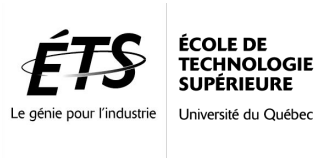
Reducing Cloud Environmental Impact is Teamwork! Enhancing User-Provider Collaboration through DevOps

Pierre Jacquet

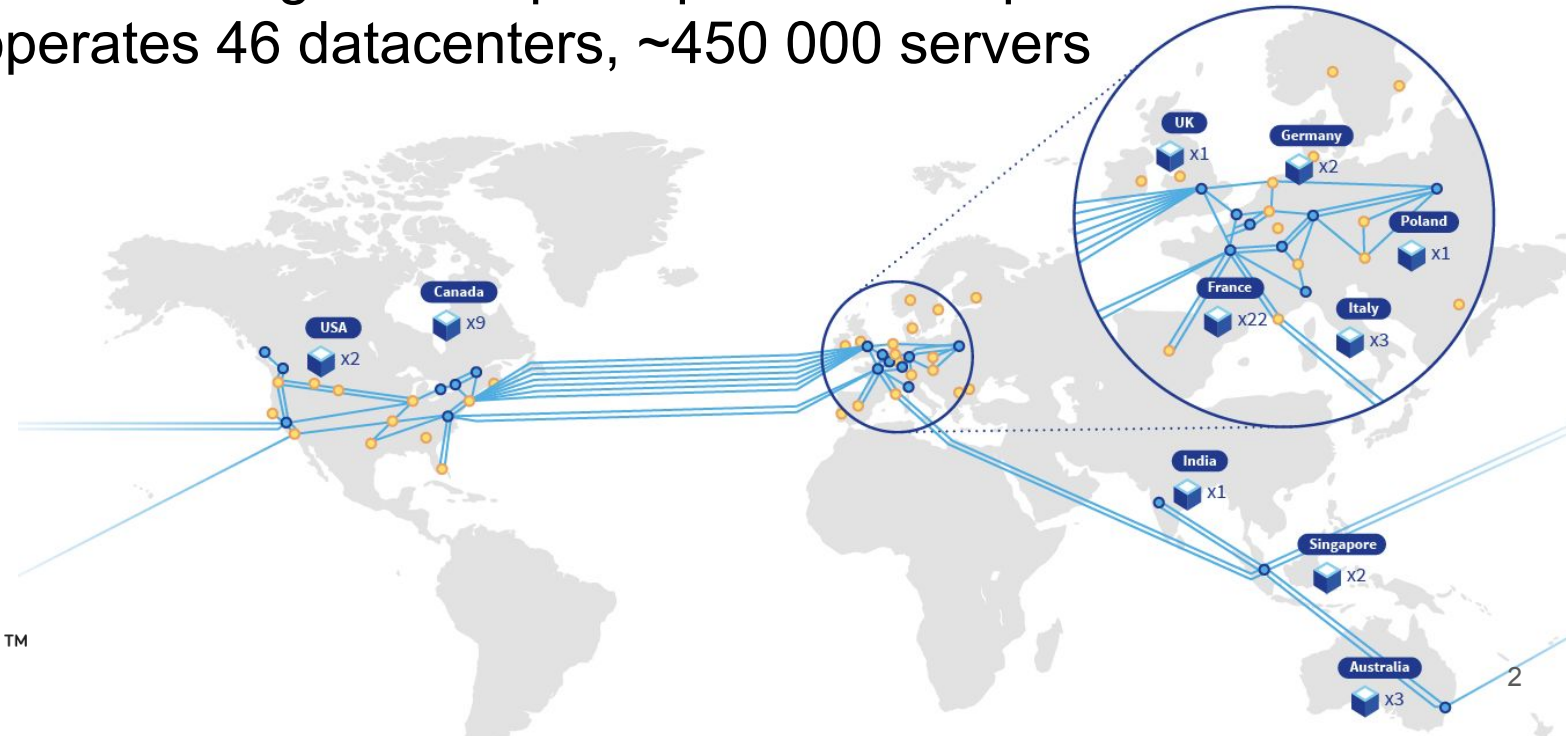
OVHcloud & ÉTS Montréal, Canada



DevOpsSustain 2026, Montréal, Québec

Introduction

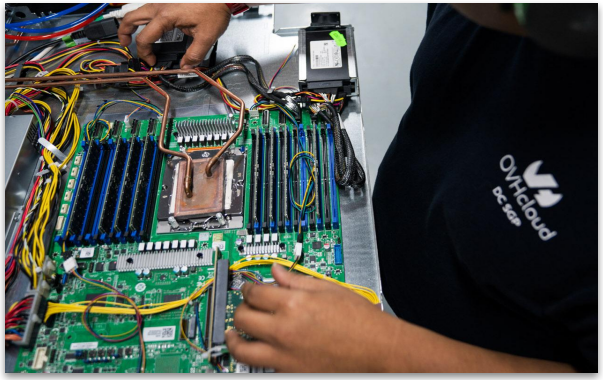
- On OVHcloud
  - One of the largest European public cloud provider
  - It operates 46 datacenters, ~450 000 servers

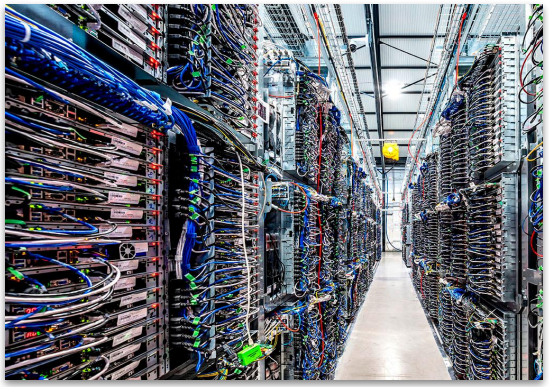


OVHcloud™

Introduction

