Exploring the Trade-offs between Energy and Performance of Federated Learning Algorithms

— First Year —

Presenter

Mai Huong Do

Advisors

Georges Da Costa
Millian Poquet

IRIT - Université Toulouse III

July 17, 2024

Table of Contents

Introduction

Thesis context

- ANR DELIGHT (a**D**vancing f**E**derated **L**earn**I**ng while reducin**G** t**H**e carbon foo**T**print) project:
  - Machine Learning (ML) is increasingly deployed across society.
  - [1]The amount of computations used to train learning models has increased 300,000 in 6 years.
  - Federated learning (FL) - one of the most growing research in ML, requires a huge supply of [2]energy, which is difficult to meet in the current context.
- Target: incorporate energy efficiency as one of the metrics of FL to push FL towards sustainability.
- This thesis is one part of ANR DELIGHT project: explore the tradeoffs between energy and performance of FL algorithms.

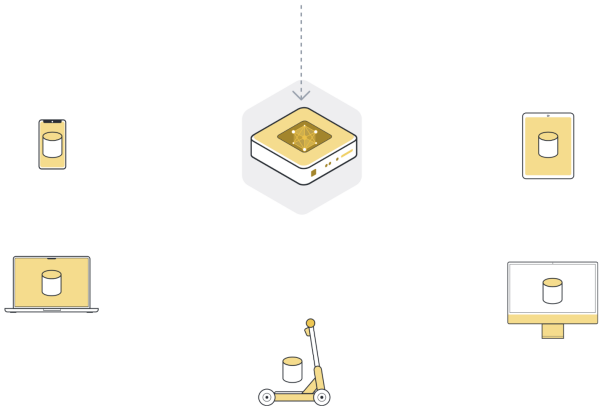
[1]Schwartz, Roy, et al. "Green ai." Communications of the ACM 63.12 (2020): 54-63.
[2]communication and computation, etc.

Objective and Planning

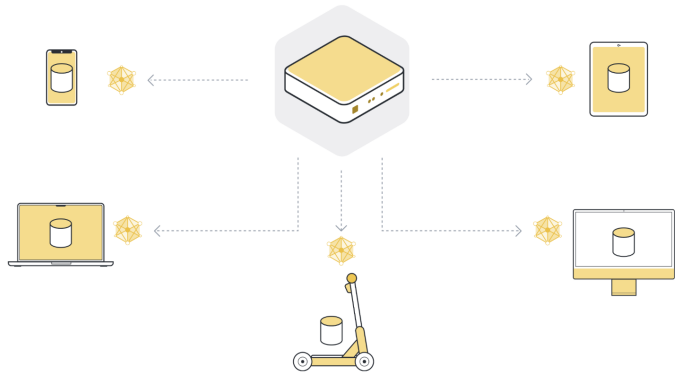
Objective: develop method for estimating energy, build a automatic framework to explore the trade-offs between Energy and FL performance.
Phases of the thesis:

- Set up an experimental environment on Grid'5000 (g5k) to gather performance and energy metrics.
  - Create a use-case for the Flower framework.
  - Build a reproducible and automated framework for obtaining metrics for this use case
- Propose, formulate energy model, and implement the different leverages.
- Explore the impact of the leverages on both energy and performance.

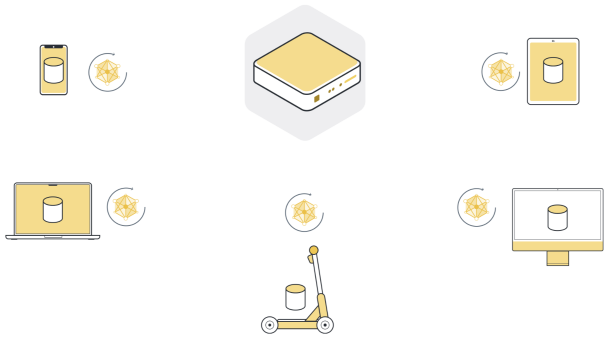
Federated Learning framework



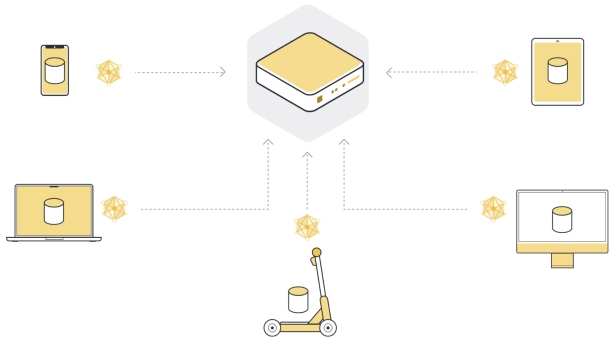
Step 1: Global model init.



