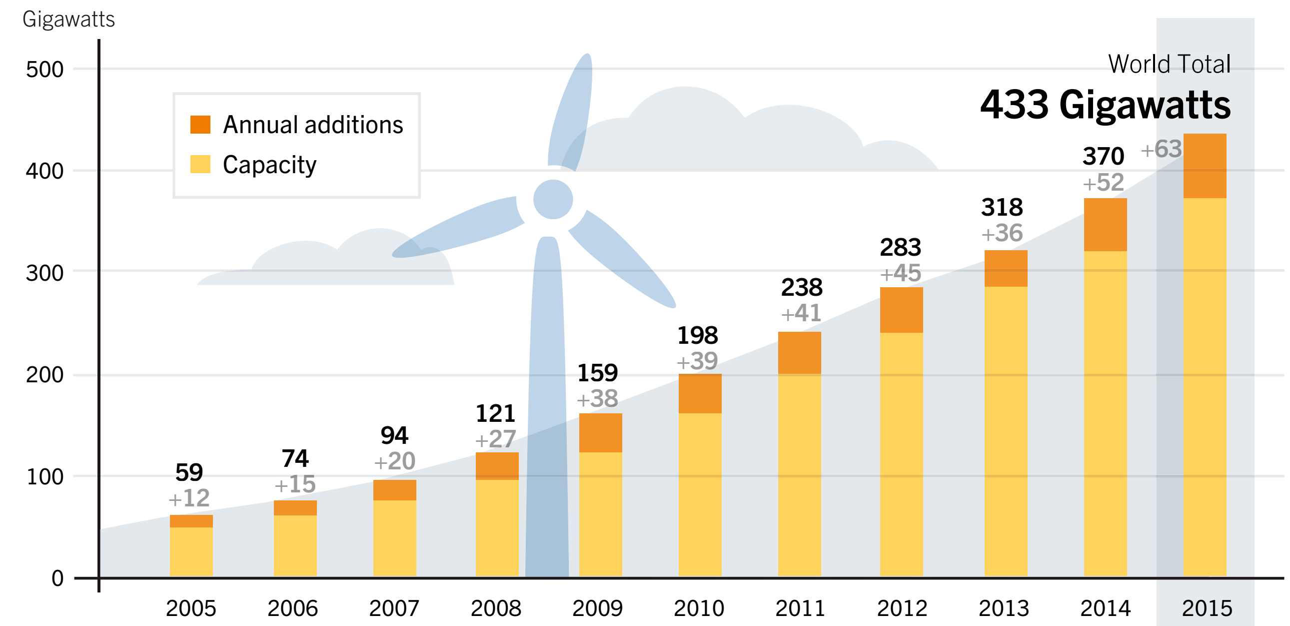
~73 billion kWh in 2020

- In the last years, the use of cloud computing has been the basis of data centers, either in a public or private fashion.

Introduction

- This migration to cloud computing increases the concern about power utilization, especially when considering renewable energy sources and its oscillation over time;

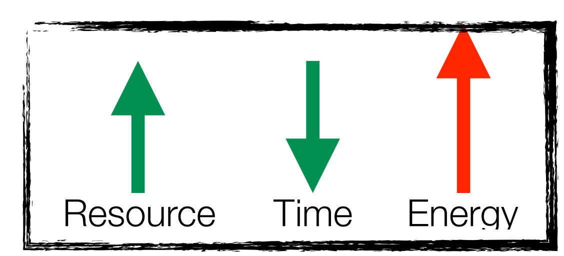




Adib, Rana, et al. "Renewables 2016 global status report." Global Status Report Renewable Energy Policy Network for the 21st Century (REN21) (2016): 272.

- Tasks submitted by users needs to be executed inside a time interval (release time and due date);

- But: When? Where? At which speed/frequency?

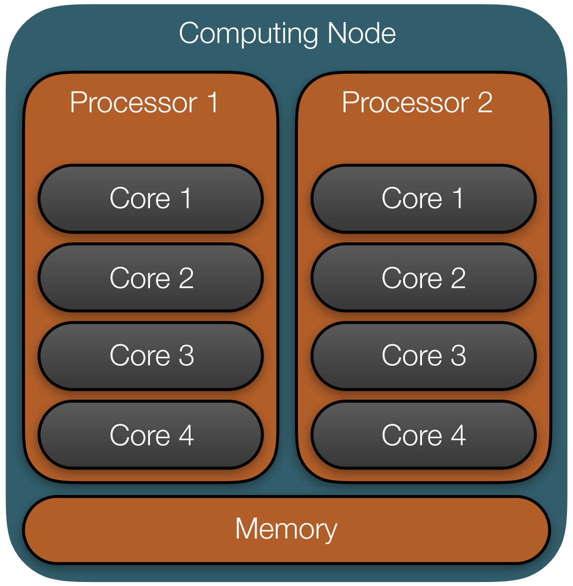



Backspace
Optimization
Shift

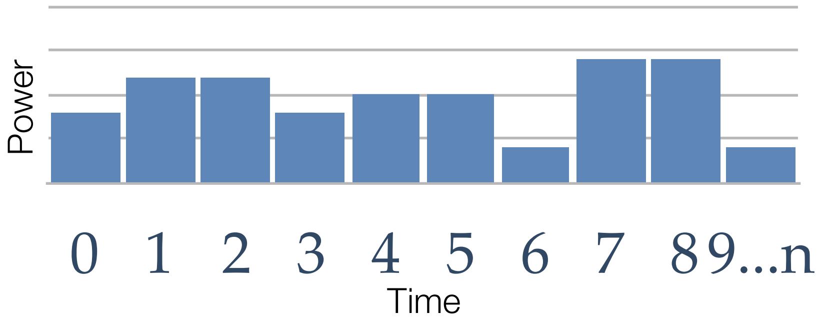
IT Optimization Module

- Set of Jobs J: $J_j=[t_{j,i};pe_{j,i};mem_{j,i}]$
    - $t_{j,i}=[t_{relase\ j,i};t_{duedate\ j,i};t_{durationj,i}]$
    - $mem_{j,i}$ = Memory requested by task $_{j,i}$
    - $pe_{j,i}$ = Number of processing elements requested by task $_{j,i}$

- Set of Machines M: $M_i=[npe_i;mem_i;P_{min\ i};P_{max\ i}]$
    - $[f_i]$ set of frequencies available
    - mem_i: Memory available in node
    - $P_{min\ i}$:Power when node is idle
    - α_i: Coefficient dependent on the machine
    - npe_i: Number of processing elements
    - $P_{max\ i}$: $g(P_{min\ i};f_{i,l};\alpha_i)$ Power with processing element at 100%