Cloud platforms are seen as efficient

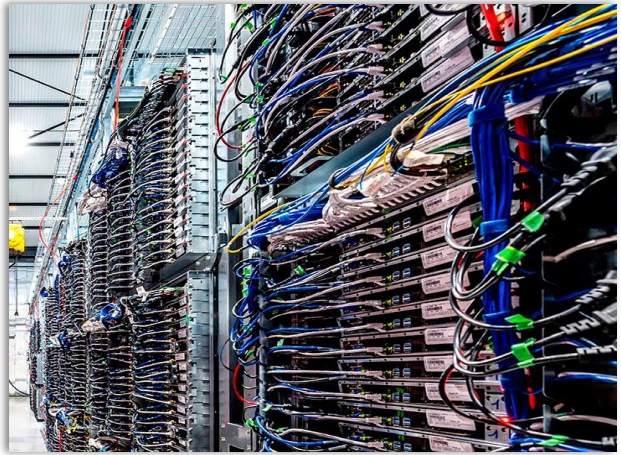




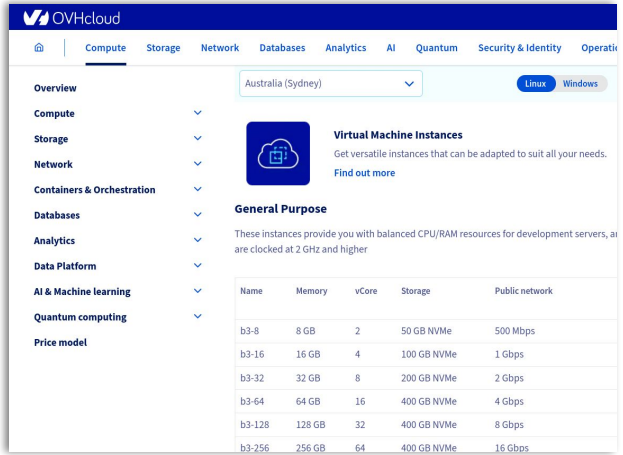
Introduction

But cloud platforms involve multiple actors

Provider side
(Architecture, product design)



User side
(Choose and consume product)



Introduction

- Are cloud platforms always as efficient as possible?
  - Are there cases where **users** do not use cloud platforms efficiently?
  - Are there cases where **providers'** choices are not efficient?
  - What can we do about it?

