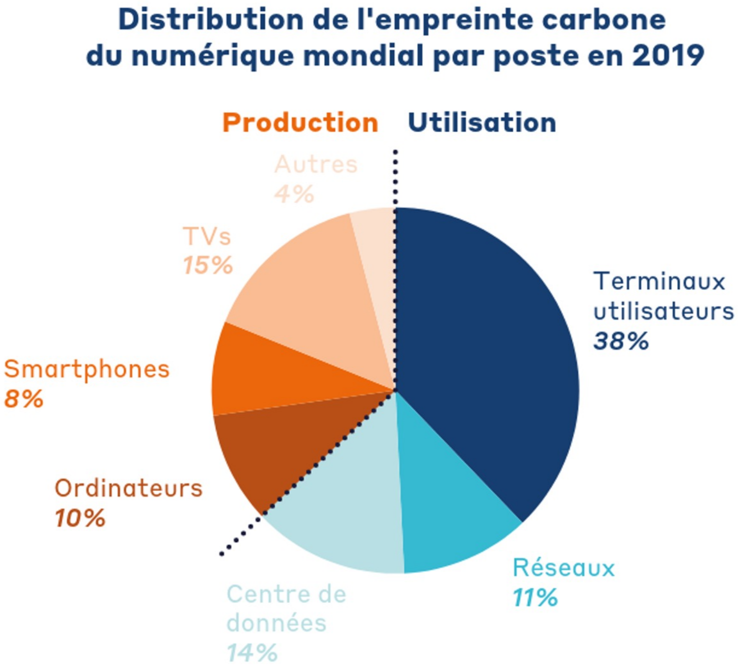
IT and Power Datacenter sizing powered by renewable energy

Manal BENAISSA - PhD Student
Supervisors : Jean marc NICOD - George DA COSTA
FEMTO-ST - UPS-IRIT

Context



78.7 millions tons of CO_2 are produced by datacenters = **2%** of global emissions

Datacenter = **14%** of the total amount of carbon production by ICT

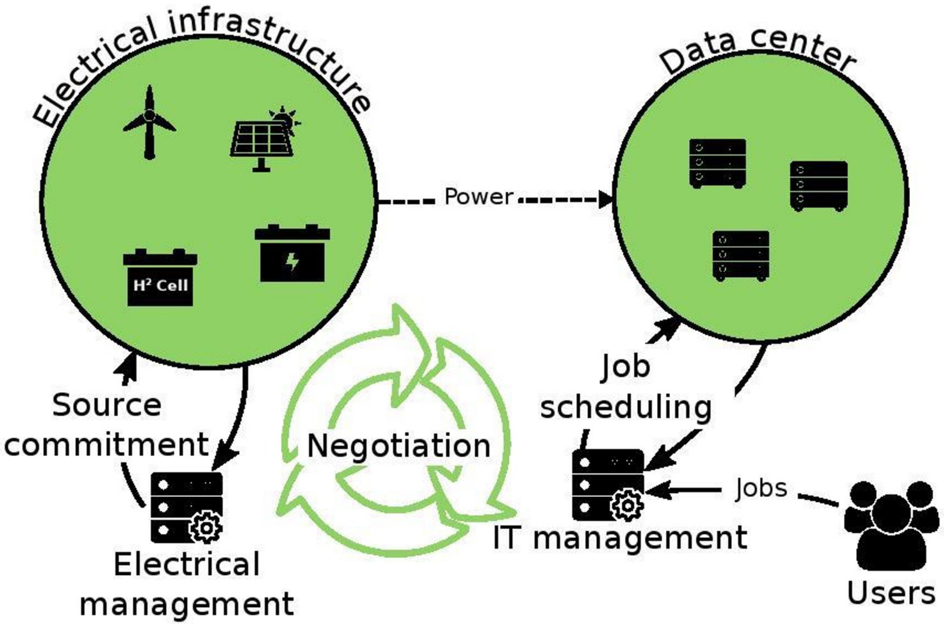
Global datacenter electricity demand in 2019 was around **200 TWh** (1 nuclear reactor = 7 TWh)

Global view of the system

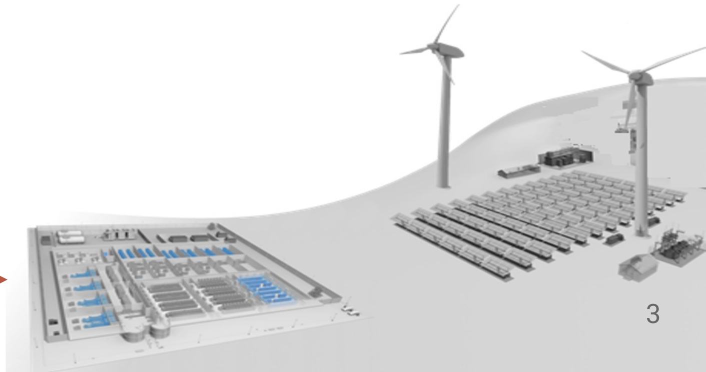
System divided into two parts : IT and Elec.

IT part : Offer an IT infrastructure to satisfy users demands

Elec part : Considering a power demand D and the weather conditions, define and manage the Electrical infrastructure.