Step 2: Send model to clients



Step 3: Local training



Step 4: Return and aggregate in global

Flower-adapt

Solve the heterogeneous issues:

- Any workload
- Any ML framework
- Any programming language

Grid'5000

- A large amount resources for experiment-driven research of computer science, includes: 10 sites, 72 clusters, 800 nodes, more than 15000 cores.
- Highly re-configurable and controllable.
- Advanced monitoring and measurement features.

Progress

Context of survey

- Range of the survey: 2019 - May 2024.
- Keywords: energy consumption, FL, trade-offs, performance.
- Current amount: 31 articles.
- Try to find reproducible works (at least open source).

Classification by objectives

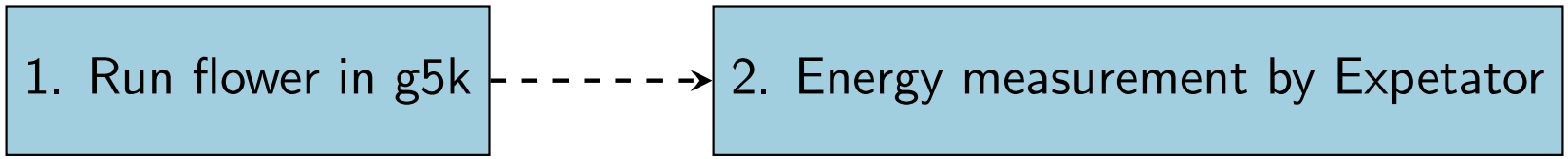
No.	Year	Energy				FL performance			Etc.			
		Scope of E		Type of Energy		Acc.	Loss	Time processing	Carbon equivalent	Bandwidth	Computing resource	Memory resource
		Client	Server	Communication	Computation							
1	2019	x		x	x	x		x			x	
2	2020	x		x			x					
3	2020	x		x	x		x	x			x	
4	2020	x		x	x			x			x	x
5	2020	x		x	x	x						
6	2021	x	x	x	x		x		x			
7	2021	x		x	x	x				x	x	x
8	2021	x		x	x		x	x				
9	2021	x		x	x	x		x		x	x	
10	2021	x		x	x		x	x		x	x	
11	2021	x		x	x	x		x			x	
12	2021	x		x	x					x	x	
13	2021	x			x		x	x				x
14	2022	x		x	x		x	x		x	x	
15	2022			x	x	x						
16	2022	x		x	x		x					
17	2022	x		x	x		x	x		x	x	
18	2022	x		x	x		x	x			x	
19	2022	x		x	x							
20	2023	x	x	x	x	x						
21	2023	x	x	x	x				x			
22	2023	x		x	x					x	x	
23	2023	x	x	x	x			x			x	
24	2023	x		x	x		x	x		x		
25	2023	x	x	x	x		x	x		x	x	
26	2023						x	x				
27	2023	x		x	x						x	
28	2024	x	x	x	x					x	x	
29	2024	x		x	x	x		x				
30	2024	x		x	x	x		x		x	x	