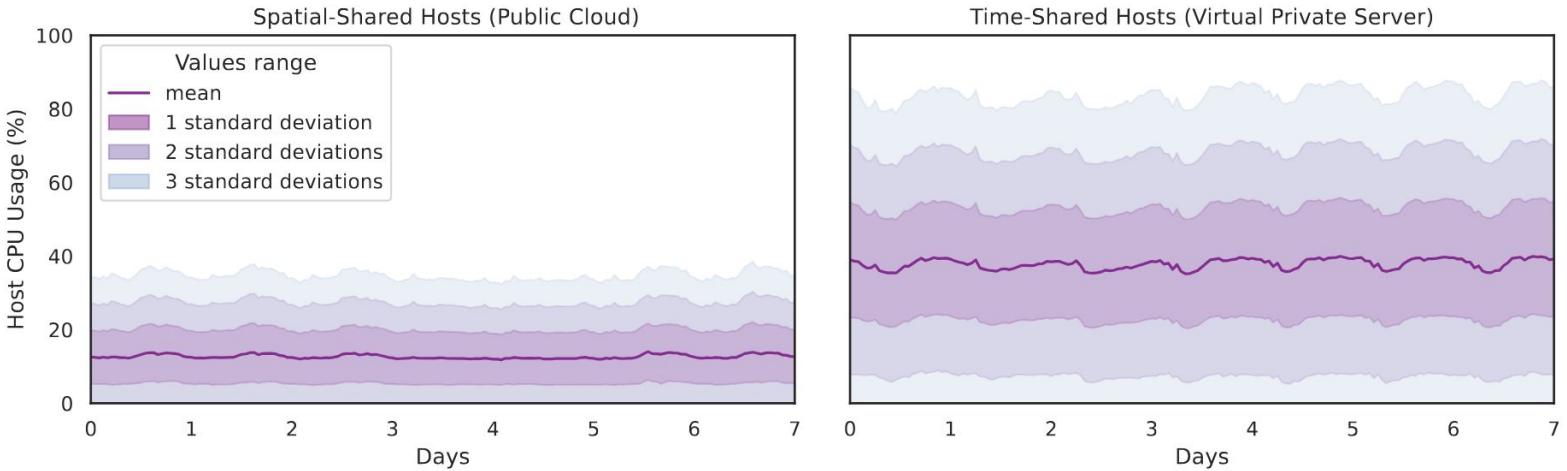
Introduction

We will explore a few needed collaborations between **Providers** and **Users**

- Improve the platform **energy efficiency**
- Limit **GPU expansion**
- Keep **VM instance** small

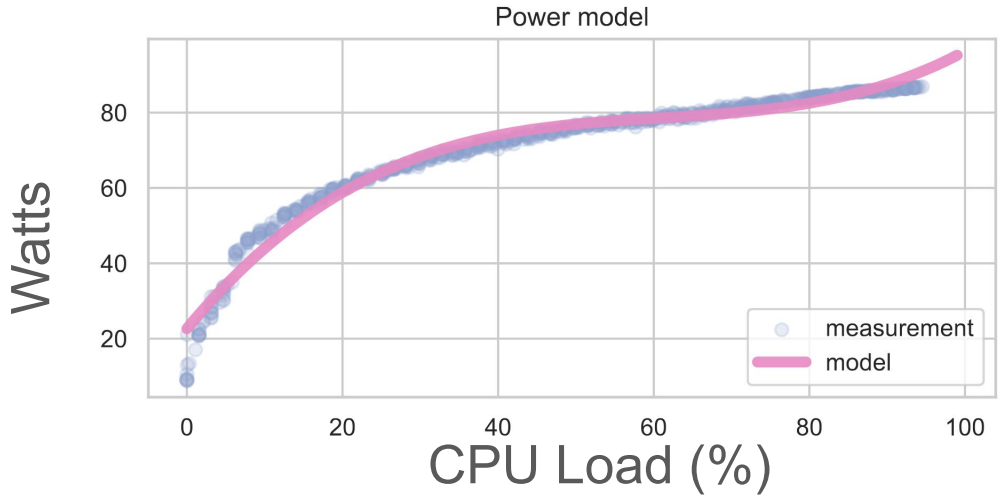
Improve **Energy efficiency**

- Servers have different **load**:
  - Hosts of dedicated products have **low usage**
  - Hosts of oversubscribed products have **higher usage**