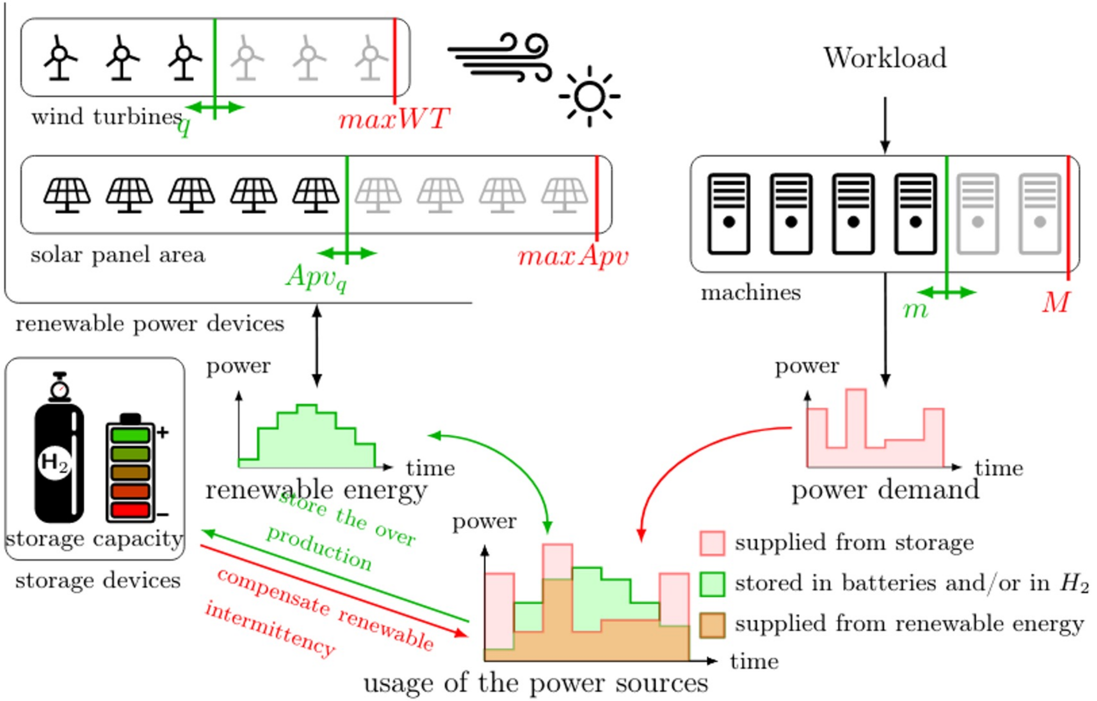
Datacenter with local renewable power



Sizing - Summary

Q : All possible configurations that satisfy the asked yearly requested energy demand of the IT infrastructure.

(0, Apv0) ∈ Q
(maxWT, 0) ∈ Q



outputs: configurations (m, q, Apv_q, storage capacity) + associated metrics, QoS, etc.

Step 1- IT Sizing : For a given workload, computes the requested number of machines *m*, and gives a *power profile* (power demand for each timeslot for a year)

Step 2 - Electrical sizing : For a given power demand (provided by the IT sizing) and a given wind speed and irradiation, gives a set Q of several configurations of wind turbines (q) & Solar Panel (Apvq), battery size (maxBCq) and H² tank size (maxLOHq)

IT Sizing

Concept

IT Sizing

Workload : Offline. No precedence constraints, preemption allowed.

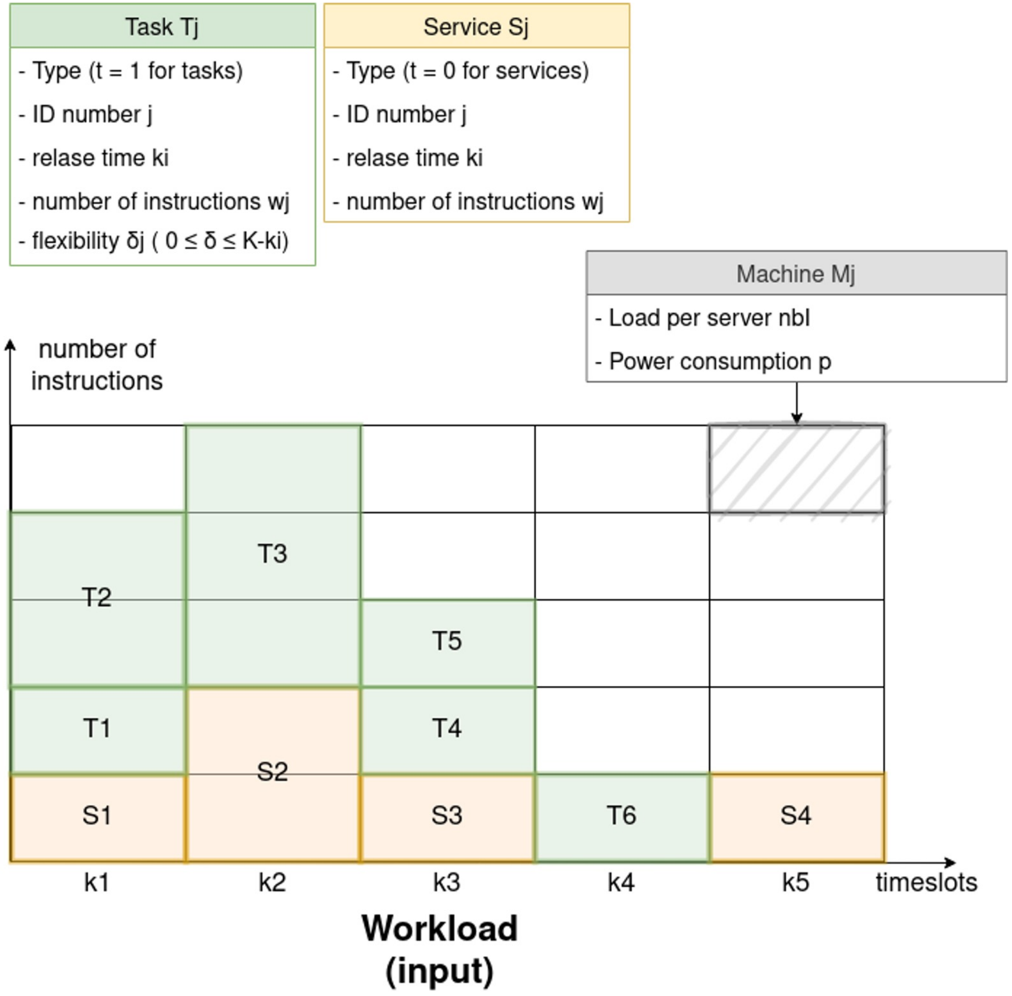
Services : Not flexible.

Tasks : Flexibles. All tasks with the same flexibility.

System : m identical machines.

Objective : Schedule all tasks by respecting flexibility constraints, before horizon K + flex.