See more

Conclusion

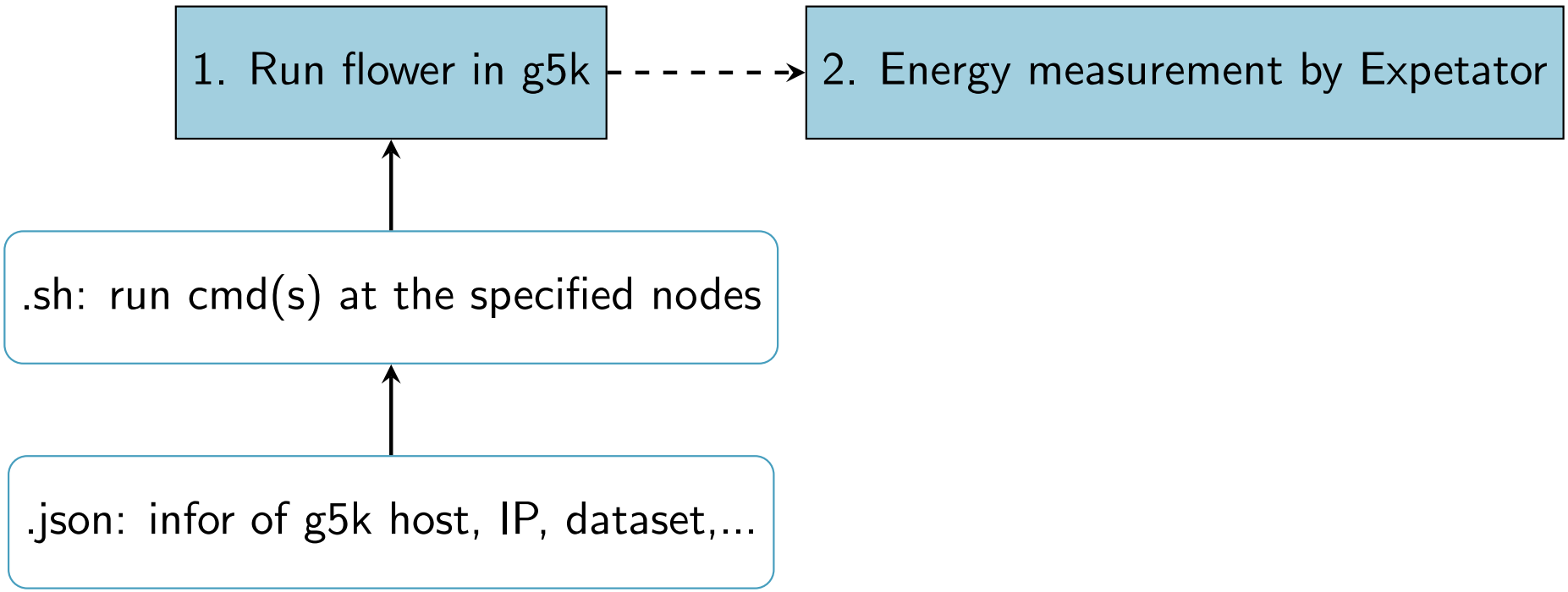
- Confirming the correctness of our project direction, which considered not only computing but also communication costs.
- Lack of modeling energy consumption in server side.
- Carbon equivalent has not really been given due attention.
- Similar to memory resources (queuing delay) is the main cause of packet loss.

Implementation



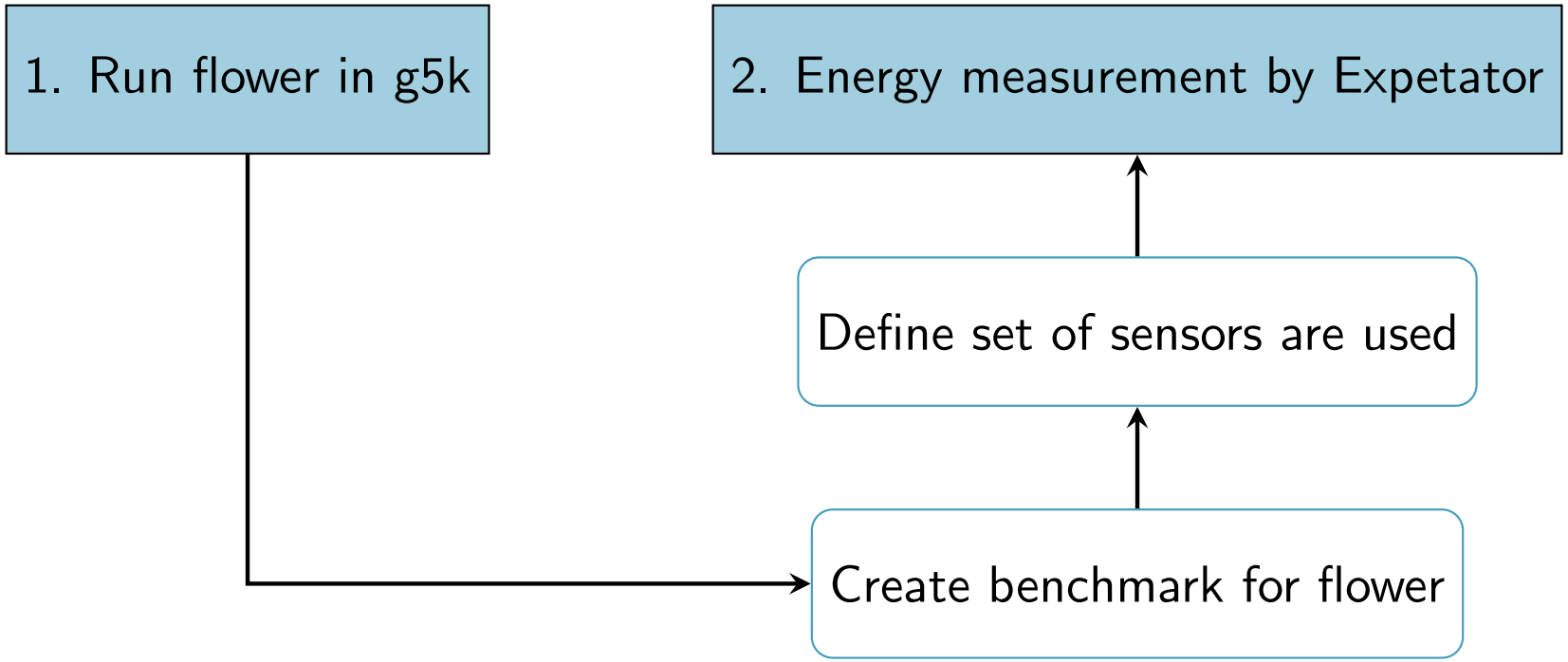
▸ Flower implement - see more ▸ Set of sensors - see more

