18th Scheduling for large-scale systems workshop

Exploring scheduling solutions for Federated Learning training

Laércio Lima Pilla (he/him)
laercio.lima-pilla@labri.fr

… but before we start discussing the topic

Who am I? 🇧🇷 🇫🇷

- CNRS research scientist (2018-)
- Aim to make **parallel applications more efficient*** through **better scheduling**
  - * faster, more energy efficient, using less or older resources, …

website

… but before we start discussing the topic

Who am I? 🇧🇷 🇫🇷

- CNRS research scientist (2018-)
- Aim to make **parallel applications more efficient*** through **better scheduling**
  - * faster, more energy efficient, using less or older resources, …

[website]

Acknowledgements

- PhD student
  - Alan Lira Nunes
- Co-advisors
  - Cristina Boeres
  - Lúcia M. A. Drummond

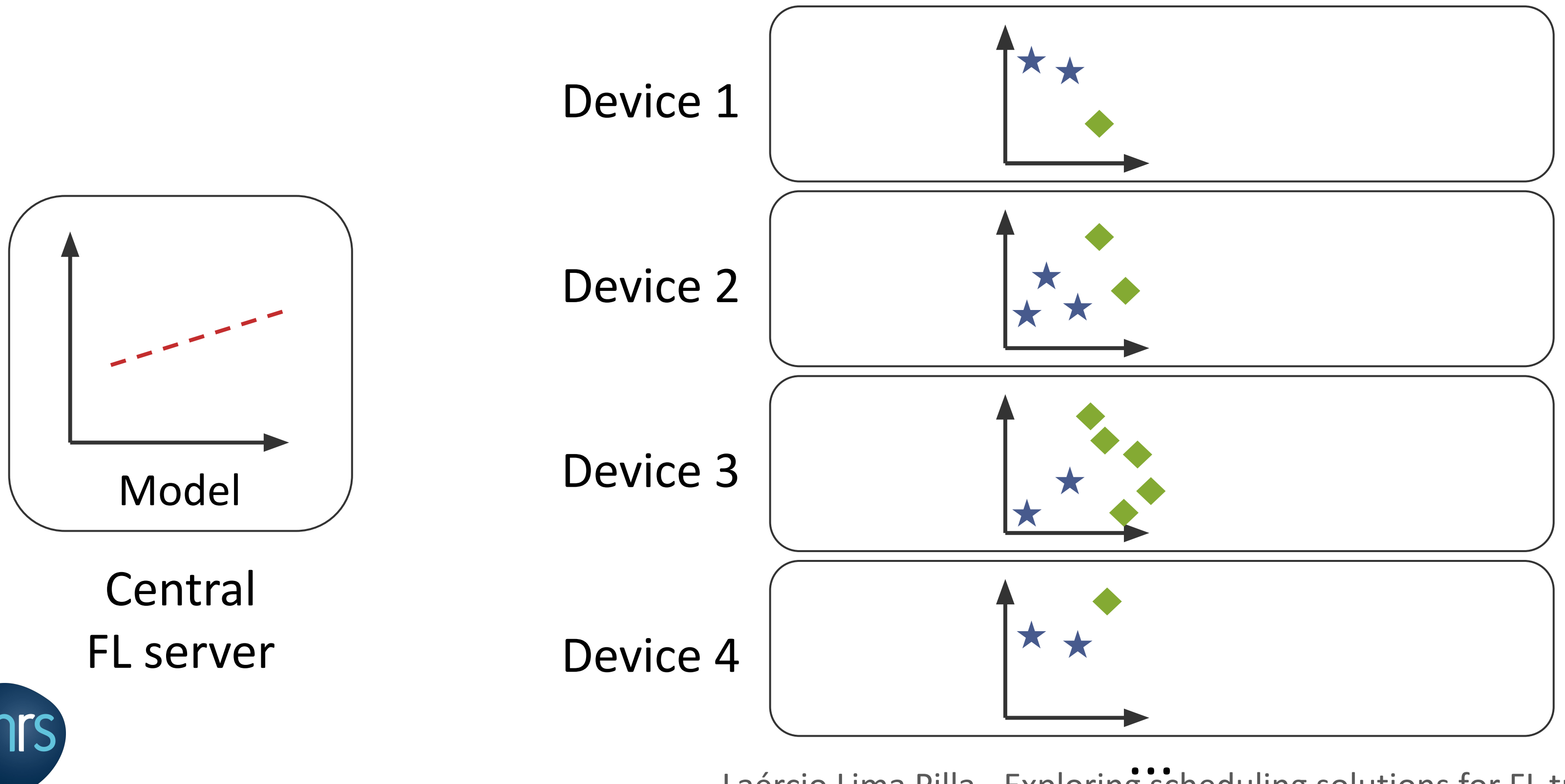
Instituto de Computação

- DecoHPC joint team

What is Federated Learning?

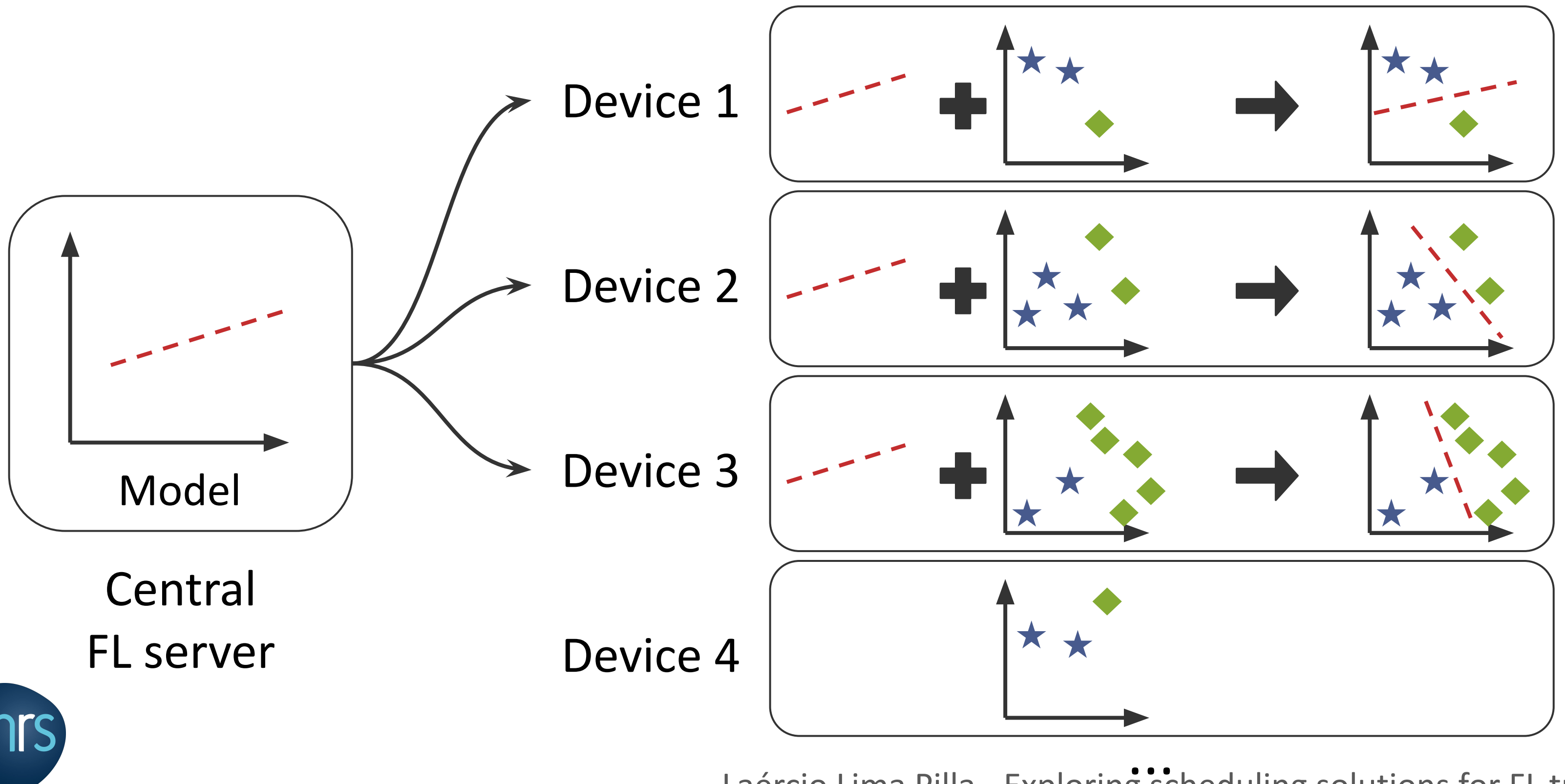
distributed learning + local data is never shared

Communication-Efficient Learning of Deep Networks from Decentralized Data



What is Federated Learning?