1 timeslot = 1 hour
K timeslot = 1 year

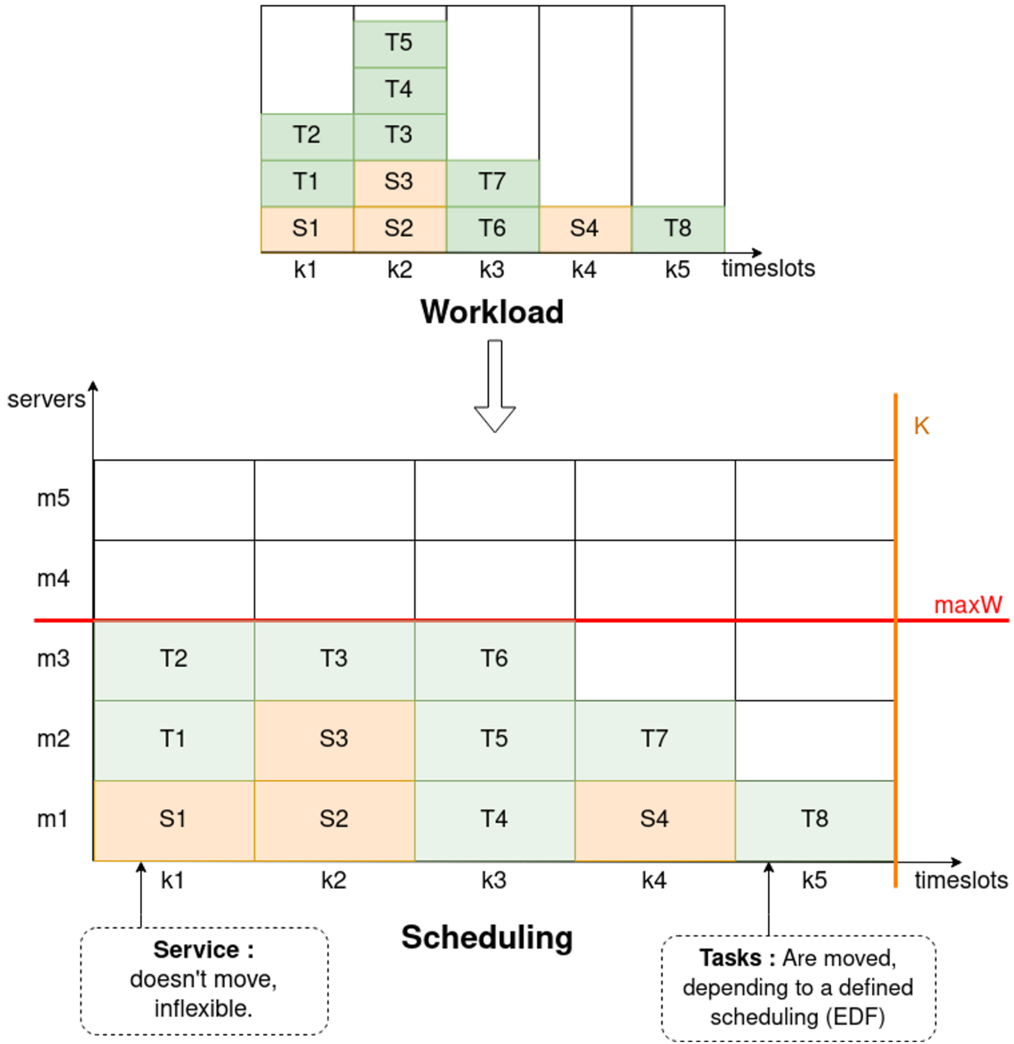


IT Sizing

Sizing of the IT part : Computation of the minimum number of servers m, given a workload divided into K timeslots and scheduling algorithm.

The value m is found by Binary search : For each m between *minM* and *maxM*, a scheduling (Earliest Deadline First) is tested.

- If all work are scheduled, decrease m.
- If not, increase m.



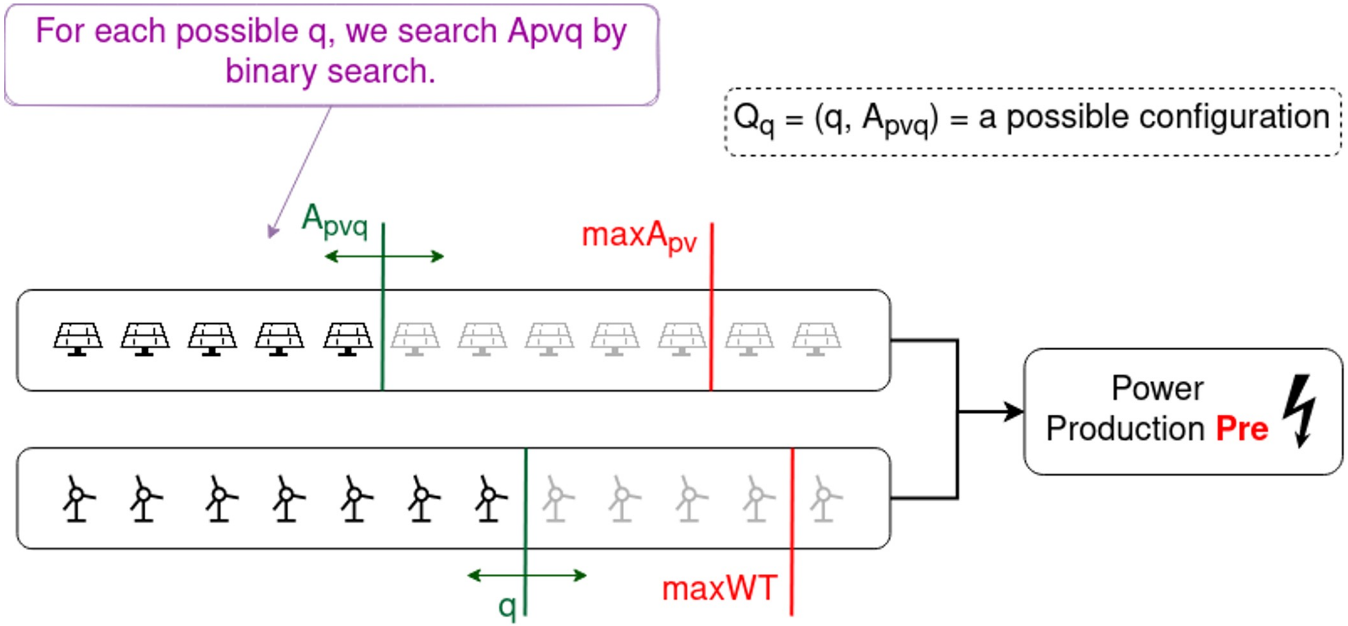
Electrical Sizing

Concept

Electrical Sizing (primary sources)

Q : All possible configurations that satisfy the asked yearly requested energy demand of the IT infrastructure.

(0, Apv0) ∈ Q
(maxWT, 0) ∈ Q



Sizing of the Electrical part : computation of the needed primary and secondary sources, given a power demand D and the weather condition.