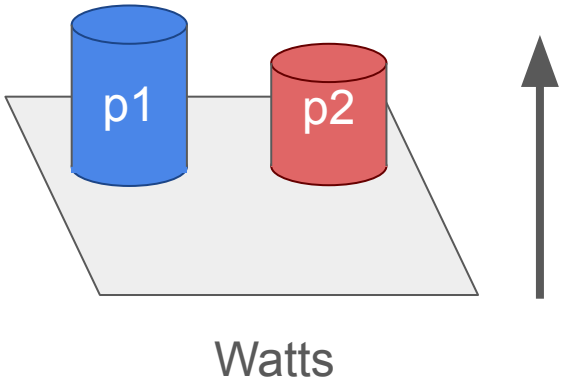
Improve **Energy efficiency**

- Servers have different **energy efficiency**:
  - Hosts of dedicated products have **low efficiency**
  - Hosts of oversubscribed products have **higher efficiency**

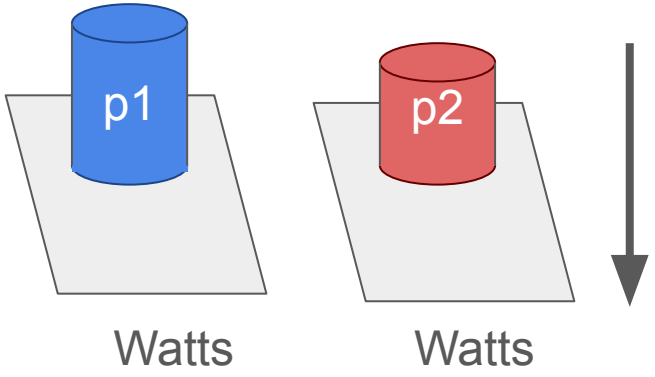


Improve **Energy efficiency**

- Carbon trackers report environmental impact
  - They rely on allocation ratios, not real measure. **Why?**



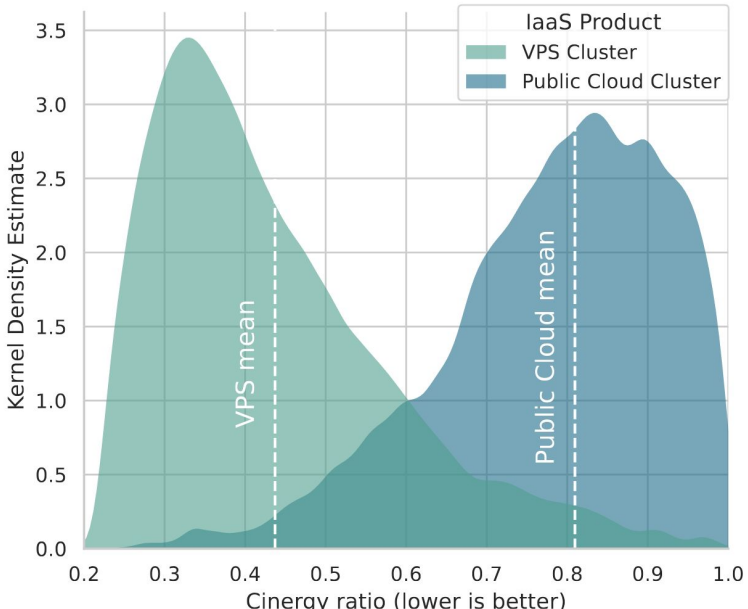
Bottom-up: Measure and dispatch
Cons: Highly contextual