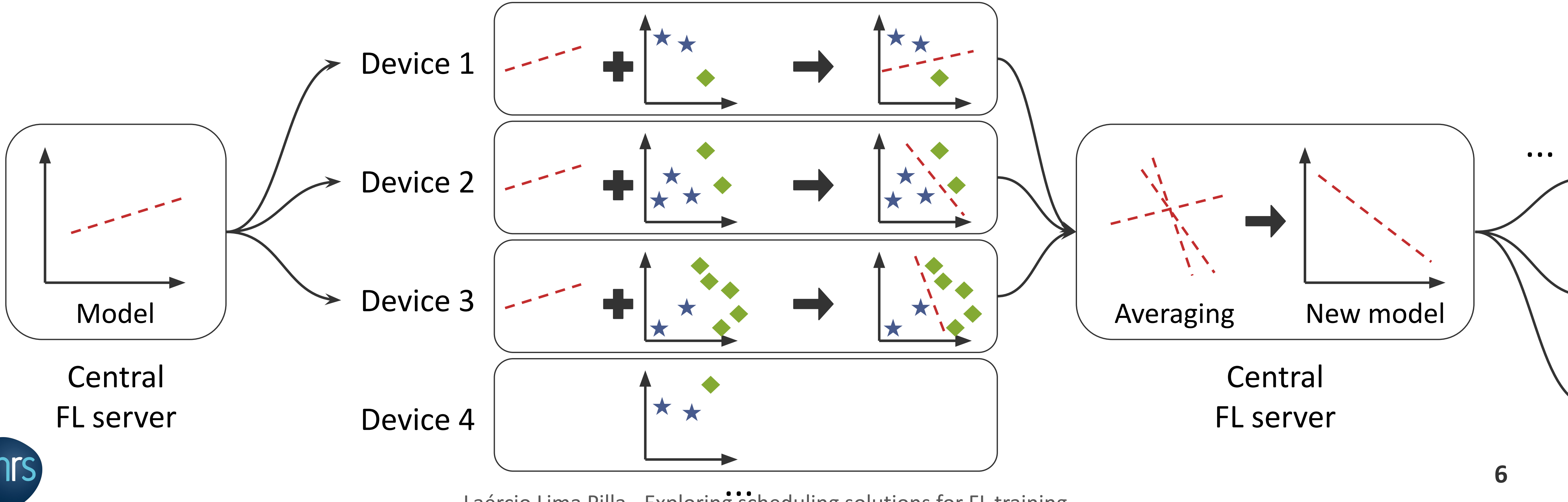
distributed learning + local data is never shared



What is Federated Learning?

distributed learning + local data is never shared

Applications: next-word prediction, on-device item ranking, cyberattack detection, medical applications



Why do we care (about optimizing it)?

Why do we care (about optimizing it)?

[...] this year edition will be focused around "Scheduling and AI" [...]

Why do we care (about optimizing it)?

Characteristics of cross-device FL

device and data heterogeneity

Issues

duration of training

energy costs and emissions of ML*

battery or energy available

convincing people to participate in training

*Qiu, Xinchi, et al. "A first look into the carbon footprint of federated learning." Journal of Machine Learning Research 24.129 (2023): 1-23.

What the problems look like?

For a training round, given

- some heterogeneous clients
- amount of data to be used for training

try to optimize

- the time
- the energy (or the emissions)
- both

it takes by

- defining how much data each client should use locally

What the problems look like?

For a training round, given

- some heterogeneous clients
- amount of data to be used for training

try to optimize

- the time
- the energy (or the emissions)
- both

it takes by

- defining how much data each client should use locally

Requires estimations of

- the amount of local data*
- functions of
  - time per data unit
  - energy per data unit

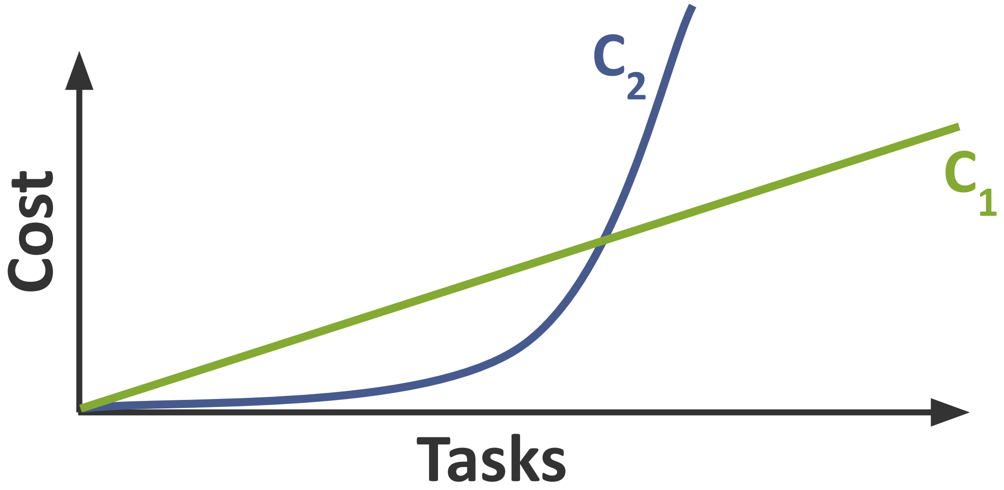
per client

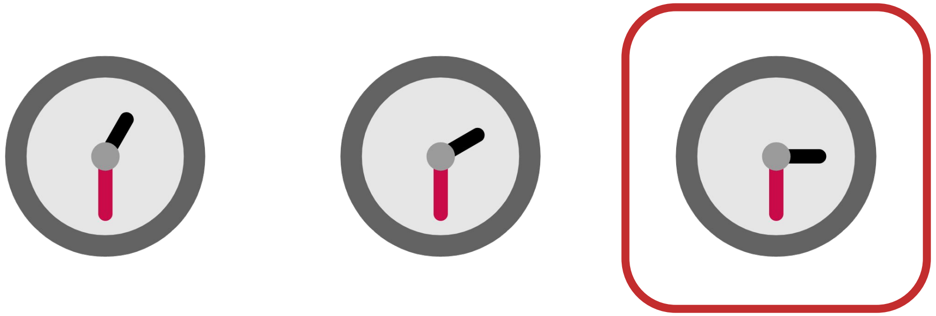
*or upper and lower limits

Optimizing time (2021)

Problem: given **T tasks** and **n resources** with different cost functions C, find an assignment that minimizes the **maximal cost C_{max}**.