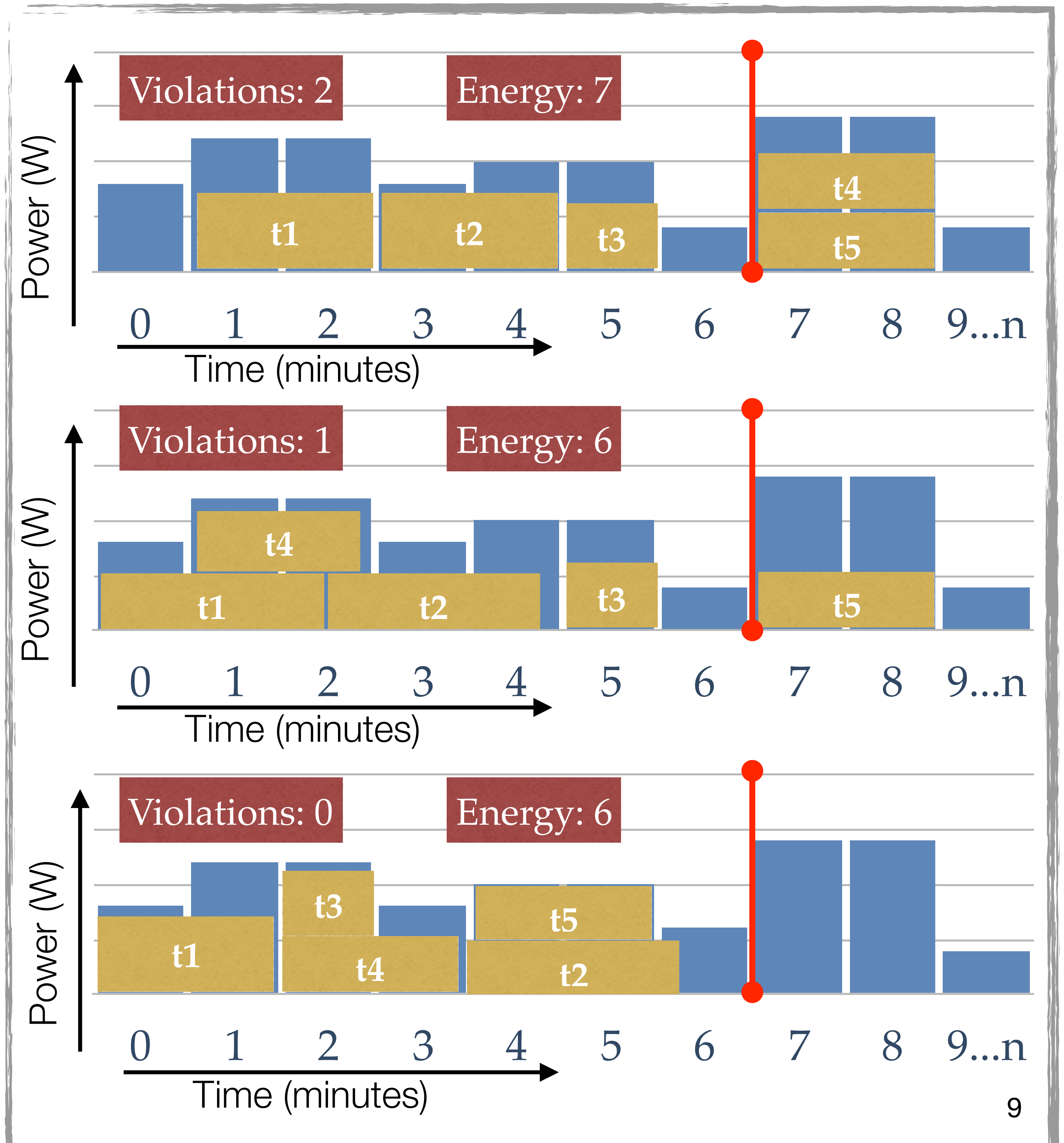
- Discrete power curves $P_{available}(t)$: power available at instant t



IT Optimization Module

Power Metric Task

- Output:
  - Which task will run where, when, at which frequency;
  - Constraints {Power, CPU, Memory}
  - Translated as a set of scheduling possibilities in the form of a power profile;
  - Associated with metrics (energy, due date violations…)



IT Optimization Module

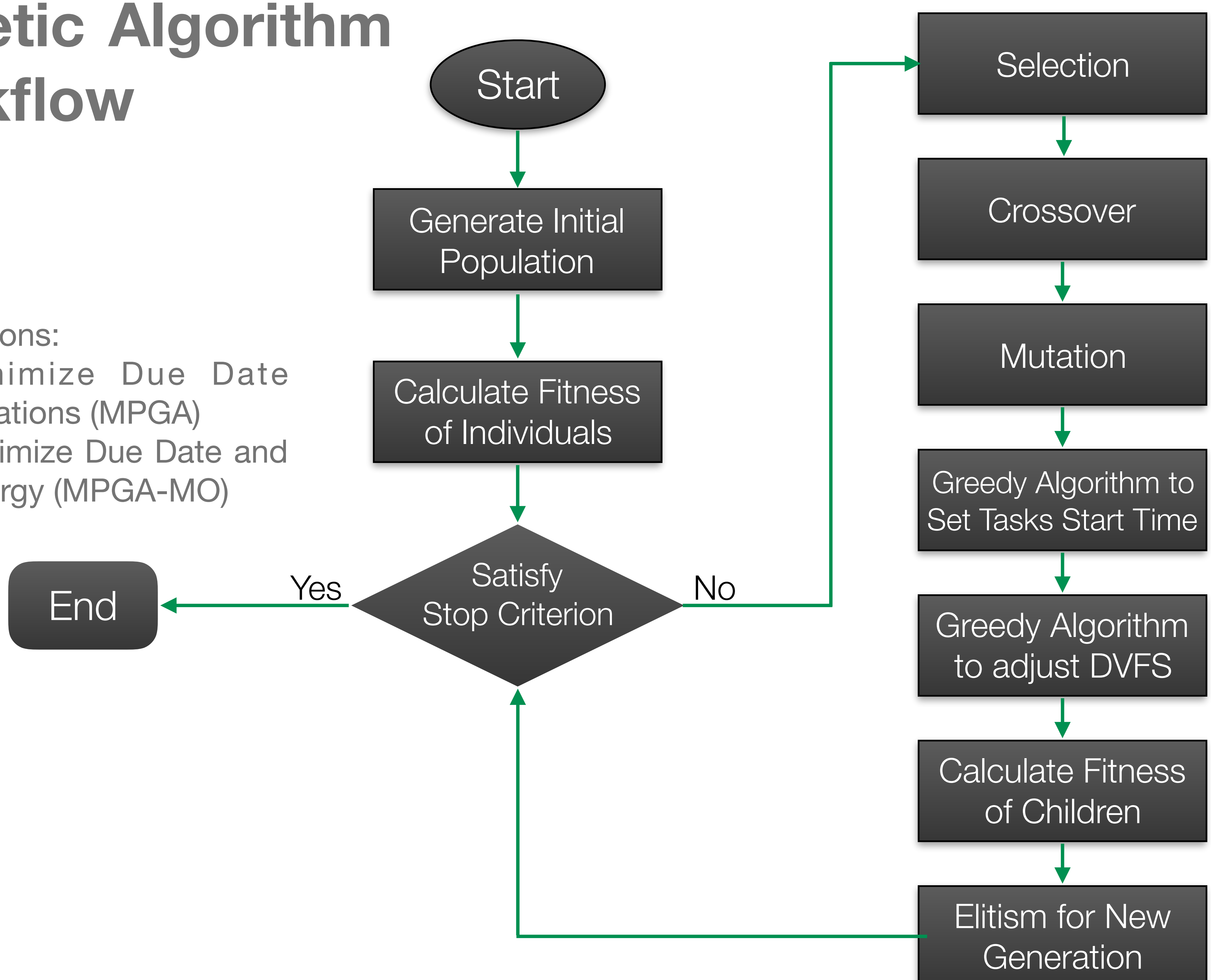
- **Greedy Heuristic (Best Fit)**:
  - Fast scheduling decisions;
  - Easy implementation;
  - Tasks can be sorted by arrival time, due date…;
  - Tasks are scheduled in a local optimal, limited by the power curve received;
  - The combinations of choices locally optimal do not always lead to an overall optimum.