Top-down: Isolated power
Cons: Ignore runtime conditions

Improve **Energy efficiency**

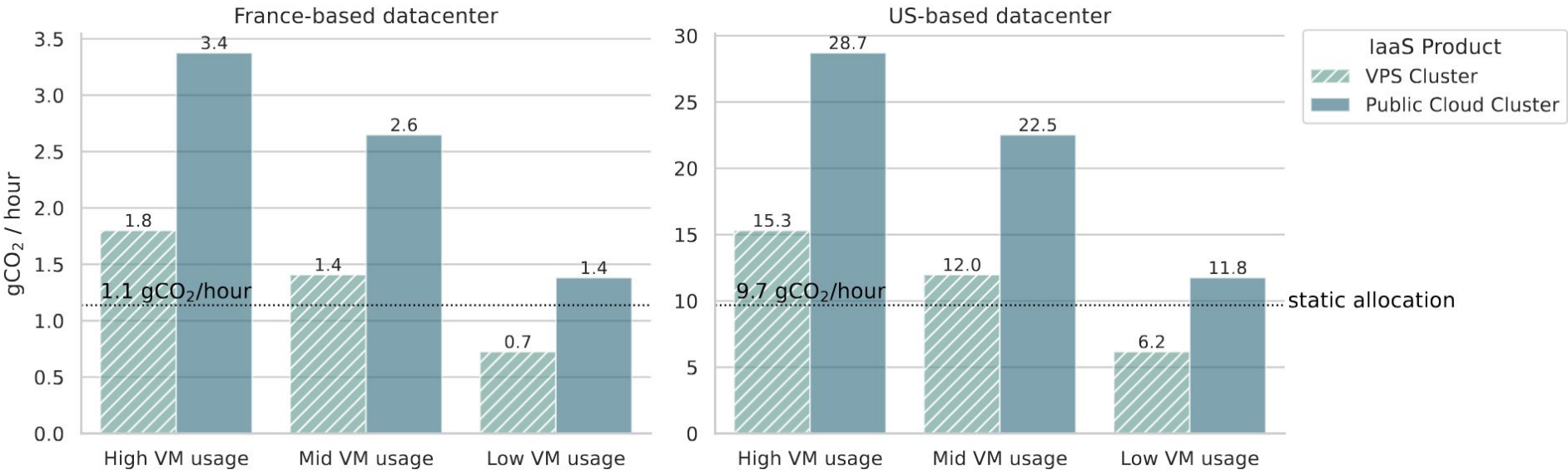
- We propose a methodology to model the isolated VM power consumption (**top-down**) while ponderating it from average runtime conditions (**bottom-up**)



$$\text{CINERGY ratio}_{VM} = \frac{\text{bottom-up model}_{VM}}{\text{top-down model}_{VM}}$$

Improve **Energy efficiency**

- Applications on carbon estimations calculators:
  Current (*static*) estimation overestimate the energy efficiency



Improve **Energy efficiency**

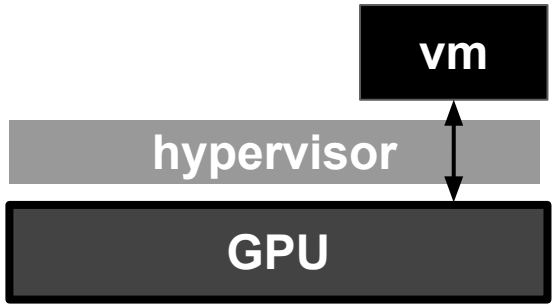
- Take-aways:
  - Runtime guarantees strongly affect cloud instance **power efficiency**
  - **Impacts** of choices and usage can be communicated through a proper **DevOps** framework