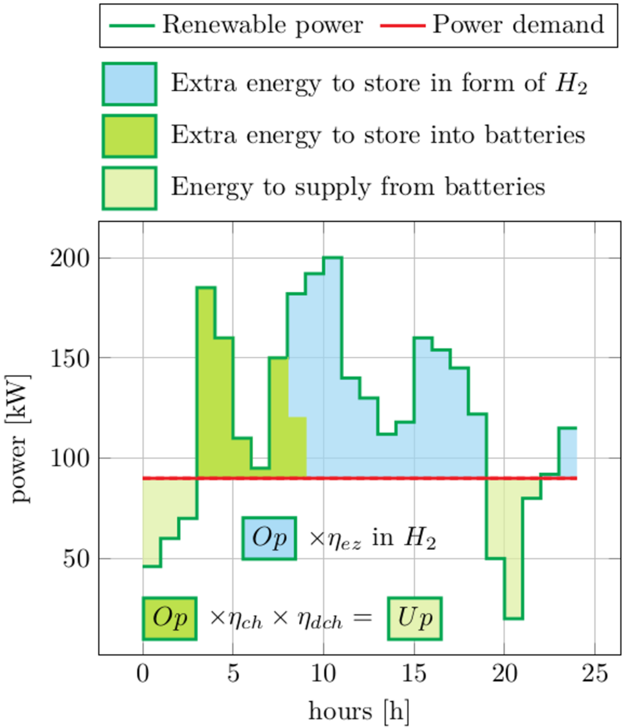
Primary sources : Wind turbines (WT) & Solar Panel (PV)

Secondary sources: Batteries (BC) & Hydrogen system (LOH).

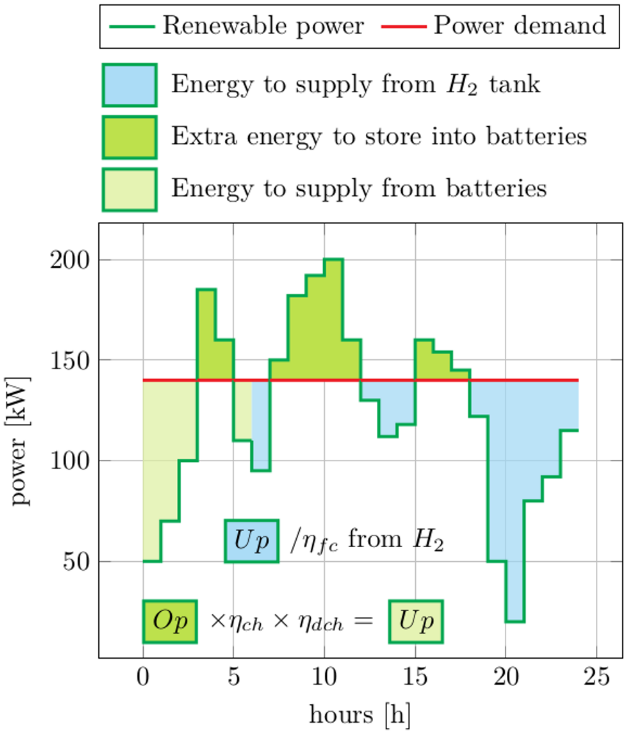
We have : $0 \leq q \leq maxWT$, where q is the number of WT, maxWT the minimal number of WT required to satisfy demand without any PV, and maxApv, the maximum area of PV needed, without any WT

Electrical Sizing (secondary sources)

Overproductive day



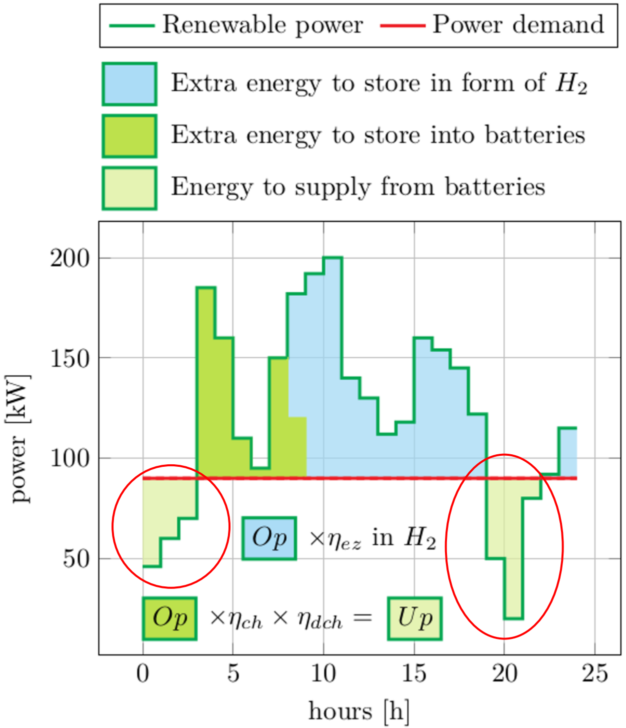
Underproductive day



Electrical Sizing (secondary sources)

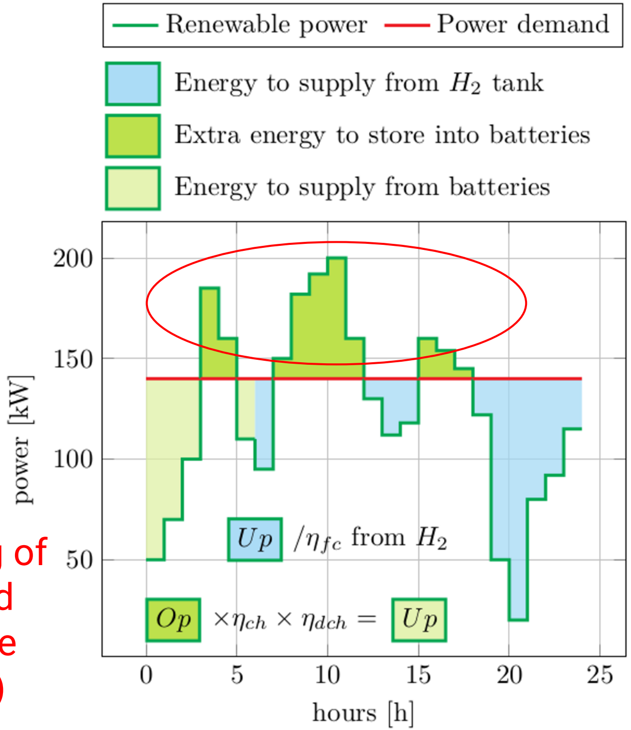
Constraint : The amount of battery at the end of the day [BC(24)] should be the same than the beginning (BC[0]) !