- **Meta-heuristic (Genetic Algorithm)**:
  - Allows to produce a large number of adapted solutions;
  - Makes it possible to approach an optimum solution;
  - Slow execution time;
  - Difficulties in setting parameters (crossover, mutation rate, population size, selection method).

IT Optimization Module

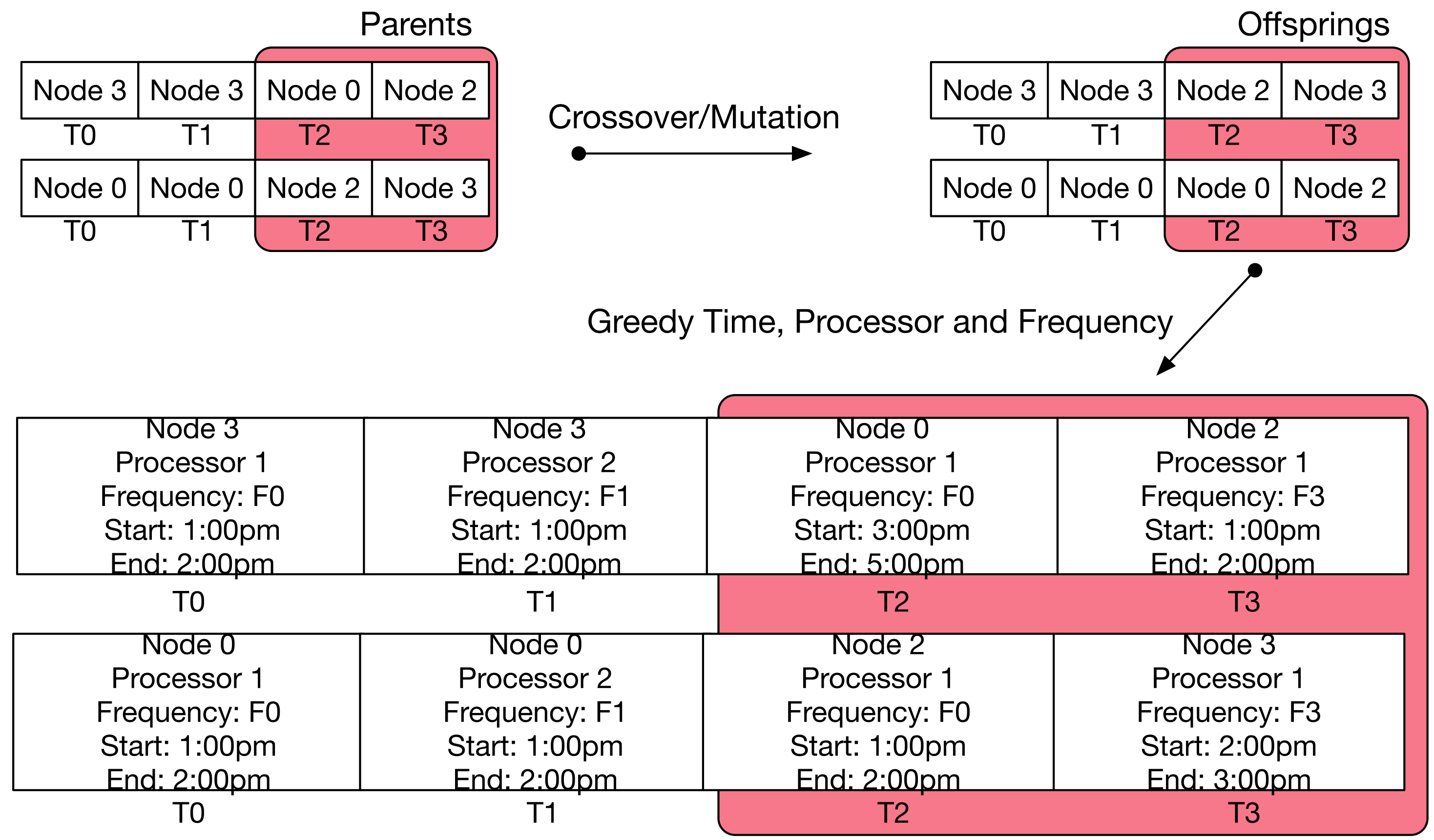
Genetic Algorithm Workflow

- 2 Variations:
  - Minimize Due Date violations (MPGA)
  - Minimize Due Date and Energy (MPGA-MO)

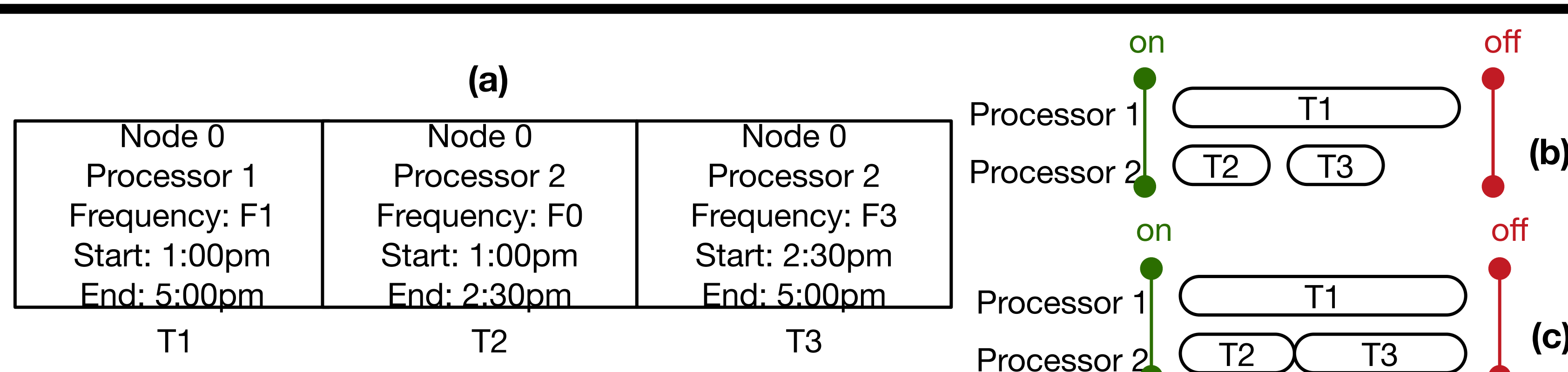


IT Optimization Module

Genetic Algorithm:

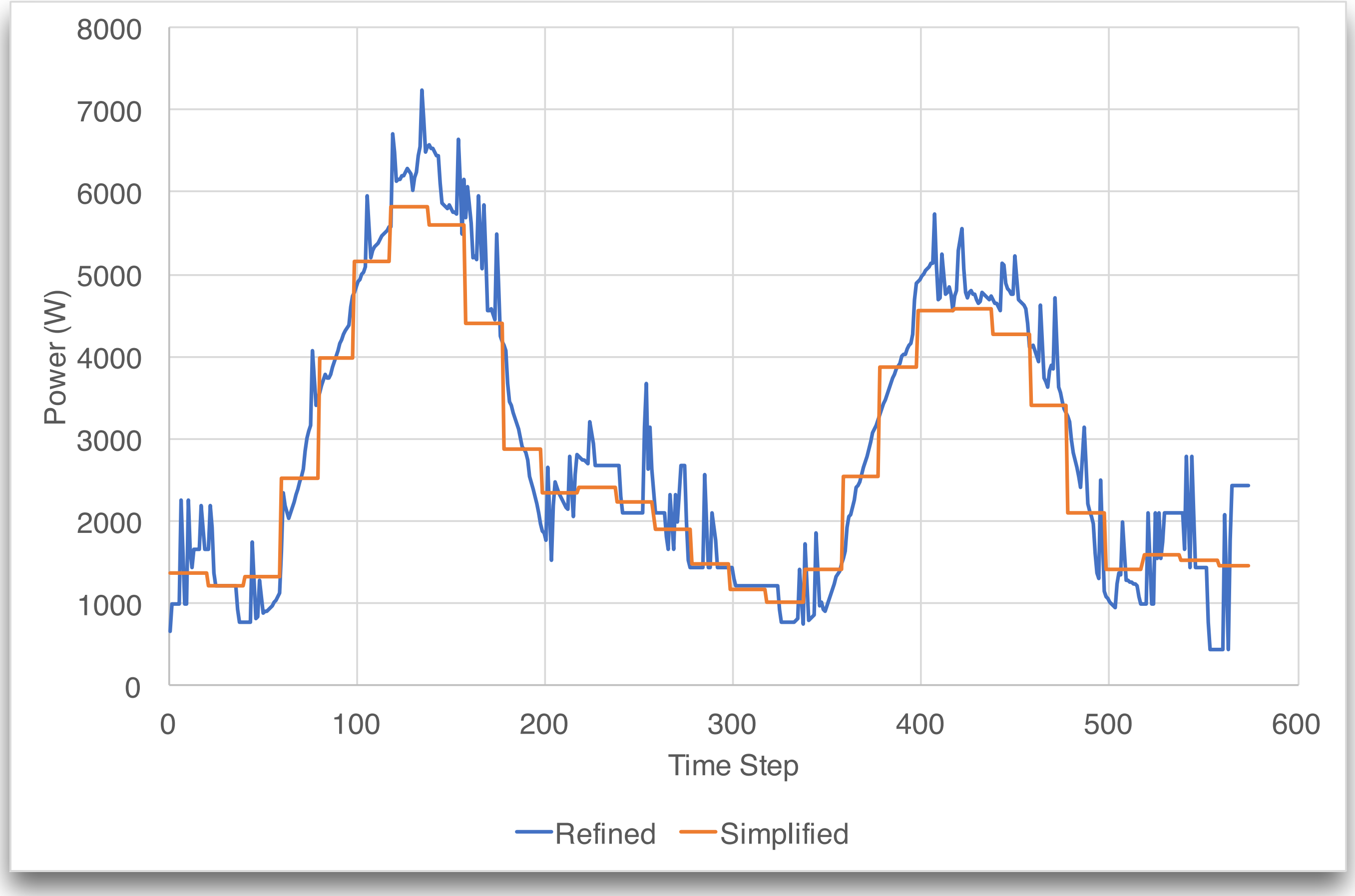


DVFS:

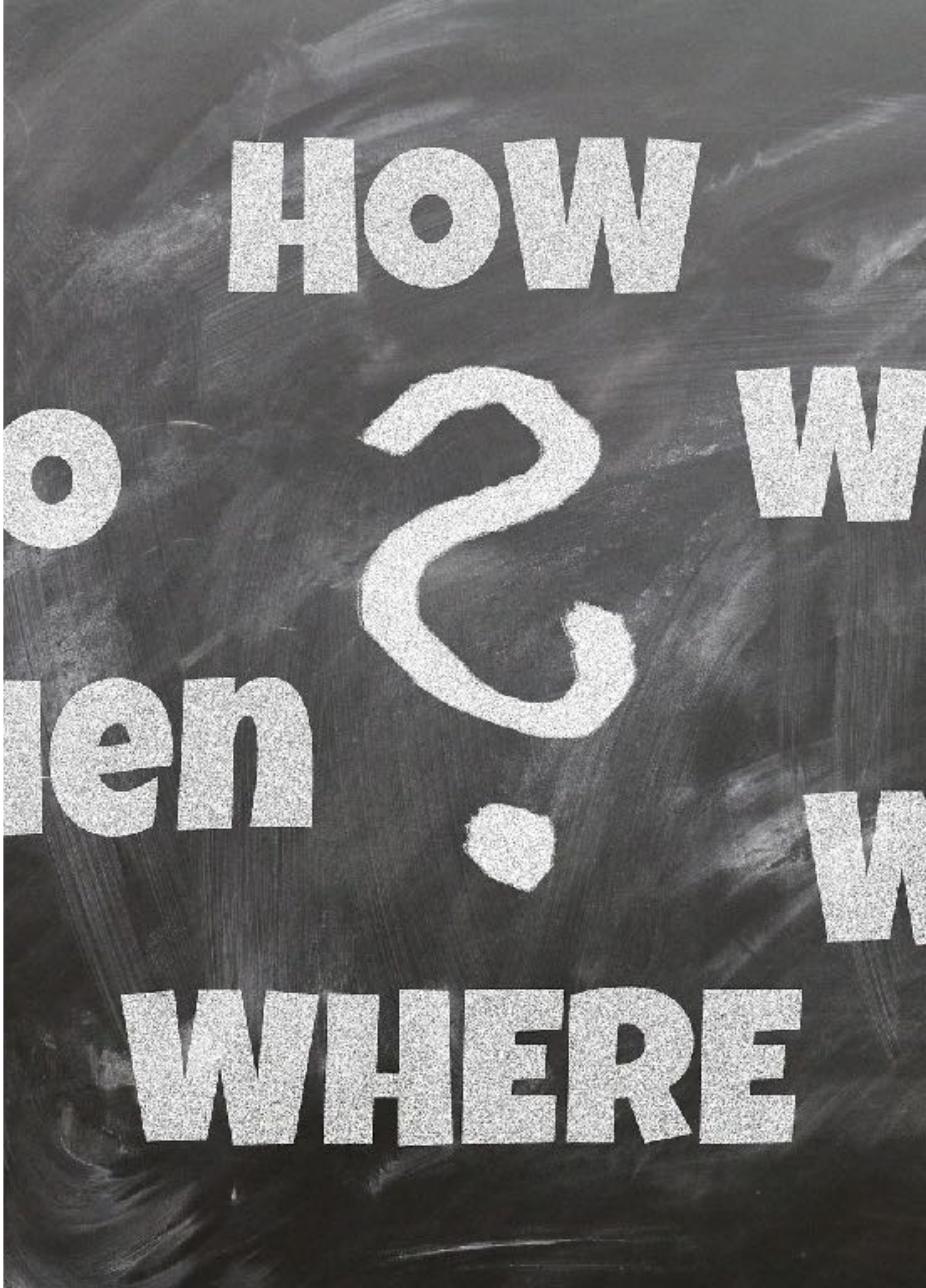


IT Optimization Module

- Two execution phases for Genetic Algorithm to improve execution time:
  - First phase provides an initial placement of tasks respecting a simplified power curve;
  - Second phase uses the power prediction with all variations to improve this initial placement, allowing the scheduling to take profit of power peaks.



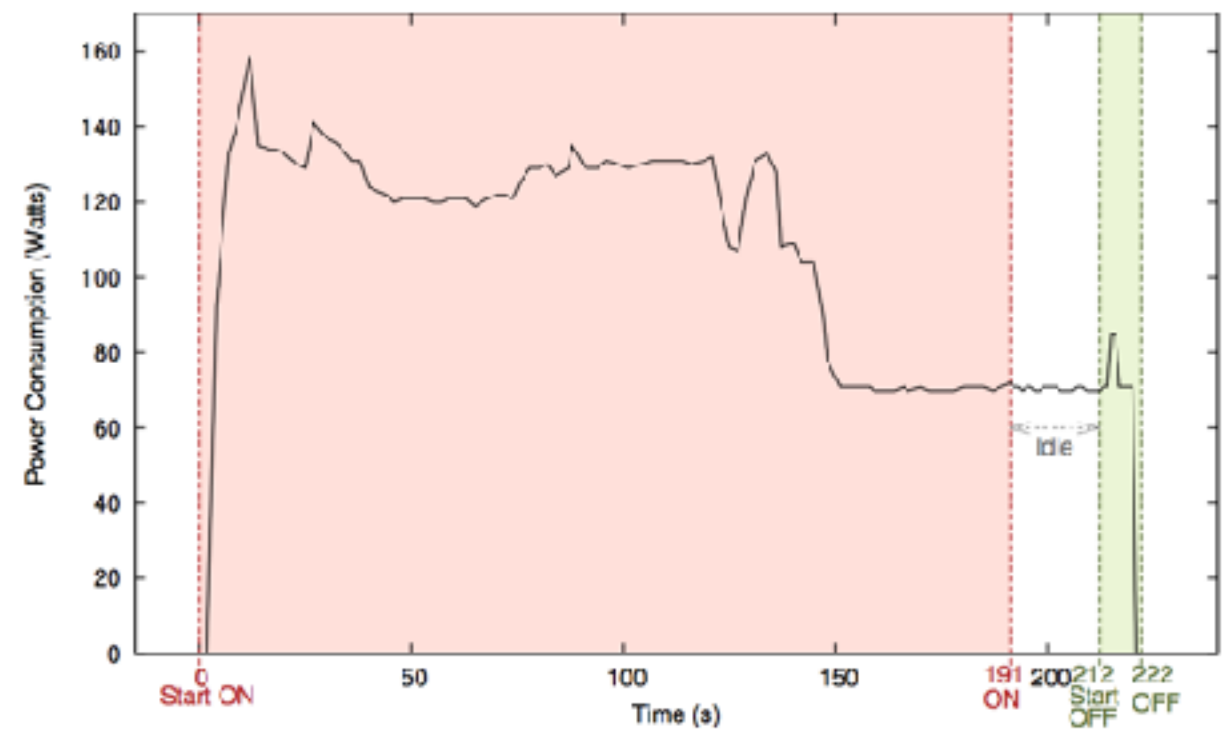
Evaluation



Evaluation

- Computing Resources:
  - Based in Paravance and Taurus (Grid5000)
  - 30 Nodes x 2 Processors x 5 Frequencies (1.2 to 2.4 Ghz);
  - + Overhead of turning on/off a node.

Source: Villebonnet, V. (2016). Scheduling and Dynamic Provisioning for Energy Proportional Heterogeneous Infrastructures (Doctoral dissertation, Université de Lyon).