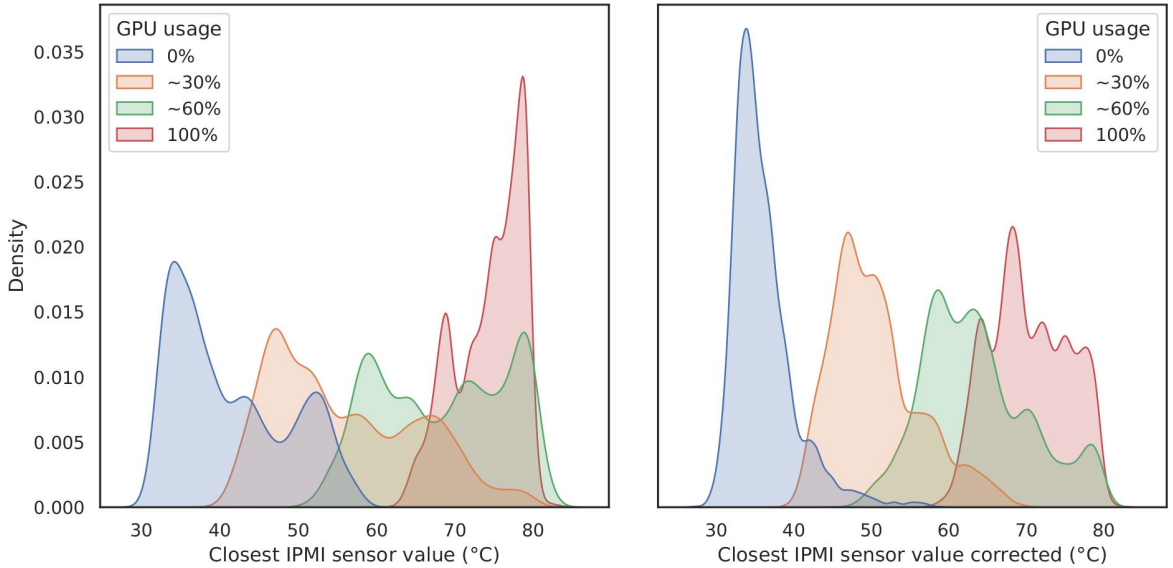
References:
- P. Jacquet et al, CINERGY: Deterministic Power Monitoring for Carbon Accounting in the Cloud, at IEEE Transactions on Cloud Computing, 2026
- P. Jacquet et al, CINERGY: Reasoning over the Worst Case Power Consumption of Cloud Virtual Machines, at IEEE CCGrid (*short paper*), 2025