Assumption: all cost functions are monotonically increasing and known

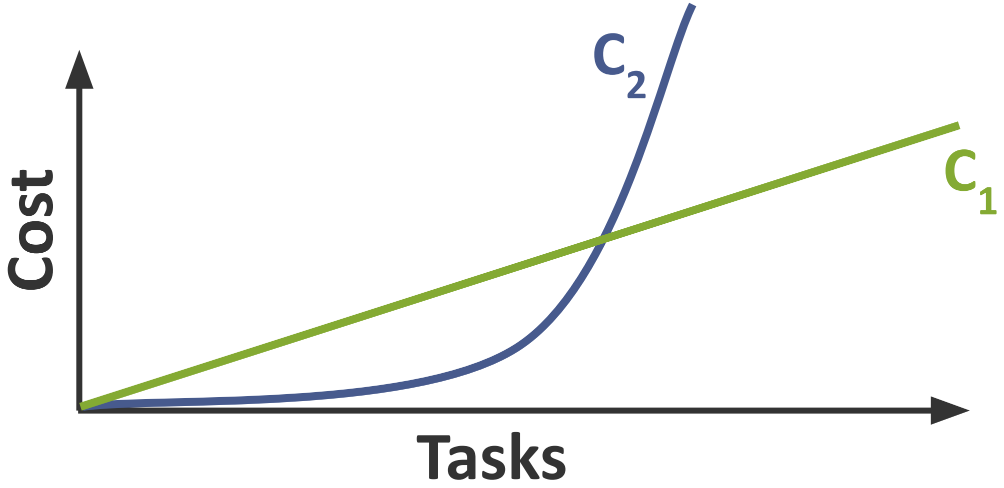




Optimizing time (2021)

Problem: given **T tasks** and **n resources** with different cost functions C, find an assignment that minimizes the **maximal cost C_{max}**.

Assumption: all cost functions are monotonically increasing and known



Find a task assignment $A_i \in \mathbb{N}$ to each resource $i \in R$ that minimizes the makespan C_{max} while assigning all tasks among the resources and respecting their lower and upper limits.

$$C_{max} := max_{i \in R}\, C_i(A_i)$$

$$\sum_{i \in R} A_i = T$$

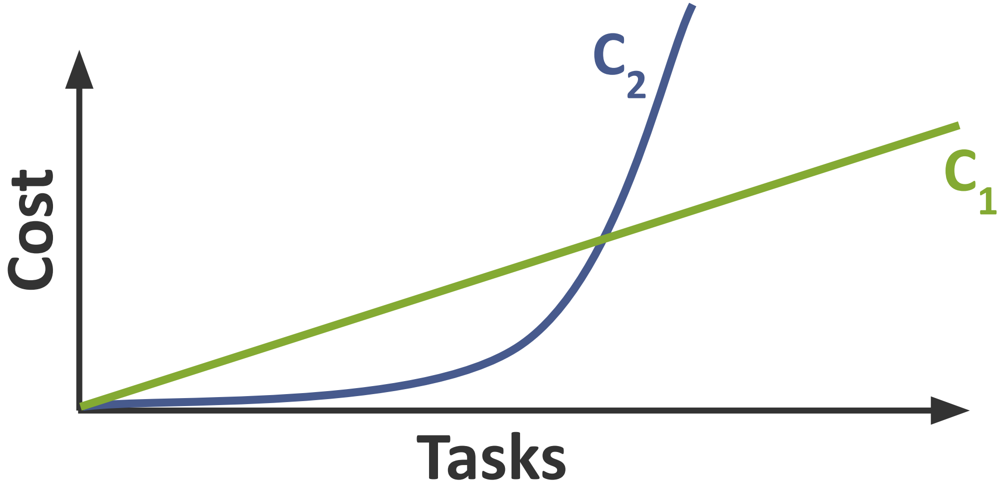
$$L_i \leq A_i \leq U_i,\ \forall i \in R$$

Pilla, Laércio Lima. "Optimal task assignment for heterogeneous federated learning devices." 2021 IEEE International Parallel and Distributed Processing Symposium (IPDPS). IEEE, 2021.

Optimizing time (2021)

Problem: given **T tasks** and **n resources** with different cost functions C, find an assignment that minimizes the **maximal cost C_{max}**.

Assumption: all cost functions are monotonically increasing and known



Solution in O(T log n): assign the **next task** to one of the resources that increases the execution time the least.

Algorithm 1: OLAR

Data: Tasks T, Resources $\mathcal{R}$, Cost functions $C_i(\cdot)$, Lower and Upper limits L_i and U_i $(i \in \mathcal{R})$

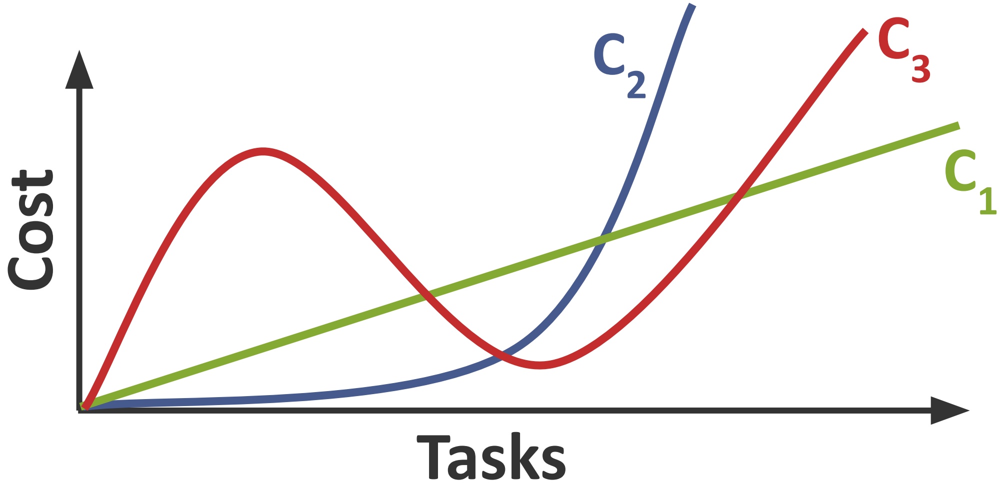
Result: Assignment of tasks to resources A_i $(i \in \mathcal{R})$

- $h \leftarrow min\text{-}heap()$ ▷ Heap sorted by cost
- **for** $i \in \mathcal{R}$ **do**
  - $A_i \leftarrow L_i$ ▷ Resources start at their lower limit
  - ▷ Checks if the resource can receive more tasks
  - **if** $A_i < U_i$ **then**
    - ▷ Inserts the cost of the next task on i
    - $h.push(C_i(A_i + 1), i)$
  - **end**
- **end**
- ▷ Main loop
- **for** t *from* $l + 1$ *to* T **do**
  - ▷ Extracts the next optimal assignment (Eq. (9))
  - $(c, j) \leftarrow h.pop()$
  - $A_j \leftarrow A_j + 1$ ▷ Assigns t to j
  - ▷ Checks if the resource can receive more tasks
  - **if** $A_j < U_j$ **then**
    - ▷ Inserts the cost of the next task on j
    - $h.push(C_j(A_j + 1), j)$
  - **end**
- **end**

Optimizing time (2022)

Problem: given **T tasks** and **n resources** with different cost functions C, find an assignment that minimizes the **maximal cost C_{max}**.

Assumption: all cost functions are ~~monotonically increasing and~~ known



Solution in $O(T^2n)$: a bit more complex

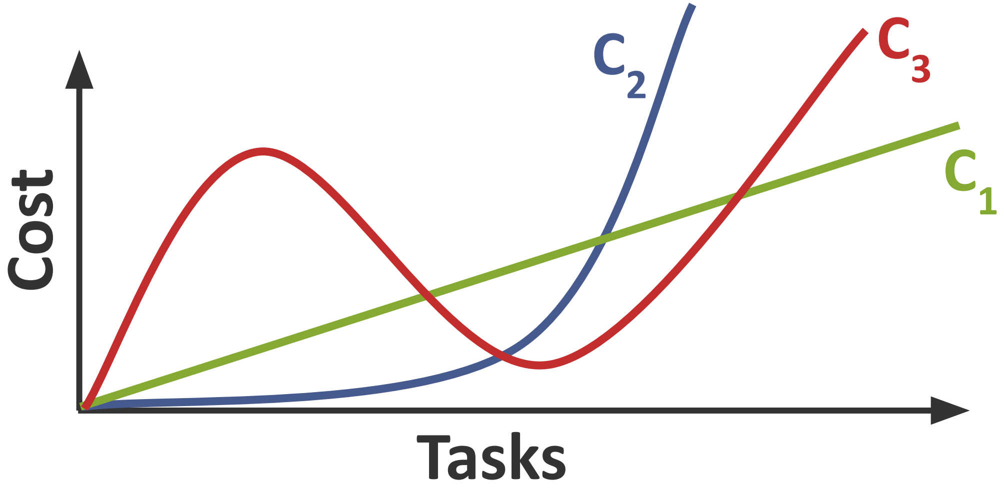
- Dynamic programming solution
- (to be shown soon) ;)
- SotA was $O(T^3n^3)$

Pilla, Laércio Lima. Optimal workload scheduling algorithm for data-parallel applications on heterogeneous platforms based on dynamic programming. Diss. CNRS; LaBRI; Inria; Université de Bordeaux; Bordeaux INP, 2022.