Implementation



▸ Flower implement - see more ▸ Set of sensors - see more

Implementation



▸ Flower implement - see more ▸ Set of sensors - see more

Results

N.Round	N.Clients	N.Server	Dataset	Time	D. Loss
1	10 train: 2 evaluate: 10	1	Cifar10	18.43 (s)	1.4628

```
mdo@hercule-1:/tmp/flower_hercule-1.lyon.grid5000.fr_1720791124_mojitos$ cat orion-3.lyon.grid5000.fr_flower_1720791142
#timestamp br0:rxp br0:rxb br0:txp br0:txb package-00 core0 dram1 package-11 core0 dram1 user nice system idle iowait irq softirq steal guest guest_nice
5470.965138854 0 0 0 0 0 0 6653 0 0 6256 0 0 0 2 0 0 0 0 0 0
5471.000093755 1 52 1 230 386470 101283 155967 381892 85506 130074 0 0 0 83 0 0 0 0 0 0
5471.100118439 1 52 1 230 659741 114557 348874 659939 104273 313079 0 0 0 240 0 0 0 0 0 0
5471.200127857 1 52 1 230 781896 157615 372128 711939 114511 314483 0 0 0 240 0 0 0 0 0 0
5471.300116694 1 52 1 230 500188 89580 350964 499089 57752 307860 0 0 0 240 0 0 0 0 0 0
5471.400108125 3 204 1 230 935681 209263 394008 850359 159339 320220 0 0 0 240 0 0 0 0 0 0
5471.500108552 1 52 1 230 627592 104304 331815 593872 72735 310531 0 0 0 240 0 0 0 0 0 0
5471.600098623 2 102 1 230 871888 195653 395595 841387 171317 322706 0 0 0 240 0 0 0 0 0 0
5471.700117289 1 52 1 230 617537 97895 336255 597472 76717 306961 0 0 0 240 0 0 0 0 0 0
5471.800109487 1 52 1 230 906966 204152 404841 841052 141213 320296 0 0 0 239 0 0 0 0 0 0
5471.900103810 1 52 1 230 640089 110621 349851 617094 89503 303268 0 0 0 241 0 0 0 0 0 0
5472.000116905 2 104 1 230 686106 121087 349042 662762 99833 316527 0 0 0 240 0 0 0 0 0 0
5472.100108144 0 0 0 0 658276 123575 377925 625227 92616 316405 0 0 0 239 0 0 0 0 0 0
5472.200116572 0 0 0 0 836383 159721 373699 818684 132867 303512 0 0 0 240 0 0 0 0 0 0
5472.300074911 0 0 0 0 808353 183981 379558 793889 161460 345472 0 0 0 240 0 0 0 0 0 0
5472.400112917 0 0 0 0 1078787 313231 424142 927976 188116 335371 0 0 0 240 0 0 0 0 0 0
5472.500123546 0 0 0 0 568909 95241 360668 539767 68066 311583 0 0 0 240 0 0 0 0 0 0
5472.600109562 0 0 0 0 546267 88160 347730 535952 67303 296494 0 0 0 240 0 0 0 0 0 0
5472.700101418 0 0 0 0 626433 108622 360730 612853 85064 317900 0 0 0 240 0 0 0 0 0 0
5472.800127204 0 0 0 0 619291 100123 351315 563112 70018 306045 0 0 0 240 0 0 0 0 0 0
5472.900126680 1 80 0 0 1163042 256517 376354 1129641 227024 324996 0 0 0 240 0 0 0 0 0 0
```