Overproductive day



give the sizing of battery needed to compensate this day (debt)

Underproductive day



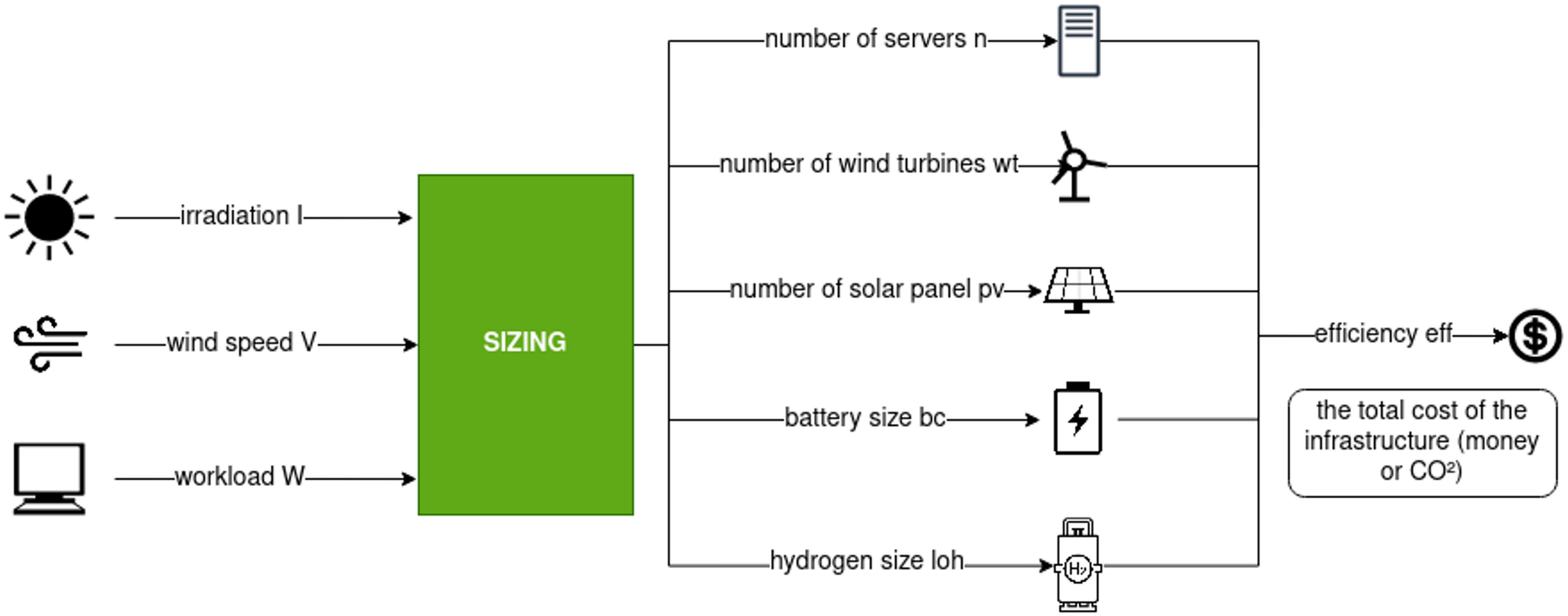
give the sizing of battery allowed to use for this day (credit)

Uncertainties

First approach of robustness study

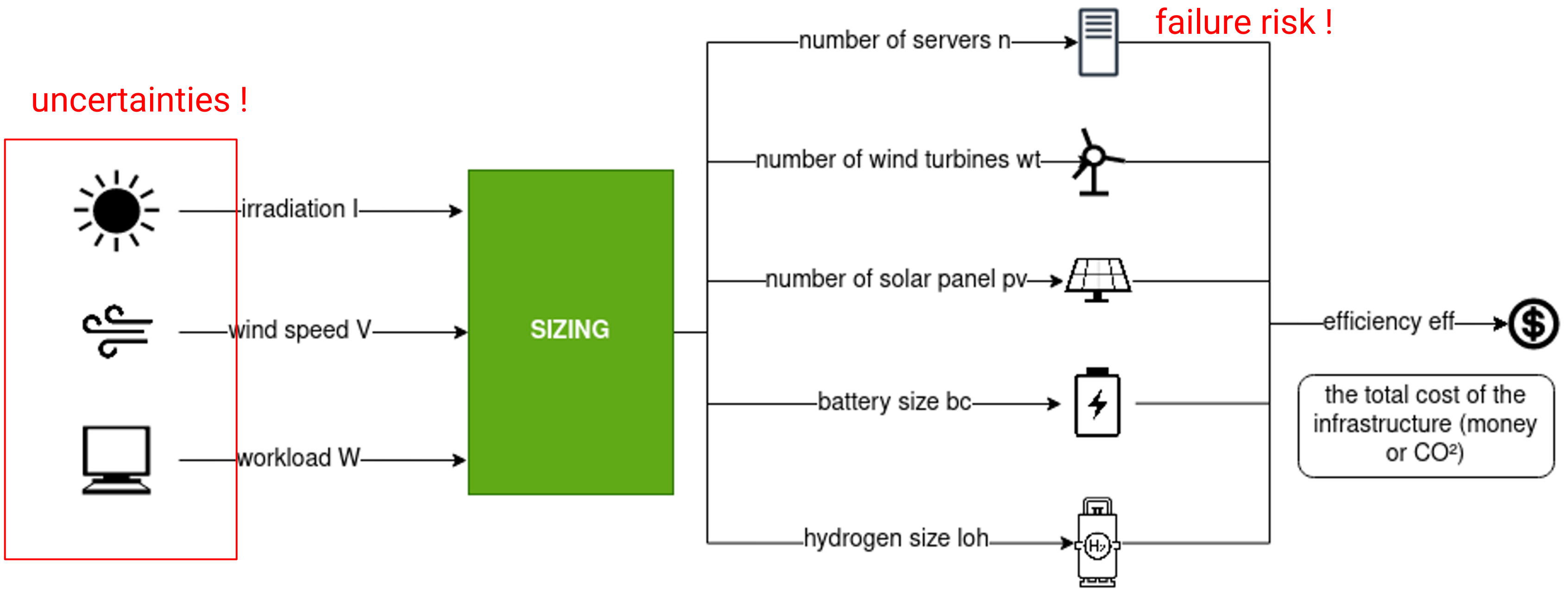
Function definition

(w, i, v) → (n, wt, pv, bc, loh, eff)