Optimizing **energy** (**2023**)

Problem: given **T tasks** and **n resources** with different cost functions C, find an assignment that minimizes the **total cost ΣE**.

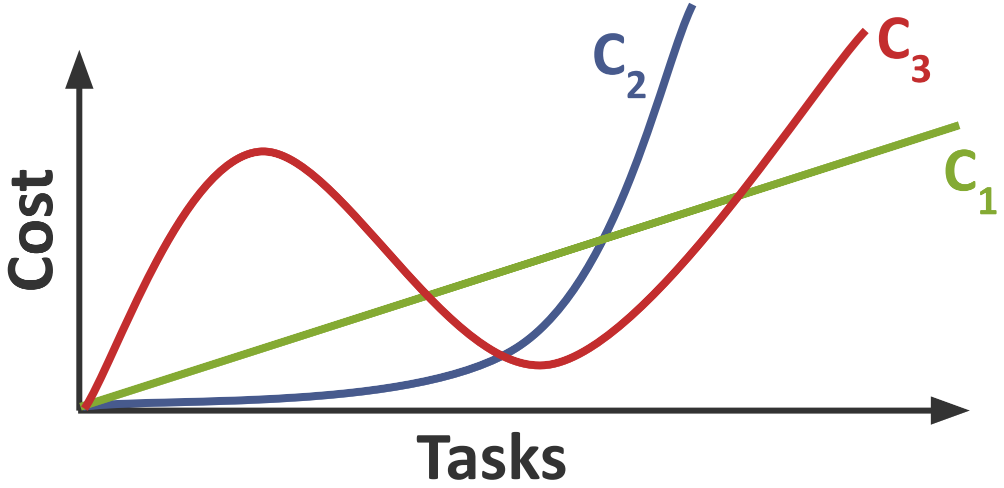
Assumption: all cost functions are known



Optimizing energy (2023)

Problem: given **T tasks** and **n resources** with different cost functions C, find an assignment that minimizes the **total cost ΣE**.

Assumption: all cost functions are known



Find a task assignment $A_i \in \mathbb{N}$ to each resource $i \in R$ that minimizes the total cost ΣE while assigning all tasks among the resources and respecting their lower and upper limits.

$$\Sigma E \coloneqq \sum_{i \in R} C_i(A_i)$$

$$\sum_{i \in R} A_i = T$$

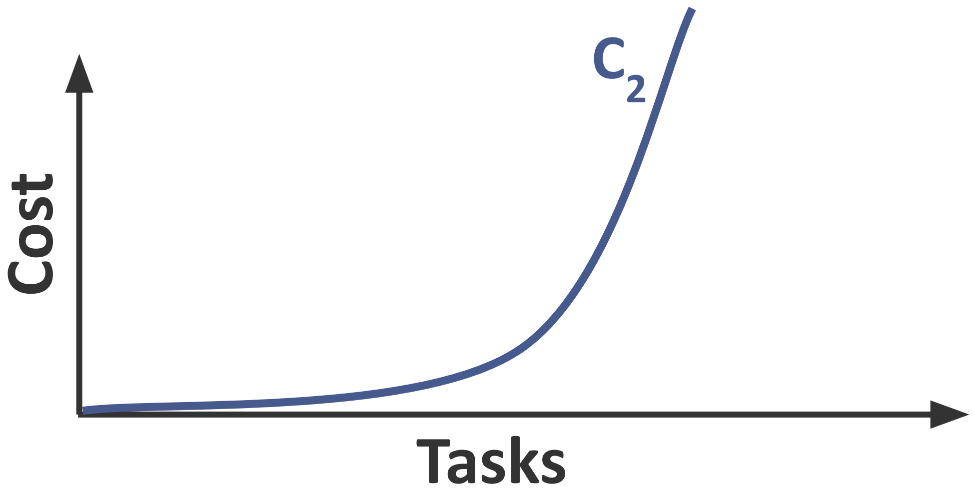
$$L_i \leq A_i \leq U_i, \ \forall i \in R$$

Pilla, Laércio Lima. "Scheduling algorithms for federated learning with minimal energy consumption." IEEE Transactions on Parallel and Distributed Systems 34.4 (2023): 1215-1226.

Optimizing energy (2023)

Problem: given **T tasks** and **n resources** with different cost functions C, find an assignment that minimizes the **total cost ΣE**.

Assumption: all cost functions are known **and superlinear**



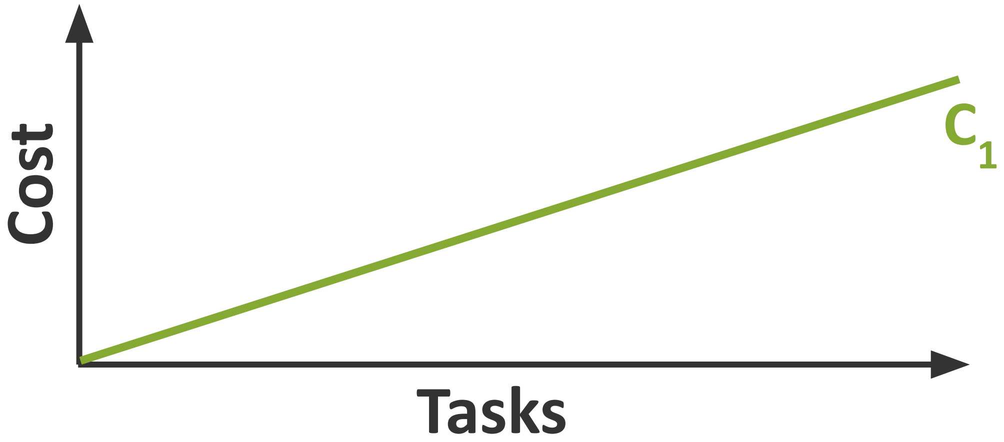
Solution in O(T log n): assign the **next task** to one of the resources that increases the total cost the least.

- similar to OLAR

Optimizing energy (2023)

Problem: given **T tasks** and **n resources** with different cost functions C, find an assignment that minimizes the **total cost ΣE**.

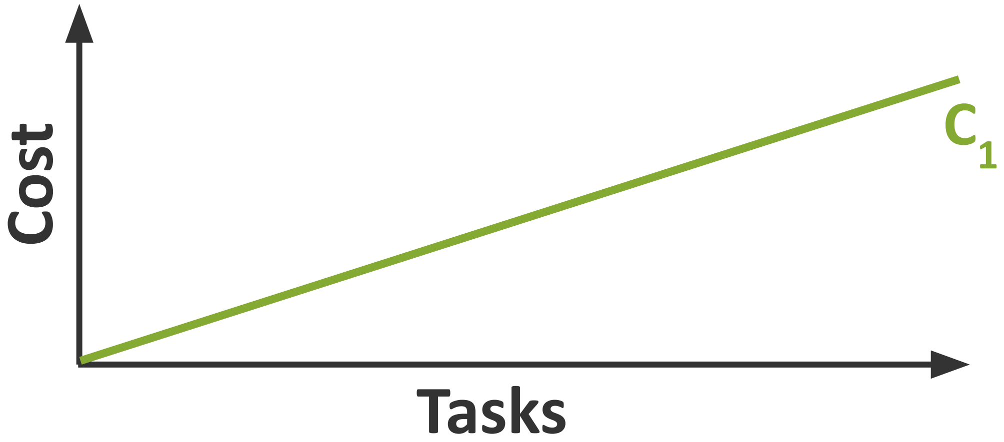
Assumption: all cost functions are known **and linear**



Optimizing energy (2023)

Problem: given **T tasks** and **n resources** with different cost functions C, find an assignment that minimizes the **total cost ΣE**.