Meeting among PhD/Master students

Name	Team/ Status	Done works	On going/Open issues	Remarks
Mai H	IRIT/ PhD	- Finish a semi-automatic framework combining Flower + Energy measurement - Deploy and test in Grid5000	- Try to change clients per round (toward clients selection) - Stability needs attention. - Need to find a solution to the problem of waiting time and synchronization when running multiple clients.	- Violation when exceed the day/night boundary of Grid'5000 –>Solved
Khaoula	LIA/ Master	- Working on client selection	None	- Not yet evaluating energy consumption - Will leave after 1 month
Ahmad	LIA/ PhD	- Implementing and testing FL schemes in the literature: FedAvg, FedBN, FedDrop, FedPMT and other variants - Suggesting and testing new schemes	None	- Not yet evaluating energy consumption
Oumayma	LAAS/ PhD	- Reducing the size of training data - Study the effect of data size on energy consumption	- How to measure the energy - Which type of energy should we consider - Violation when exceeding the limit of Grid'5000	- Open issues were solved during the meeting

Future works

- Framework automating & improving stability.
- Test with other FL strategies, multi-choice modification.
- Keep updating the survey and check their source code (framework).
- Propose a new idea to improve the framework of monitor energy consumption & FL performance.

Working situation

Difficulties encountered

- Many quality publications in this area have been published, increasing competition.
- There is still a lack of professionalism in using tools to help systematize research.
- It's time to come up with a proposal.

Publications

I have no one with SEPIA yet. However, I have 1 journal, 2 inter-confs, 1 domestic-conf during my Master's in Korea.

- Do, H. M., Tran, T. P., & Yoo, M. (2023). Deep Reinforcement Learning-Based Task Offloading and Resource Allocation for Industrial IoT in MEC Federation System. *IEEE Access*.
- Do, H. M., & Yoo, M. (2023). Delay Optimization for Augmented Reality Service using Mobile Edge Computing Federation system. In *2023 14th International Conference on Information and Communication Technology Convergence (ICTC)*
- Do, H. M., & Yoo, M. (2023). Energy Consumption Optimization in Mobile Edge Computing Federation based Deep Reinforcement Learning. In *Korean Society of Communication Studies Conference Proceedings*, 1834-1835.
- Do, H. M., & Yoo, M. (2022). Delay Optimization in Mobile Edge Computing Federation using Task Offloading and Resource Allocation. In *2022 13th International Conference on Information and Communication Technology Convergence (ICTC)* (pp. 767-770). IEEE.

Workshop

ITRC 2023

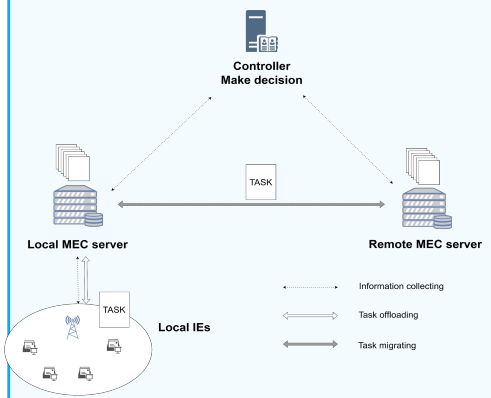
Deep Reinforcement Learning-based Task Offloading and Resource Allocation for Industrial IoT in MEC Federation System

Soongsil University

Introduction

- Propose a task offloading and resource allocation framework for IIoT system with MEC federation
- Formulate an optimization problem for both energy consumption and latency of our system model.
- Propose the DDPG-PER-based-RA algorithm to solve the optimization problem.
- Conduct the simulation to evaluate the performance of our proposed.