Limit **GPU expansion**

Are GPUs well-sized on IaaS ? We looked at their temperature

Passthrough

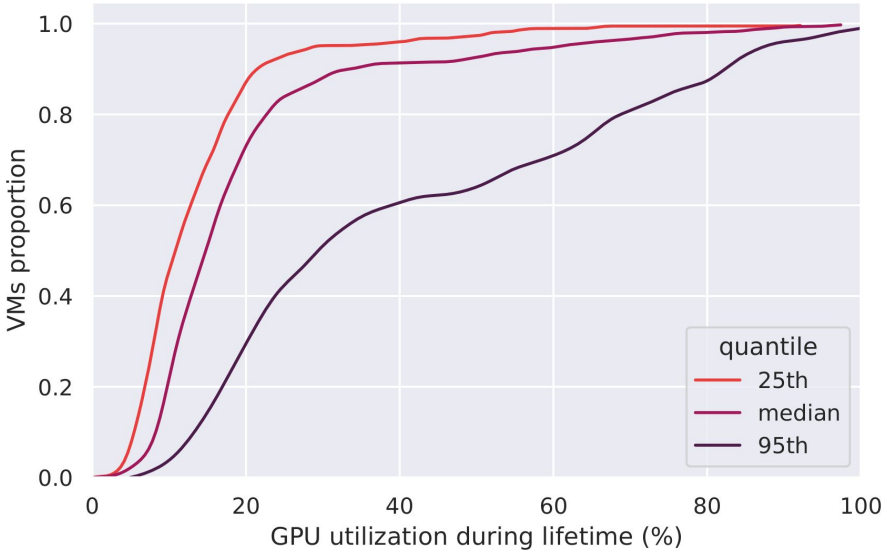


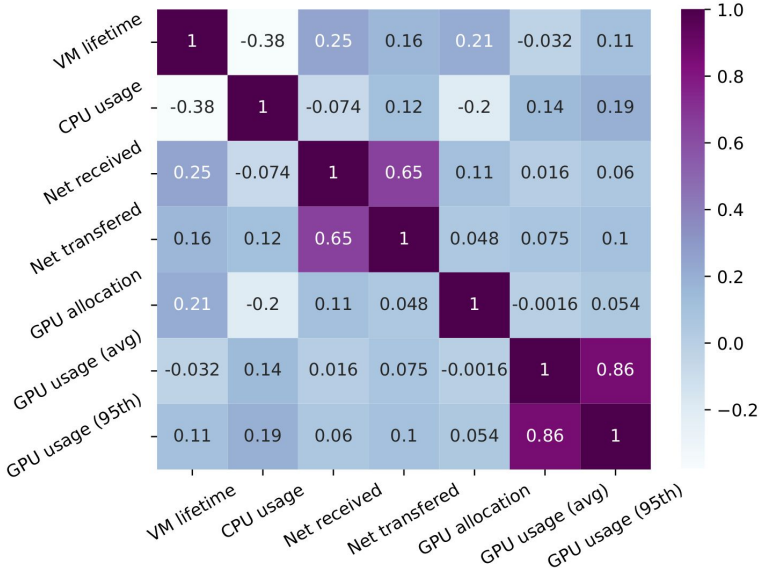


4-A100 servers temperature corrected

Limit **GPU expansion**

We applied our temperature-based monitoring to an IaaS H100 platform





GPU resources are largely underused

Poor correlation with VM features

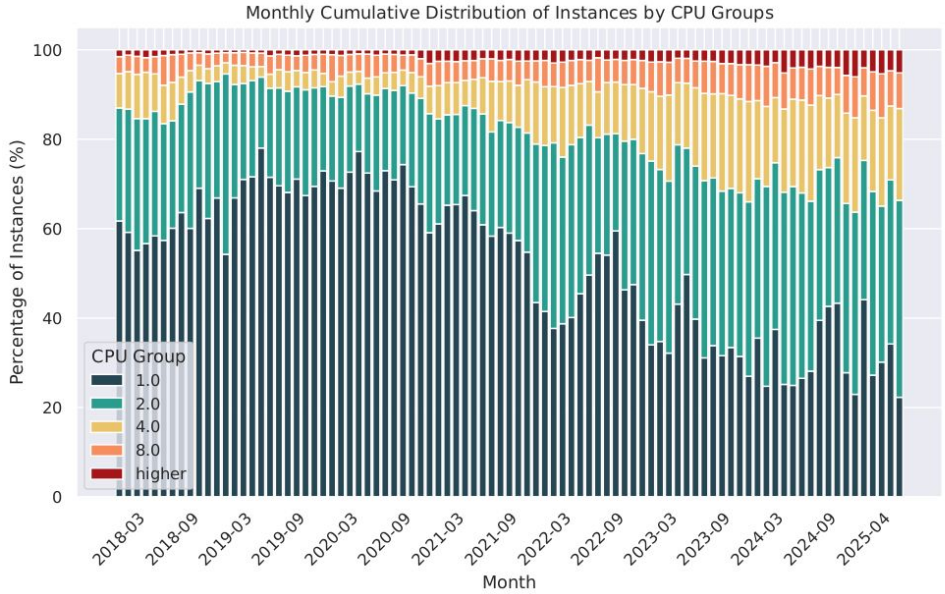
Limit **GPU expansion**

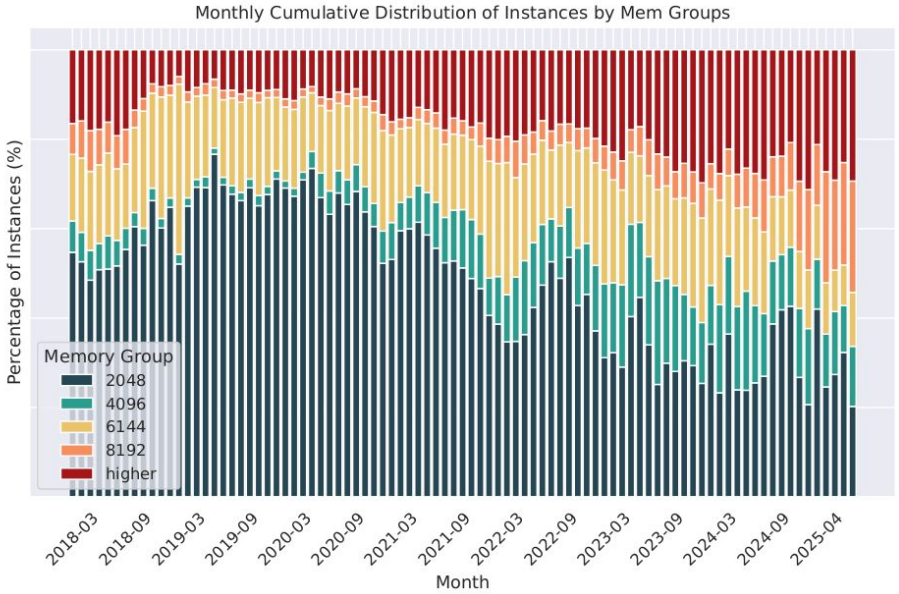
- Take-aways:
  - Inferred GPU compute **usage is far below capacity**, questioning the relevance of the chosen accelerator type