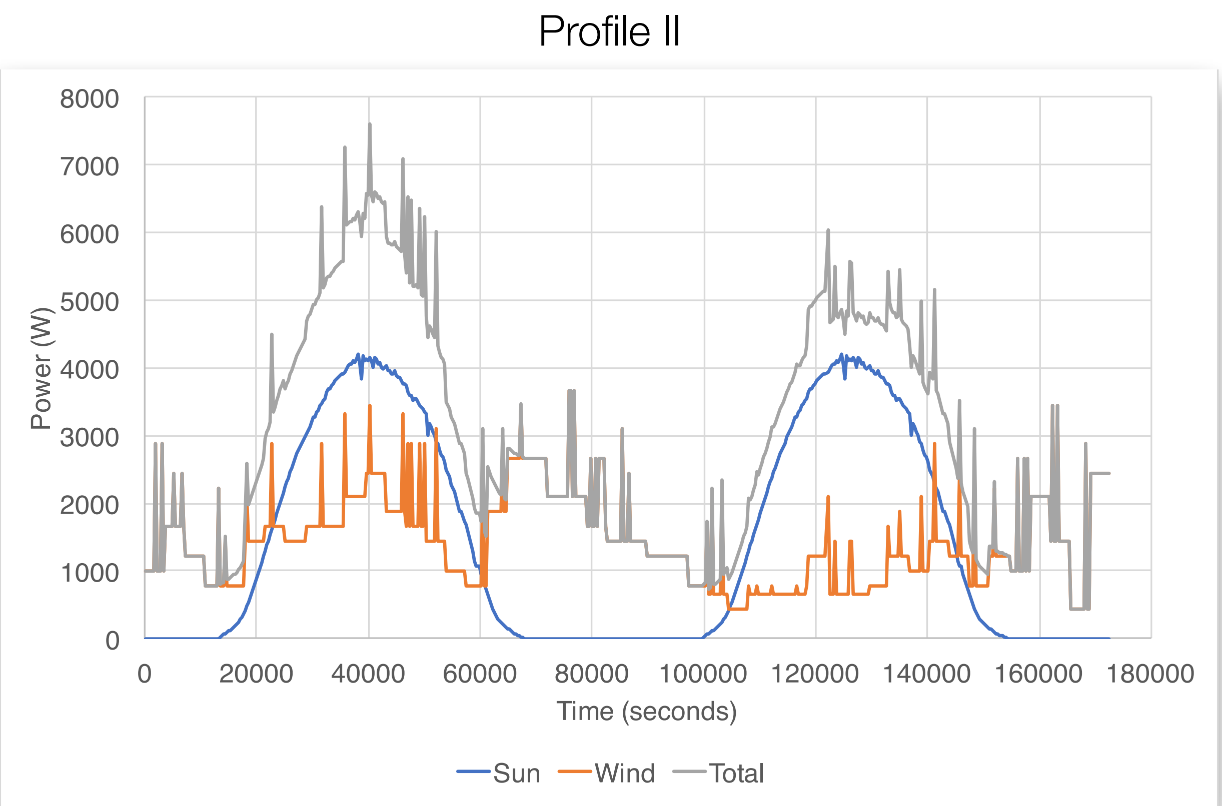
Paravance	**Freq. (GHz)**	2.4	2.16	1.92	1.68	1.44	1.2
	Power (W)	200.5	165.10	136.76	114.69	98.10	86.22
Taurus	**Freq. (GHz)**	2.3	2.07	1.84	1.61	1.38	1.15
	Power (W)	223.7	189.03	161.28	139.67	123.43	111.79

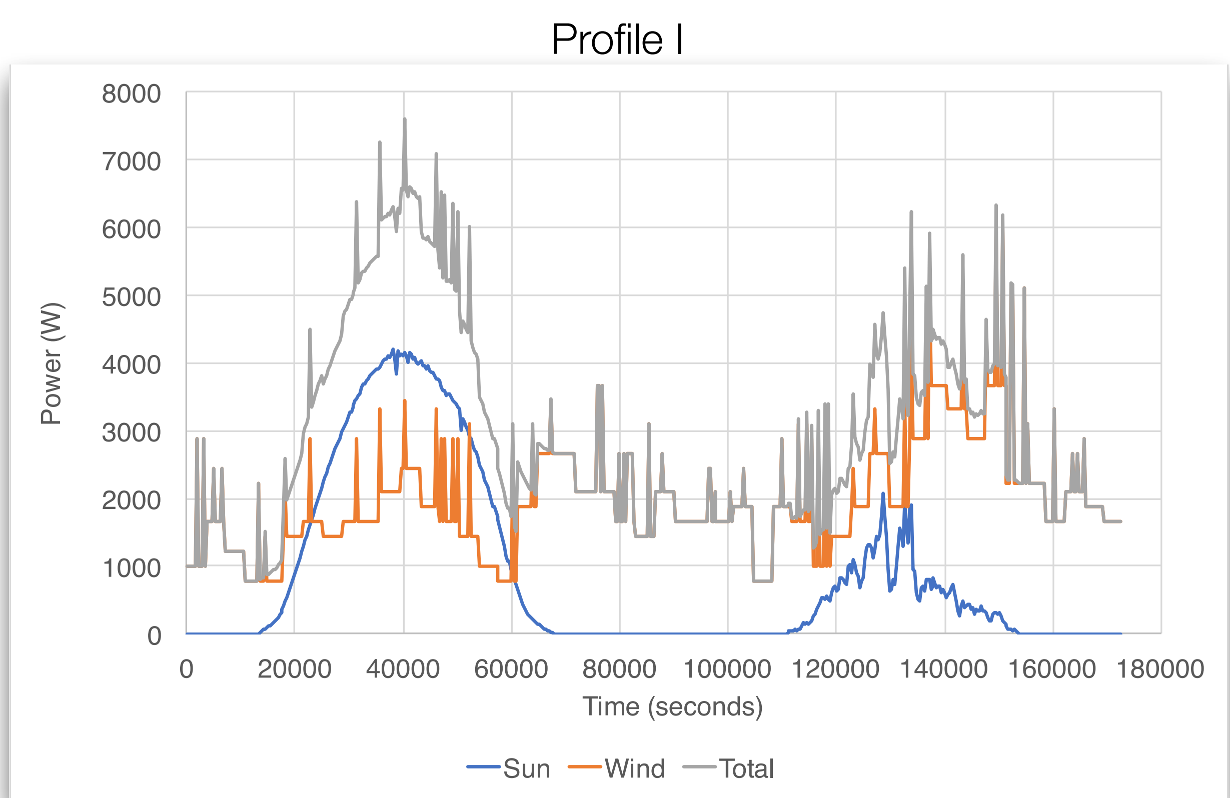
$$P_h = \begin{cases} P_h^{(\text{idle})} + \sum_{h=1}^{C_h} run_{h,p} \cdot P_h^{(\text{dyn})} \cdot (f_{h,p})^3 & \text{if } s_h = on \\ 0 & \text{otherwise} \end{cases}$$