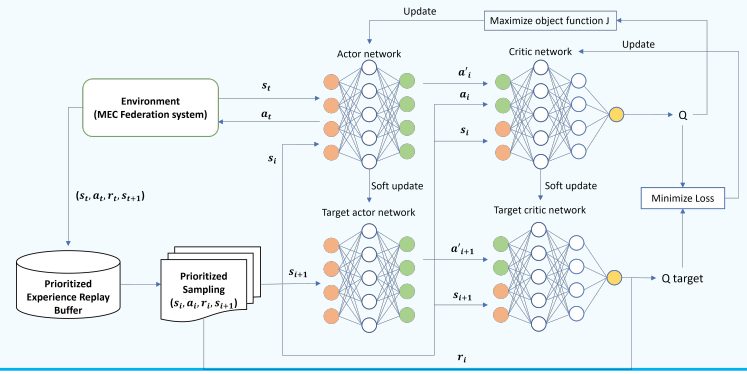
Proposed system model & Problem Formulation



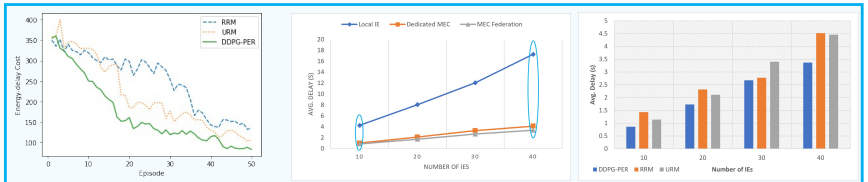
- **Objective**: Minimize the average energy-delay cost per task of whole system (IEs + fed servers) in T time slots.
- **Constraints:**
  i. Power allocate
  ii. Memory resource allocate
  iii. Bandwidth resource of wireless link allocate
  iv. Offloading decision

Methodology

- **Define**
  - **State**: status of current whole system
  - **Action**: resource allocation, task offloading decision
  - **Reward**: (each t) negative for the average cost per task
- **DDPG-PER-based-RA framework**



Results



Training model convergence comparison | Impact of number IEs on system model comparison | Impact of number IEs on training model comparison

Teaching

- No previous experience at a university.
- Plan to teach next semester (MPI, Monitors, or Petri network)

Training (14H recorded)

- ASR day - IRIT: PhD students present and discuss thesis topics (done).
- MOOC 1 (Reproducible research: methodological principles for transparent science): done
  - Understand the challenges and difficulties of reproducible research.
  - Gitlab for version tracking and collaborative work (used).
  - Notebook: Jupyter (used), RStudio, or Org-Mode (used).
  - Write a notebook to combine data analysis and its documentation effectively (COVID-19 analysis).
- MOOC 2 (Reproducible Research II: Practices and tools for managing computations and data): not finish.

Thank you!