Assumption: all cost functions are known **and linear**



Solution in O(n log n): assign the **most tasks possible** to one of the resources with the least cost per task.

Algorithm 3: MarCo.

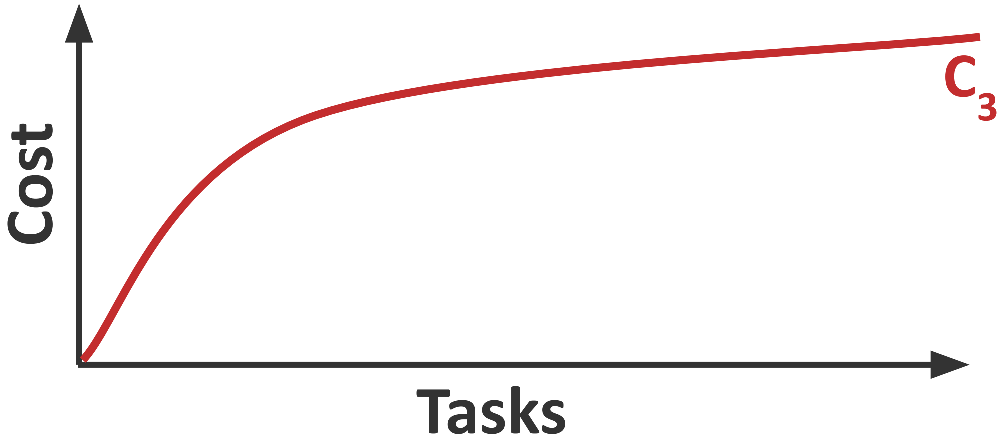
Input: Set of resources $\mathcal{R}$, number of tasks to schedule T, set of upper limits $\mathcal{U}$, set of cost functions $\mathcal{C}$.
Output: Optimal schedule X.

for all $i \in \mathcal{R}$ **do**
  $x_i \leftarrow 0$ ▷ *All resources start without any tasks.*
end for
$t \leftarrow 0$
while $t < T$ **do**
  $k \leftarrow \arg\min_{i \in \mathcal{R}, x_i \neq U_i} M_i(1)$
  $x_k \leftarrow \min(U_k, T - t)$ ▷ *Assigns the most tasks possible.*
  $t \leftarrow t + x_k$
end while
return X

Optimizing energy (2023)

Problem: given **T tasks** and **n resources** with different cost functions C, find an assignment that minimizes the **total cost ΣE**.

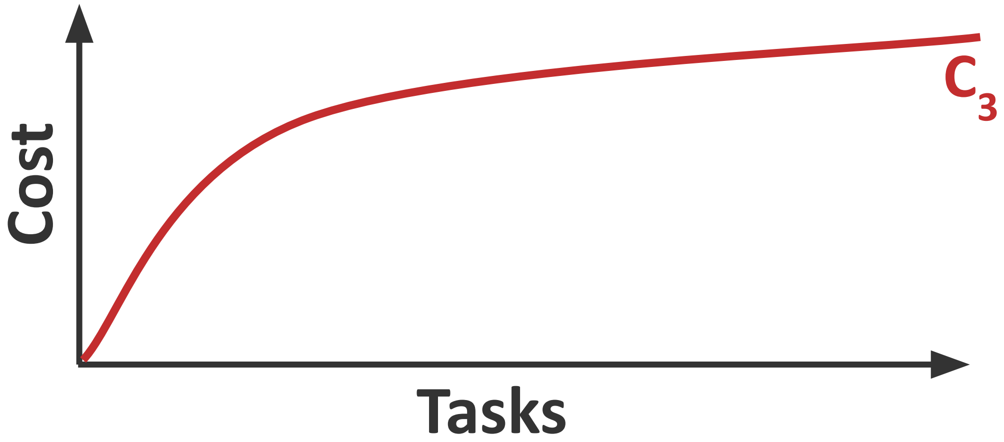
Assumption: all cost functions are known **and sublinear**



Optimizing energy (2023)

Problem: given **T tasks** and **n resources** with different cost functions C, find an assignment that minimizes the **total cost ΣE**.

Assumption: all cost functions are known **and sublinear**



Solution without upper limits in O(n): assign **all tasks** to one of the resources with the least total cost.

Algorithm 4: MarDecUn.

Input: Set of resources $\mathcal{R}$, number of tasks to schedule T, set of upper limits $\mathcal{U}$, set of cost functions $\mathcal{C}$.
Output: Optimal schedule X.

for all $i \in \mathcal{R}$ **do**
  $x_i \leftarrow 0$ ▷ *All resources start without any tasks.*
end for
$k \leftarrow \arg\min_{i \in \mathcal{R}} C_i(T)$
$x_k \leftarrow T$ ▷ *Assigns all tasks to the same resource.*
return X

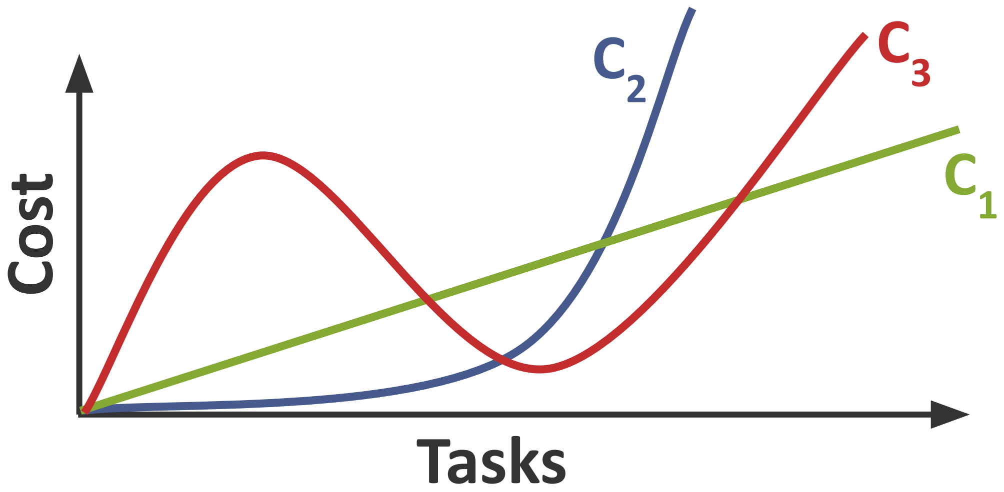
Solution with upper limits in O(Tn²)

- More complex (and related to our previous mystery) ;)

Optimizing energy (2023)

Problem: given **T tasks** and **n resources** with different cost functions C, find an assignment that minimizes the **total cost ΣE**.

Assumption: all cost functions are known



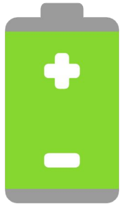
Solution in $O(T^2n)$: **$(MC)^2$MKP -** Knapsack w/ DP implementation

- **M**ultiple-**c**hoice
  - One item per group (resource)
- **M**inimum-**c**ost
  - Minimizes cost (instead of maximizing profit)
- **M**aximum **K**napsack **P**acking
  - Fills the knapsack as much as possible

Optimizing **time and energy (2024)**

Problems: given **T tasks** and **n resources** with different cost functions **P and E**, find an assignment that

- minimizes the **maximal cost C_{max}** and then **total cost ΣE**
  - Minimal Makespan and Energy Consumption (**MEC**)
- or **vice-versa within a deadline D**
  - Minimal Energy Consumption and Makespan under Time Constraint (**ECMTC**)



$$C_{\max} := \max_{i \in \mathcal{R}} P_i(x_i) \quad (1)$$

$$\Sigma E := \sum_{i \in \mathcal{R}} E_i(x_i) \quad (2)$$

$$\begin{aligned} &\text{lex } \min_X \; C_{\max}, \Sigma E && (3a)\\ &\text{subject to } \sum_{i \in \mathcal{R}} x_i = T, && (3b)\\ &x_i \in A_i, \quad \forall i \in \mathcal{R} && (3c) \end{aligned}$$