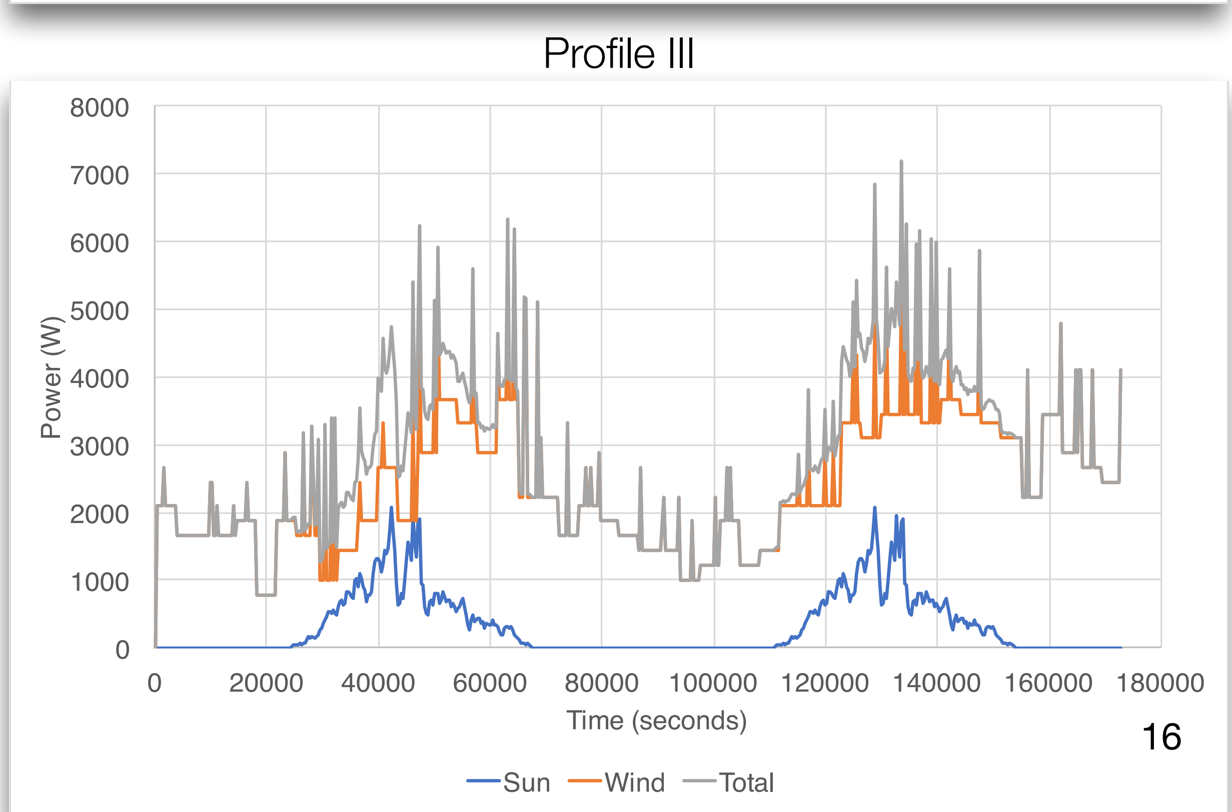
T. Mudge, "Power: A first-class architectural design constraint," Computer, vol. 34, pp. 52– 58, 2001.

Evaluation

- 2 days simulation with DCWorms:
  - 3 power profiles
  - 3 workloads (Google Based)
    - 234, 569 and 1029 tasks (known at beginning of execution)