Number

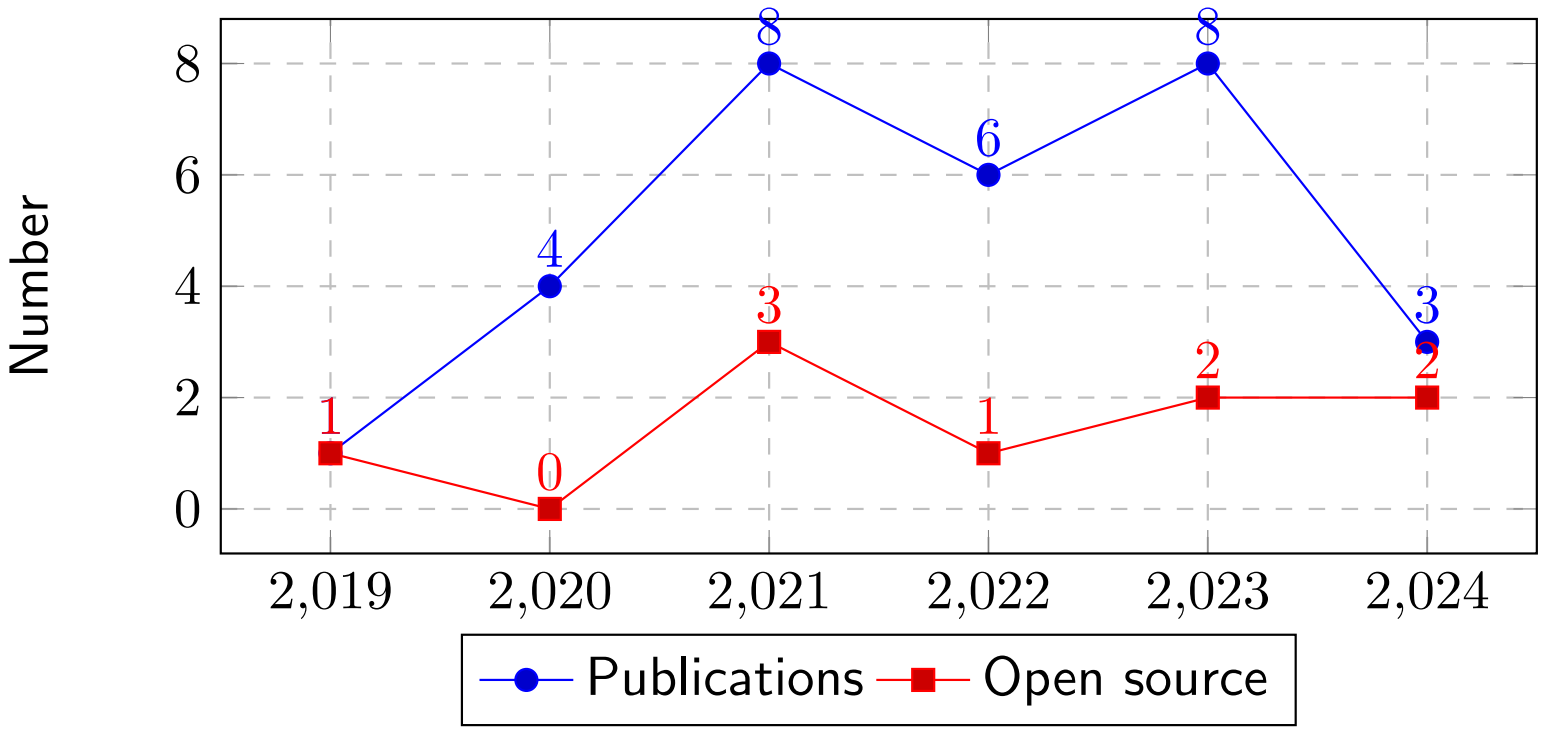


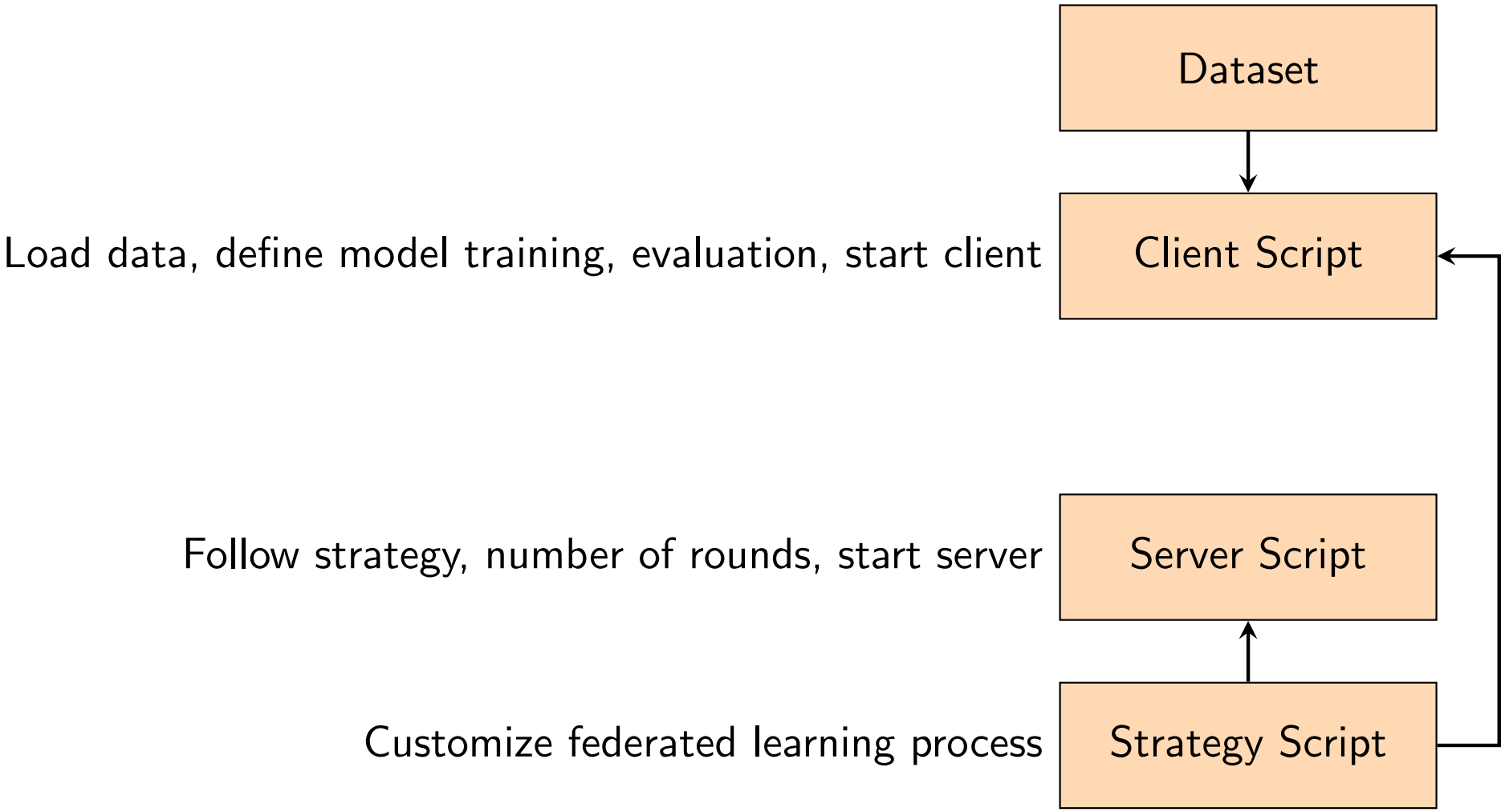
Figure: Number of Articles and Open source Frameworks Published Each Year