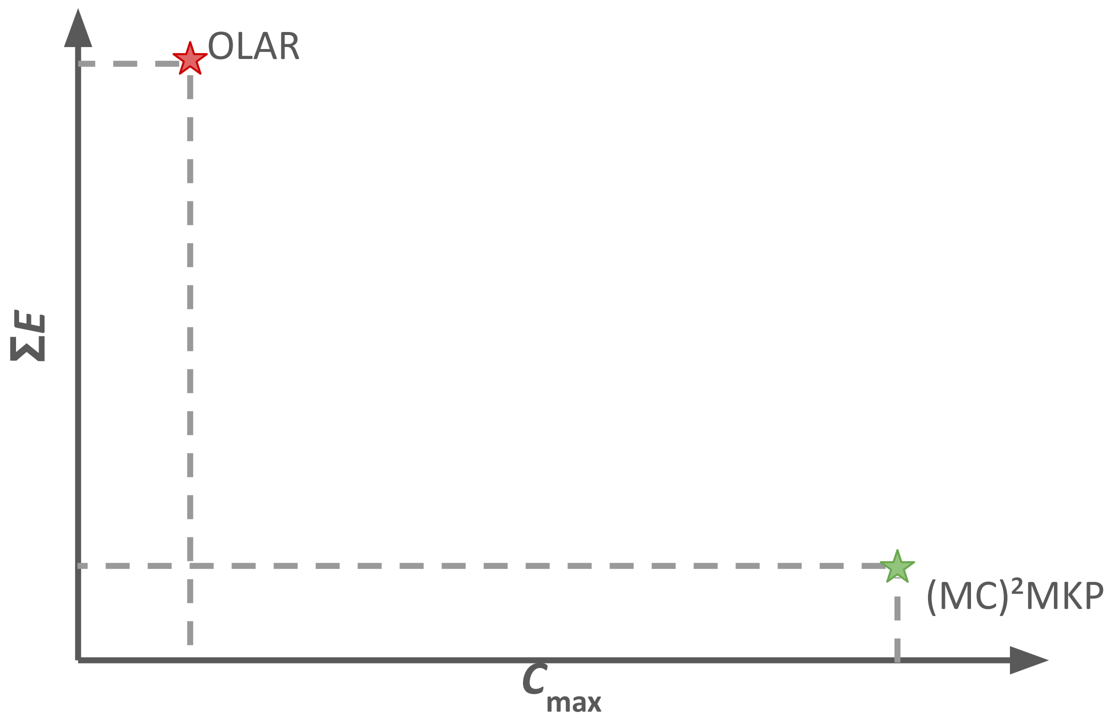
$$\begin{aligned} &\text{lex } \min_X \; \Sigma E, C_{\max} && (4a)\\ &\text{subject to } \sum_{i \in \mathcal{R}} x_i = T, && (4b)\\ &C_{\max} \le D, && (4c)\\ &x_i \in A_i, \quad \forall i \in \mathcal{R} && (4d) \end{aligned}$$

Nunes, Alan L., et al. "Optimal time and energy-aware client selection algorithms for federated learning on heterogeneous resources." 2024 IEEE 36th International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD). IEEE, 2024.

Optimizing time and energy (2024)

Problems: given **T tasks** and **n resources** with different cost functions **P and E**, find an assignment that

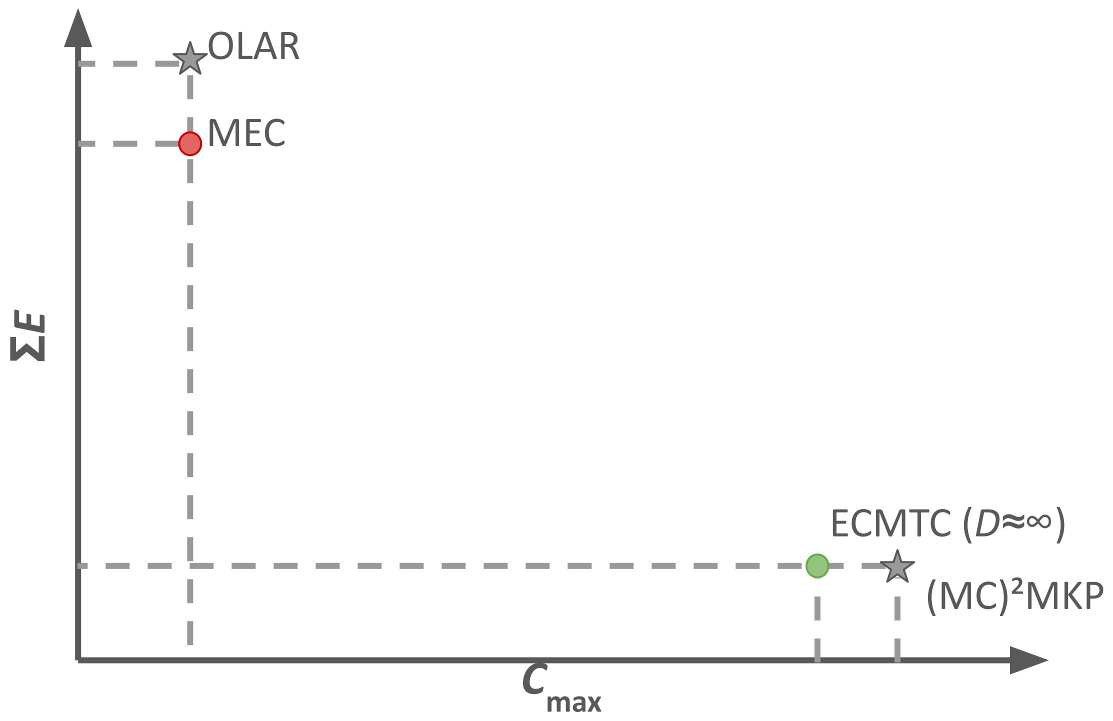
- minimizes the **maximal cost C_{max}** and then **total cost ΣE**
  - Minimal Makespan and Energy Consumption (**MEC**)
- or **vice-versa within a deadline D**
  - Minimal Energy Consumption and Makespan under Time Constraint (**ECMTC**)



Optimizing time and energy (2024)

Problems: given **T tasks** and **n resources** with different cost functions **P and E**, find an assignment that

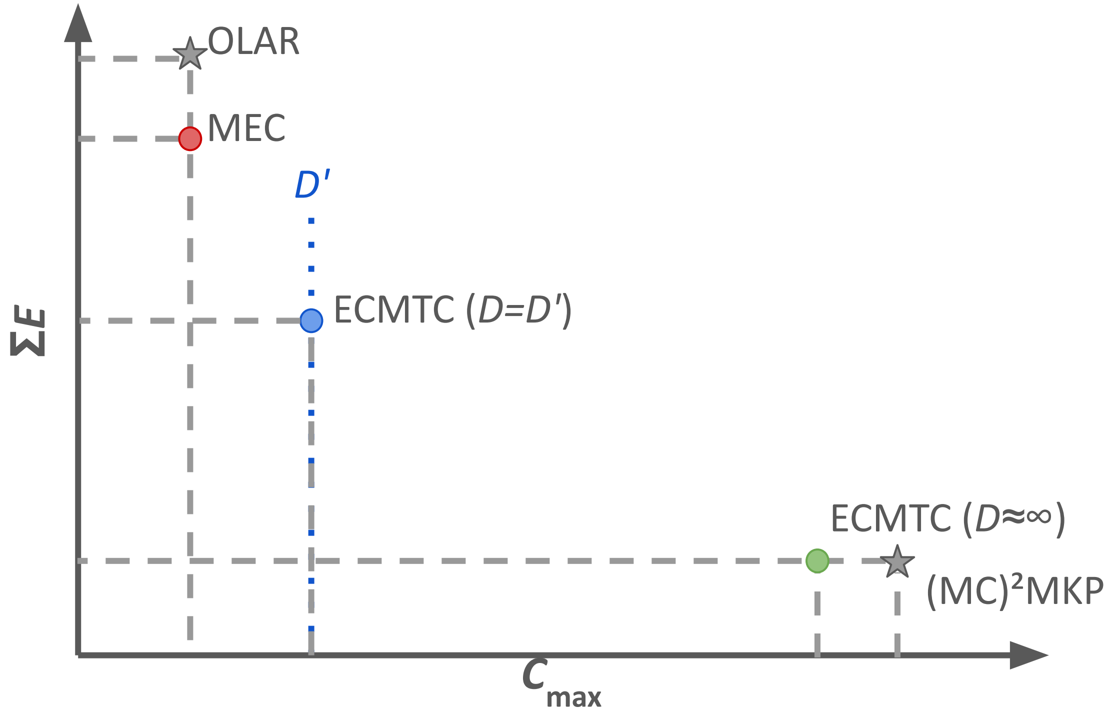
- minimizes the **maximal cost C_{max}** and then **total cost ΣE**
  - Minimal Makespan and Energy Consumption (**MEC**)
- or **vice-versa within a deadline D**
  - Minimal Energy Consumption and Makespan under Time Constraint (**ECMTC**)