Function definition

(w, i, v) → (n, wt, pv, bc, loh, eff)



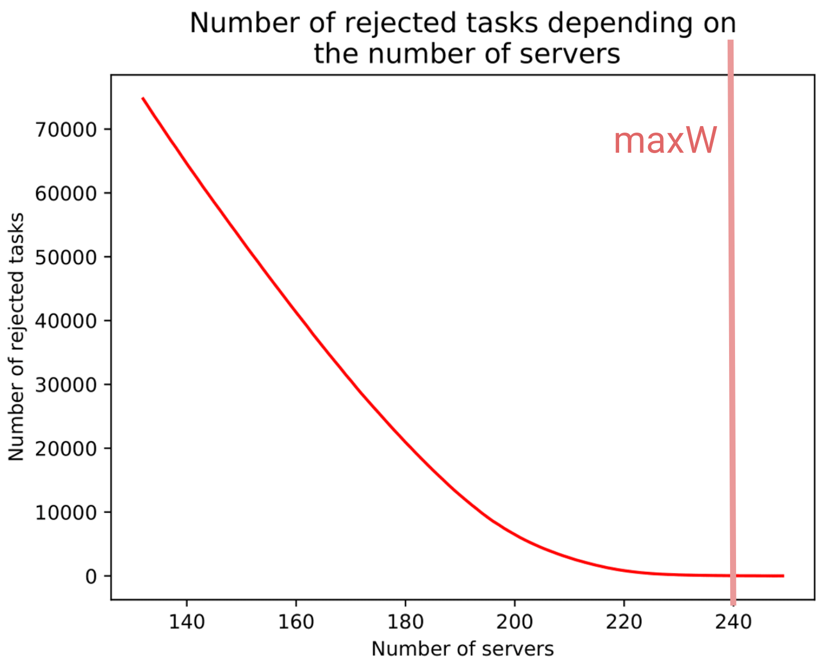
Sensitivity Analysis

Idea : play with the outputs to see how the sizing responds.

Method :

- **Select one configuration**
- **Modify** one parameter (number of servers for example)

Aim : A linear search of the sensitivity. By decreasing **n,** the number of servers, how the others parameters (wt, pv…) react ? See if any law is visible (decreasing by one machine decrease drastically wt, pv...)



(w, i, v) ⟶ (n, wt, pv, bc, loh, eff)

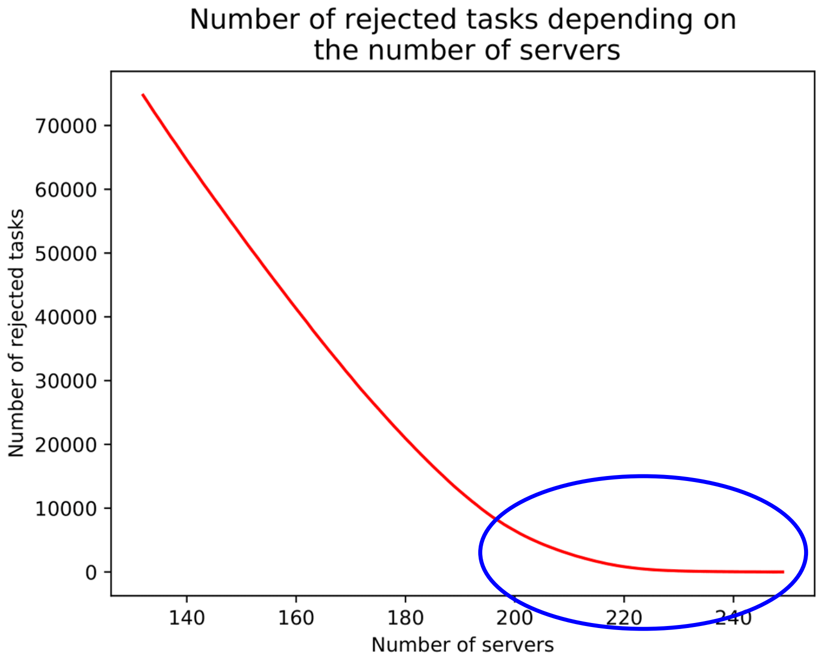
Sensitivity Analysis

Idea : play with the outputs to see how the sizing responds.

Method :

- **Select one configuration**
- **Modify** one parameter (number of servers for example)

Aim : A linear search of the sensitivity. By decreasing **n,** the number of servers, how the others parameters (wt, pv…) react ? See if any law is visible (decreasing by one machine decrease drastically wt, pv...)



Decreasing by 20% the number of machine doesn't change the number of rejected tasks but it can change drastically the number of required WT, and in the same way, the total efficiency (by counting n)