Source code available

Article	Proposal	Dataset	Language
1	Non-convex → decomposing sub problems	MNIST	Python
6	Framework for the analysis of energy and carbon footprints	MNIST FL radar	Python
11	Transfer to convex → use Lagrange duality	MNIST	Matlab
12	Non-convex → decomposing sub problems	MNIST FENIST Synthetic	Python
15	Closed-form expression for the expected convergence rate → the optimal power → optimize user selection and loss	MNIST	Matlab
22	Framework for the analysis of energy and carbon footprints in distributed and FL	MNIST CIFAR	Python
27	Intelligent Participant Selection (IPS): improve resource diversity Staleness-Aware Aggregation (SAA): improve resource efficiency	CIFAR10, OpenImage StackOverflow, Reddit, GG speech	Python
28	Pseudo-polynomial optimal solution Multiple-Choice Minimum-Cost Maximal Knapsack Packing Problem 4 algorithms for scenarios	Generate	Python
30	Pareto boundary for the convergence rate Nash bargaining solution and analyzing the derived convergence rate	MNIST	Python

Go back

Flower implement

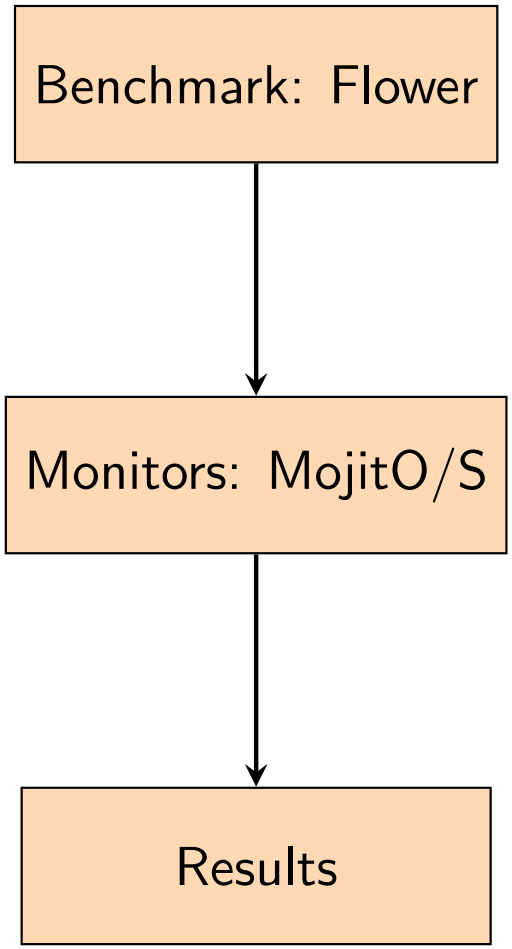


Go back

Energy measurement - Expetator

Run the Flower framework

Define set of sensors are used for measurement



▸ Energy parameters - see more

Energy measurement - Expetator

rxp	number of received packets
rxb	number of received bytes
txp	number of sent packets
txb	number of sent bytes
package	entire sockets
core0	or Power Plane 0, all processor cores on the socket
dram	RAM
idle	no activate status
user	CPU

Go back