Optimizing time and energy (2024)

Problems: given **T tasks** and **n resources** with different cost functions **P and E**, find an assignment that

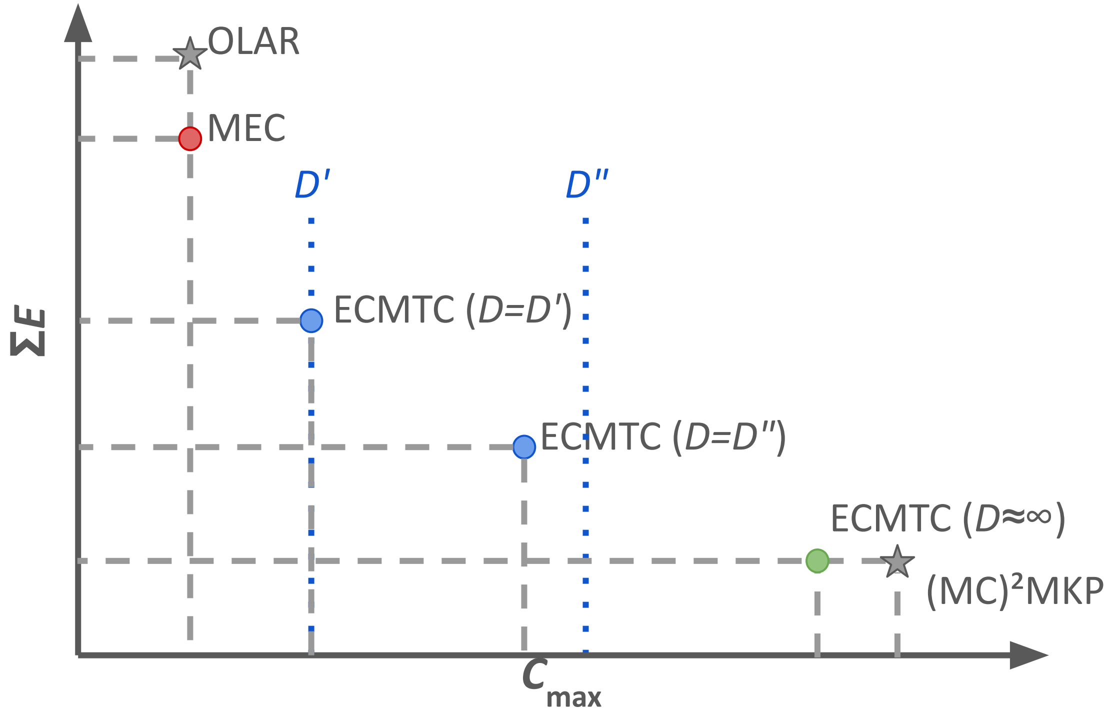
- minimizes the **maximal cost C_{max}** and then **total cost ΣE**
  - Minimal Makespan and Energy Consumption (**MEC**)
- or **vice-versa within a deadline D**
  - Minimal Energy Consumption and Makespan under Time Constraint (**ECMTC**)



Optimizing time and energy (2024)

Problems: given **T tasks** and **n resources** with different cost functions **P and E**, find an assignment that

- minimizes the **maximal cost C_{max}** and then **total cost ΣE**
  - Minimal Makespan and Energy Consumption (**MEC**)
- or **vice-versa within a deadline D**
  - Minimal Energy Consumption and Makespan under Time Constraint (**ECMTC**)



All solutions are Pareto optimal!

Optimizing time and energy (2024)

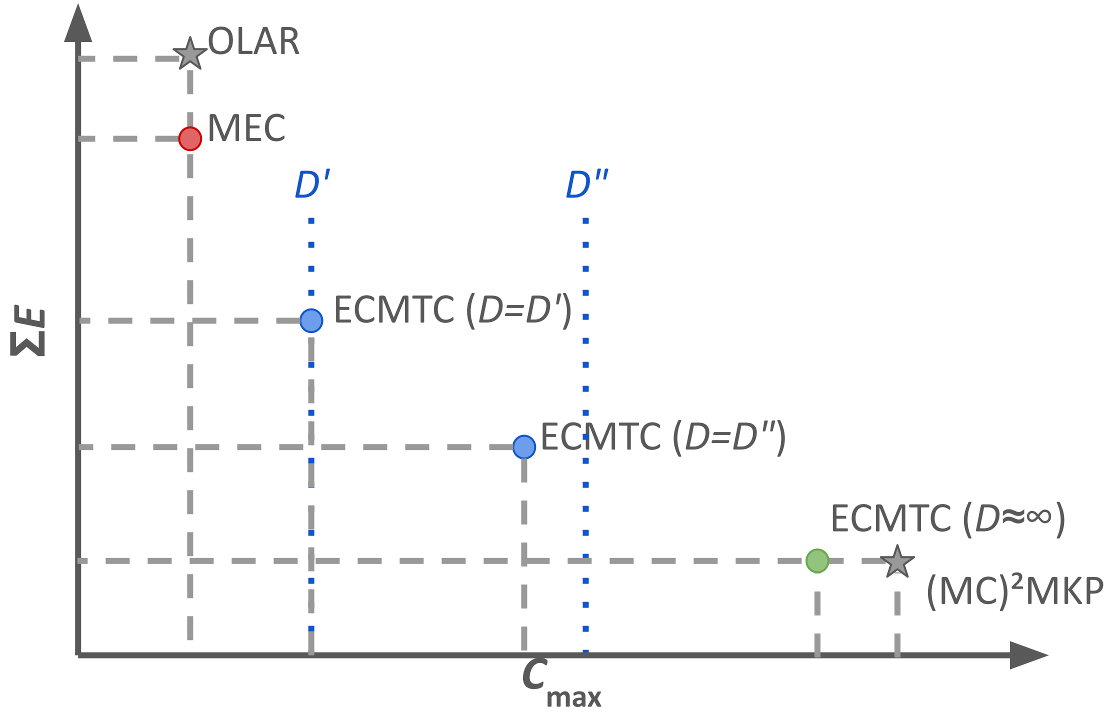
Solutions in $O(T^2n)$ based on $(MC)^2MKP$

MEC

- Find minimal makespan, set it as deadline
- Filter assignments based on deadline
- $(MC)^2MKP$

ECMTC

- Filter assignments based on deadline
- As $(MC)^2MKP$, but solve ties for the minimal total time



Optimizing **over multiple rounds** (**2025**)

General problem:

1. get an optimal schedule
2. reuse it over multiple rounds
3. only see "the same" data
4. the model requires more rounds to converge/achieve a given accuracy
5. total **time and energy increase**

Optimizing over multiple rounds (2025)

General problem:

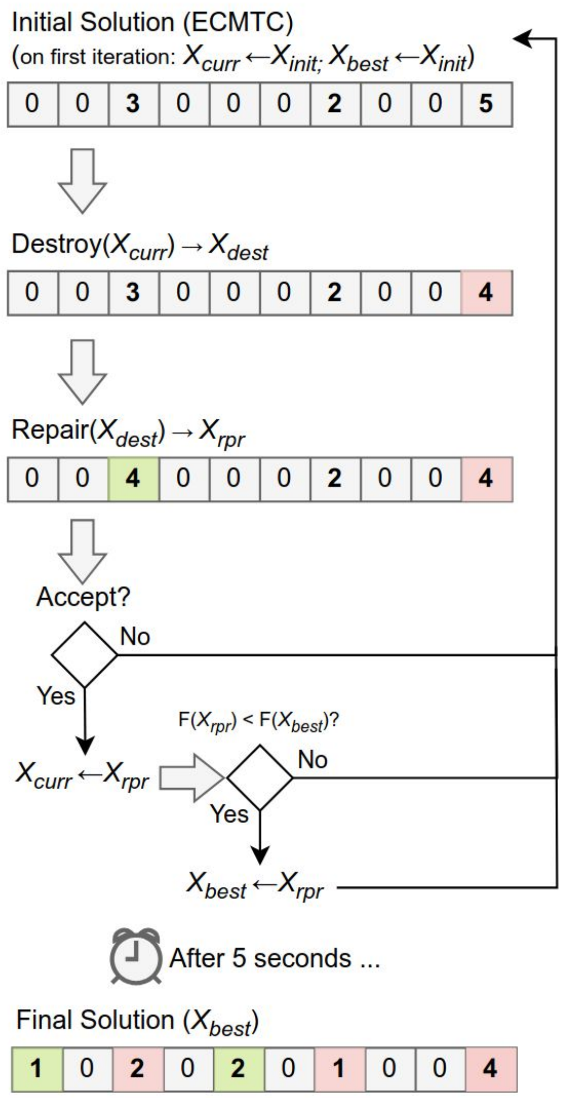
1. get an optimal schedule
2. reuse it over multiple rounds
3. only see "the same" data
4. the model requires more rounds to converge/achieve a given accuracy
5. total **time and energy increase**

Possible approaches:

- Select a random subset of resources before computing each schedule
- **Change a schedule to add diversity**
  - **Large Neighborhood Search**
- etc

Nunes, Alan L., et al. "MetaCS-FL: A Metaheuristic-Based Framework for Client Selection in Federated Learning Systems." to be submitted.

Optimizing over multiple rounds (2025)



Possible approaches:

- Select a random subset of resources before computing each schedule
- **Change a schedule to add diversity**
  - **Large Neighborhood Search**
- etc

Nunes, Alan L., et al. "MetaCS-FL: A Metaheuristic-Based Framework for Client Selection in Federated Learning Systems." to be submitted.

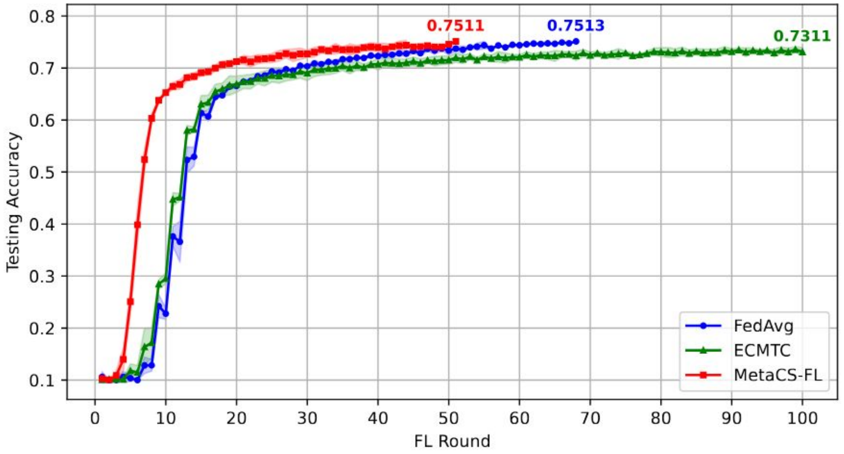
Optimizing over multiple rounds (2025)

Preliminary results: M_{total} = time; Σ_{total} = energy; S_{fair} = Jain's Fairness Index; figures -> rounds to accuracy

TABLE VI: Performances for CIFAR-10 IID. NaN means the method did not achieve the target testing accuracy.

CS Approach	M_{total}	Σ_{total}	S_{fair}	RoA@0.75
FedAvg (Baseline)	10,696.79	2,077,926.08	0.98	62 ± 5.4
ECMTC	8,615.93 (−19.45%)	2,019,645.24 (−2.8%)	0.45 (−53.85%)	NaN
MetaCS-FL	3,762.31 (−64.83%)	897,893.30 (−56.79%)	0.67 (−31.54%)	46 ± 3.6

Units: M_{total} in seconds; Σ_{total} in joules.

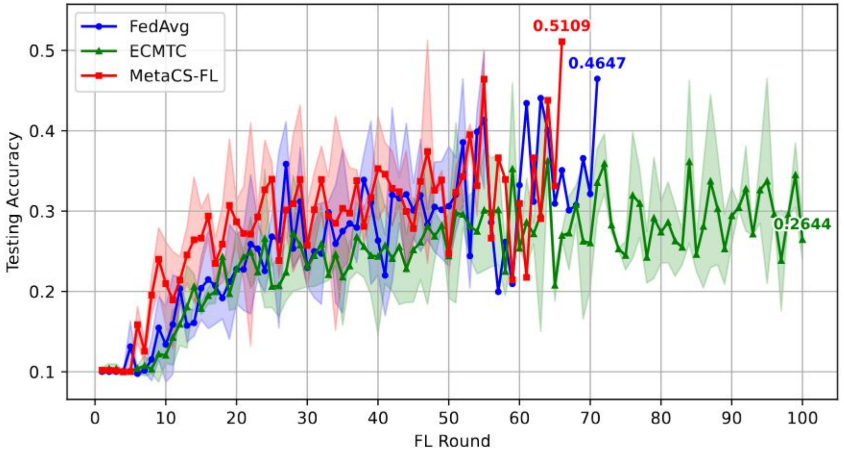


(a) CIFAR-10 IID

TABLE VII: Performances for CIFAR-10 non-IID. NaN means the method did not achieve the target testing accuracy.

CS Approach	M_{total}	Σ_{total}	S_{fair}	RoA@0.45
FedAvg (Baseline)	10,010.54	1,976,410.31	0.98	59 ± 8.3
ECMTC	8,389.58 (−16.19%)	1,822,920.56 (−7.77%)	0.46 (−53.12%)	NaN
MetaCS-FL	5,106.6 (−48.99%)	1,151,546.44 (−41.74%)	0.71 (−27.76%)	56 ± 7.8

Units: M_{total} in seconds; Σ_{total} in joules.



(b) CIFAR-10 non-IID

Concluding remarks

We take FL and optimize its performance and/or energy consumption by controlling how much work each device should do.

Advantages of our approaches

★ Given the required information, we can find optimal solutions
  ○ Time, energy, both
★ It should be easier to control how much work to give to a resource than to control other aspects of the resources
  ○ e.g. DVFS

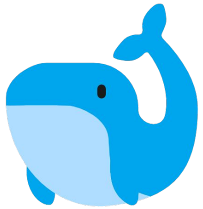
Limitations

❖ We require more information from the resources
❖ Optimizing for one objective can lead to worse outcomes for other objectives

Concluding remarks

We take FL and optimize its performance and/or energy consumption by controlling how much work each device should do.

- Can we get energy or carbon-equivalent emissions **information**?
  - Should we? (privacy issues)
  - Can we trust this information?
- **How far from the best performance** (%) do I accept to be if it improves my energy consumption (and vice-versa)?
- How to optimize considering more options (e.g., edge device offloading)?
- **White whales**
  - Optimize energy if monotonically increasing [faster than $O(T^2n)$]?
  - Optimize time given an energy budget in $O(T^2n)$ or less?



That’s all folks!