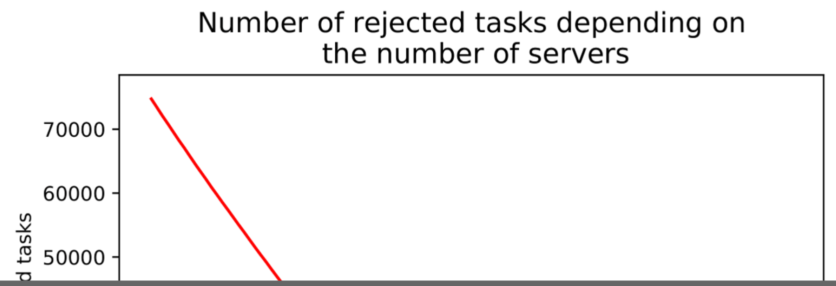
Sensitivity Analysis

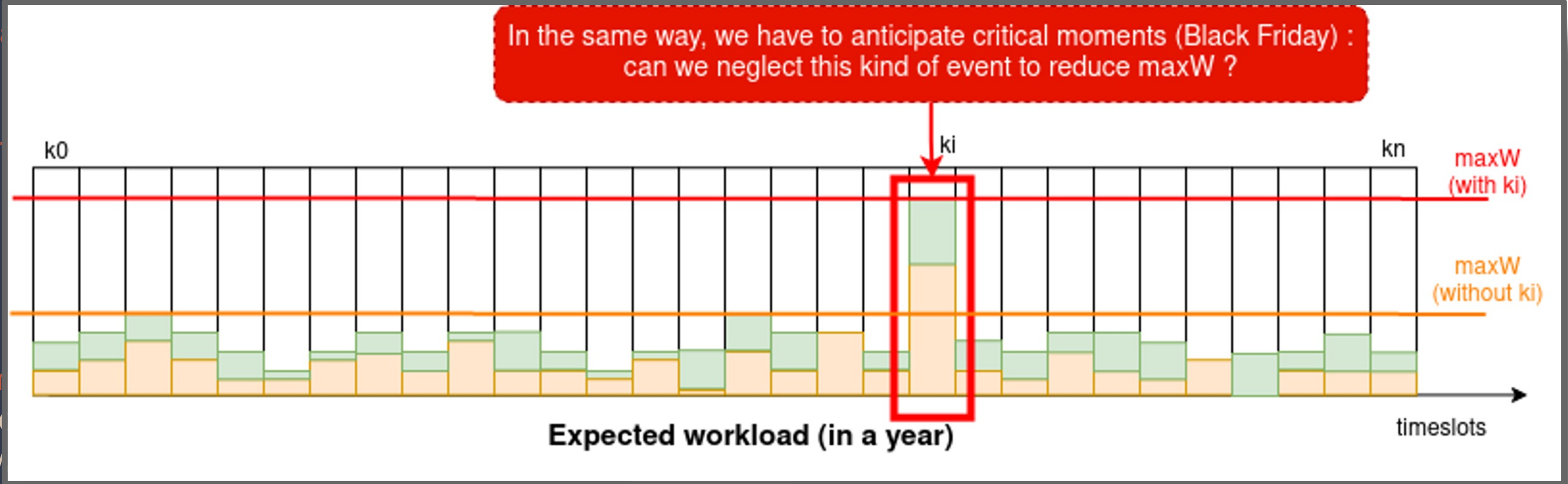
Ide
the

Me

-
-

Ain
By
hov
react ? See if the pareto law (80/20 law) is visible (decreasing by one machine decrease drastically wt, pv...)





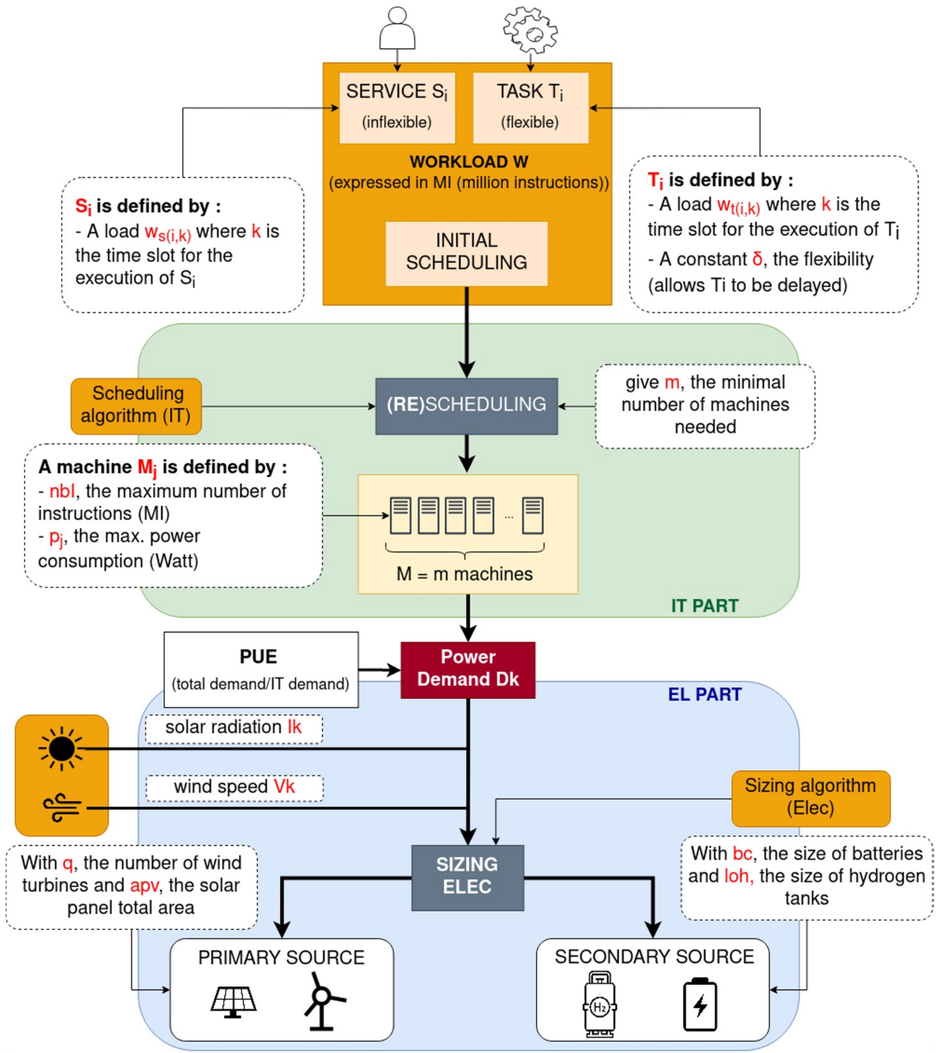
doesn't change the number of rejected tasks but it can change drastically the number of required WT, and in the same way, the total efficiency (by counting n)

Conclusion

Sizing divided into two parts : IT and Elec.

IT part : Considering a workload W, finding the minimum number of servers maxW needed

Elec part : Considering a power demand D and the weather conditions, define Q : All possible configurations that satisfy the asked yearly requested energy demand of the IT infrastructure.