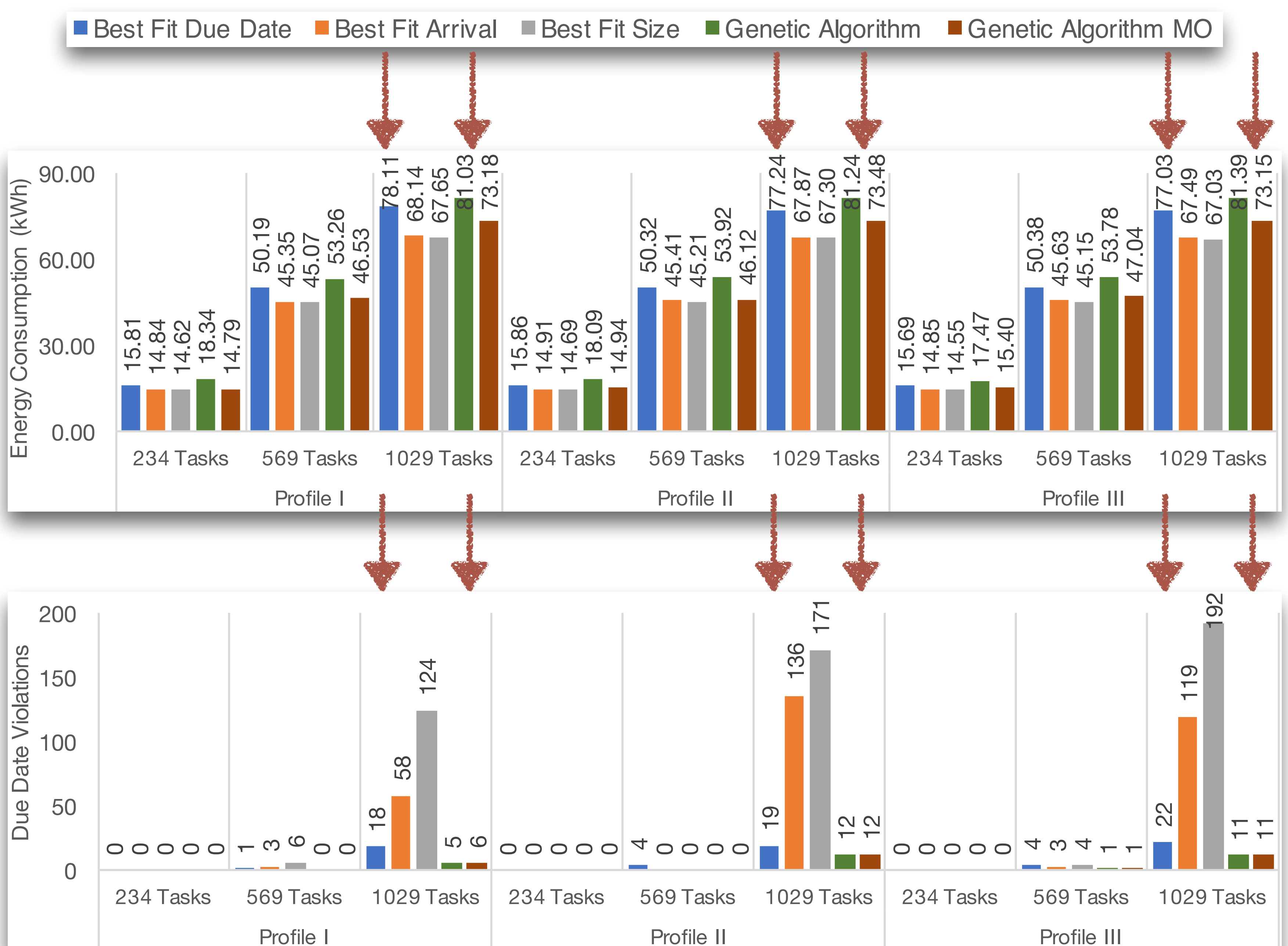






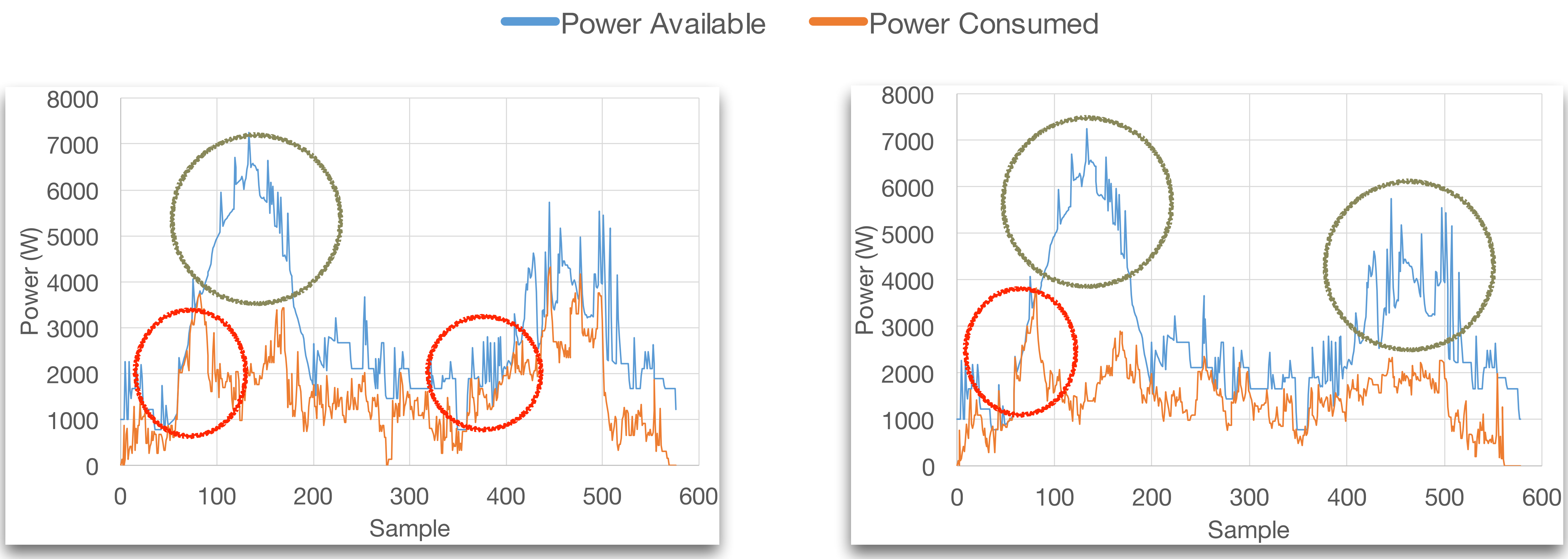
Results

Results



Best Fit Due Date
Best Fit Arrival
Best Fit Size
Genetic Algorithm
Genetic Algorithm MO
Energy Consumption (kWh)
90.00
60.00
30.00
0.00
15.81 14.84 14.62 18.34 14.79
50.19 45.35 45.07 53.26 46.53
78.11 68.14 67.65 81.03 73.18
15.86 14.91 14.69 18.09 14.94
50.32 45.41 45.21 53.92 46.12
77.24 67.87 67.30 81.24 73.48
15.69 14.85 14.55 17.47 15.40
50.38 45.63 45.15 53.78 47.04
77.03 67.49 67.03 81.39 73.15
234 Tasks
569 Tasks
1029 Tasks
Profile I
Profile II
Profile III
Due Date Violations
200
150
100
50
0
0 0 0 0 0
1 3 6 0 0
18 58 124 5 6
0 0 0 0 0
4 0 0 0 0
19 136 171 12 12
0 0 0 0 0
4 3 4 1 1
22 119 192 11 11

Results



Best Fit Due Date

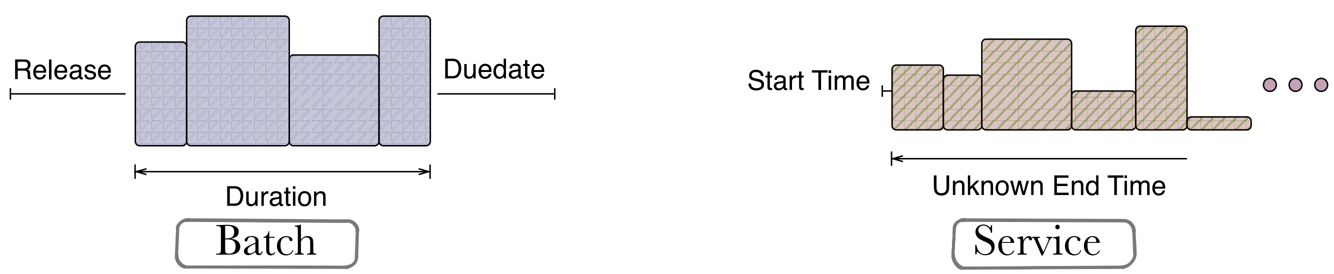
Genetic Algorithm MO

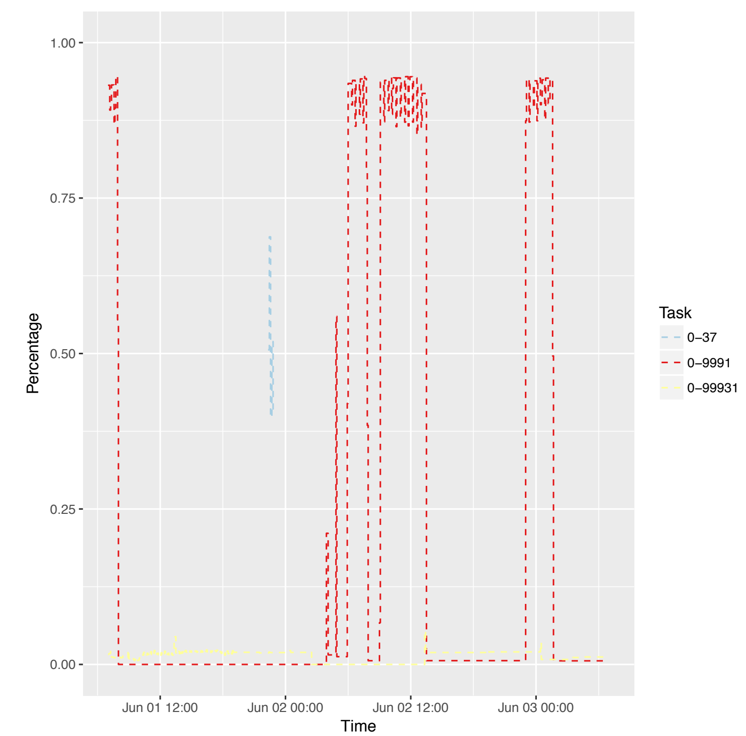
Conclusion

- Different algorithms that aims to minimize due date violations while respecting power and resource constraints.
- Provide scheduling possibilities translated into power profiles with associated metrics for decision modules.
- Integration between power production and IT load;
- Just one segment of DataZero, which contains more modules that interacts with IT.

Future Works

- Scheduling of mixed workload batch and services;
- Phases based tasks:





- Pre-evaluate workload/time available to choose algorithm;
- Online scheduling + weather events (quick reaction).

Green Days@Toulouse v2.0 (Edition 2018)

IT Optimization Under Renewable Energy Constraint

Gustavo Rostirolla

gustavo.rostirolla@irit.fr

Stephane Caux, Paul Renaud-Goud, Gustavo Rostirolla, Patricia Stolf. *IT Optimization for Datacenters Under Renewable Power Constraint (regular paper). Euro-Par*, *Turin*, *27/08/2018 - 31/08/2018*.