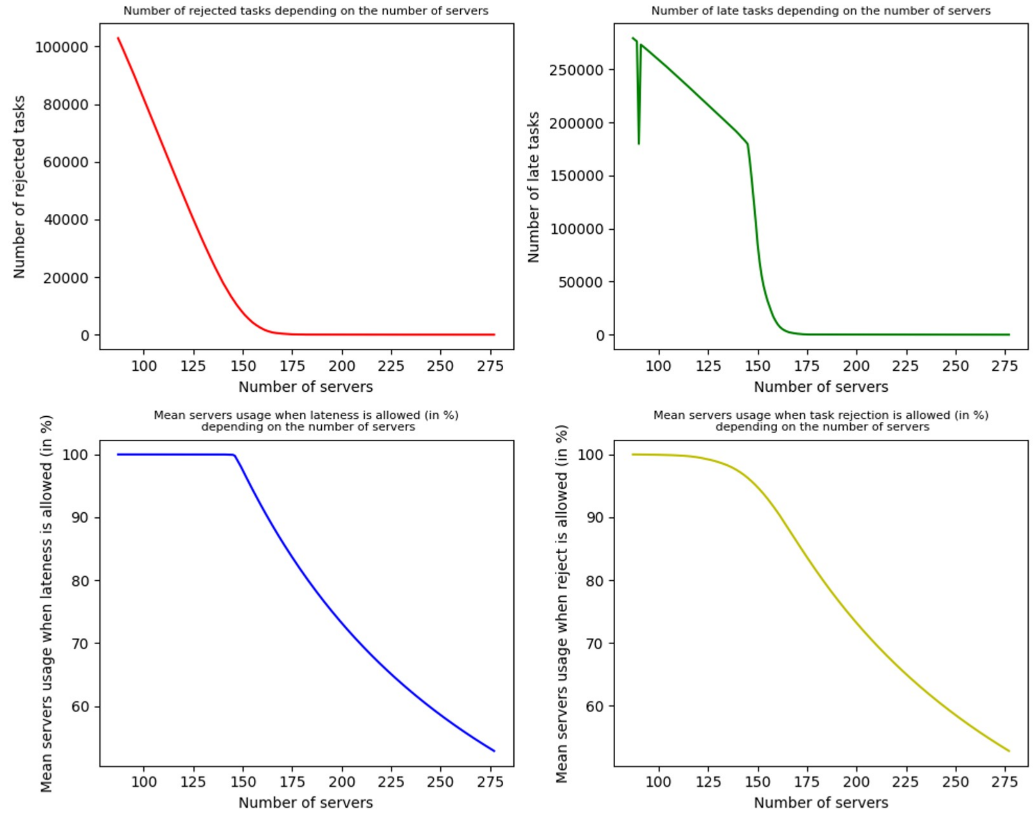
Thanks for your attention

Questions ?

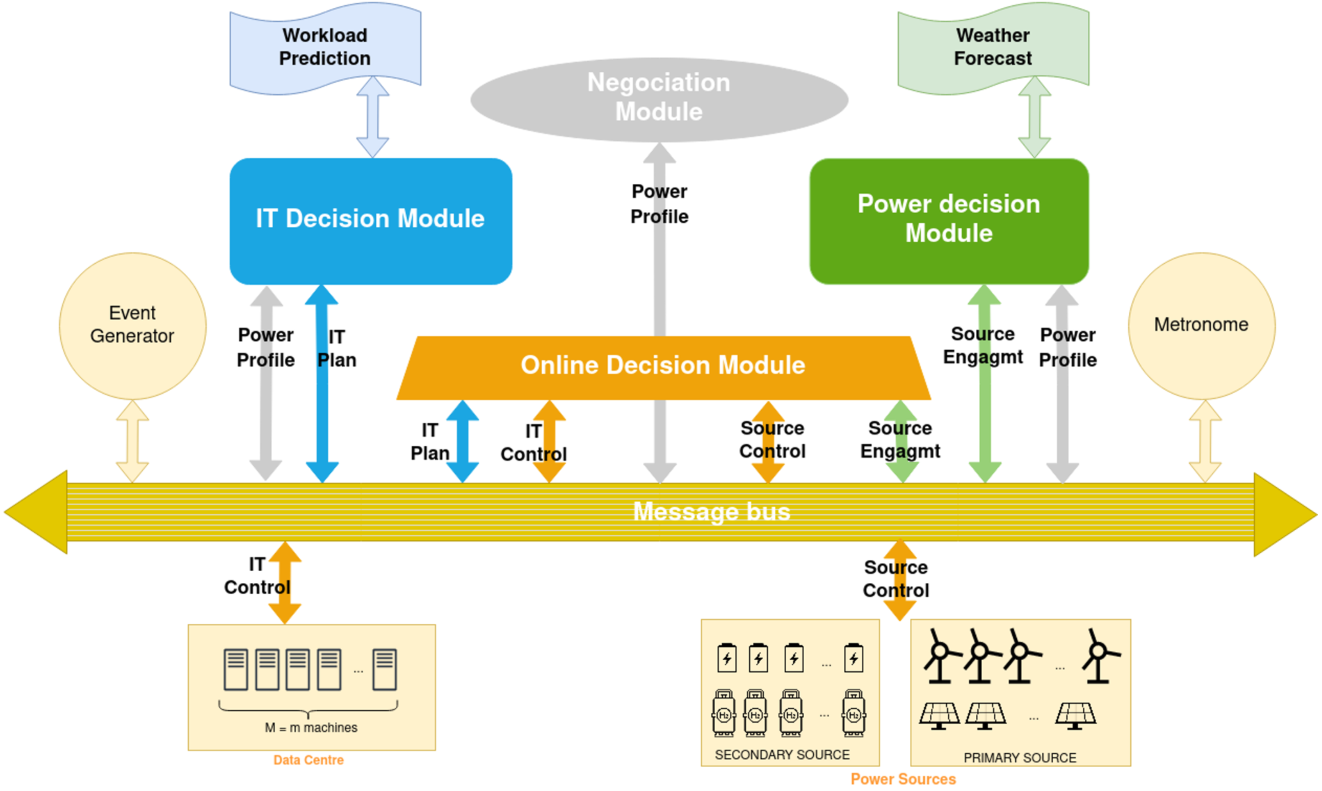
Rejection vs Lateness

Two strategies used :

- **Rejection (left) :** If deadline reached, the task is completely rejected. This strategy tends to reject huge tasks and schedule the smallest ones.
- **Lateness (right):** Task scheduled in any case, even if deadline reached. The delay is propagated to next tasks, until a empty timeslot.



Middleware


Workload
Prediction
Negociation
Module
Weather
Forecast
IT Decision Module
Power
Profile
Power decision
Module
Event
Generator
Power
Profile
IT
Plan
Online Decision Module
Source
Engagmt
Power
Profile
Metronome
IT
Plan
IT
Control
Source
Control
Source
Engagmt
Message bus
IT
Control
Source
Control
M = m machines
Data Centre
SECONDARY SOURCE
PRIMARY SOURCE
Power Sources