  - **DevOps** could help:
    - Feedback based on black-box monitoring
    - Benchmark-as-a-Service

Reference:
- P. Jacquet et al, Untangling GPU Power Consumption: Job-Level Inference in Cloud Shared Settings, at EuroSys, 2026

VM instances evolution

Cloud workloads consolidation is key to limiting impact…

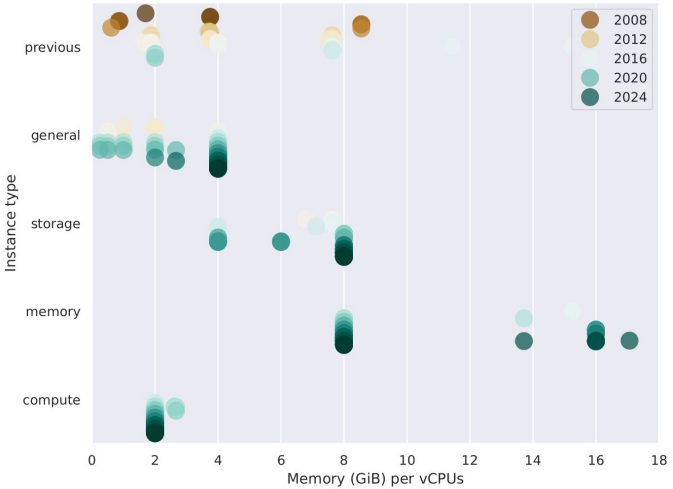


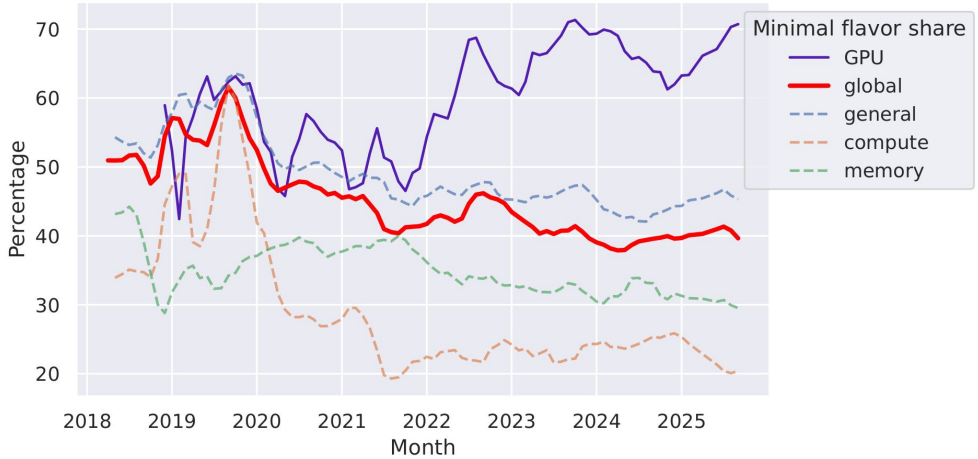


But, the average active instance **doubled in size** from 2019 to 2025

VM instances evolution

Who drives larger VMs: the catalog or user choices?

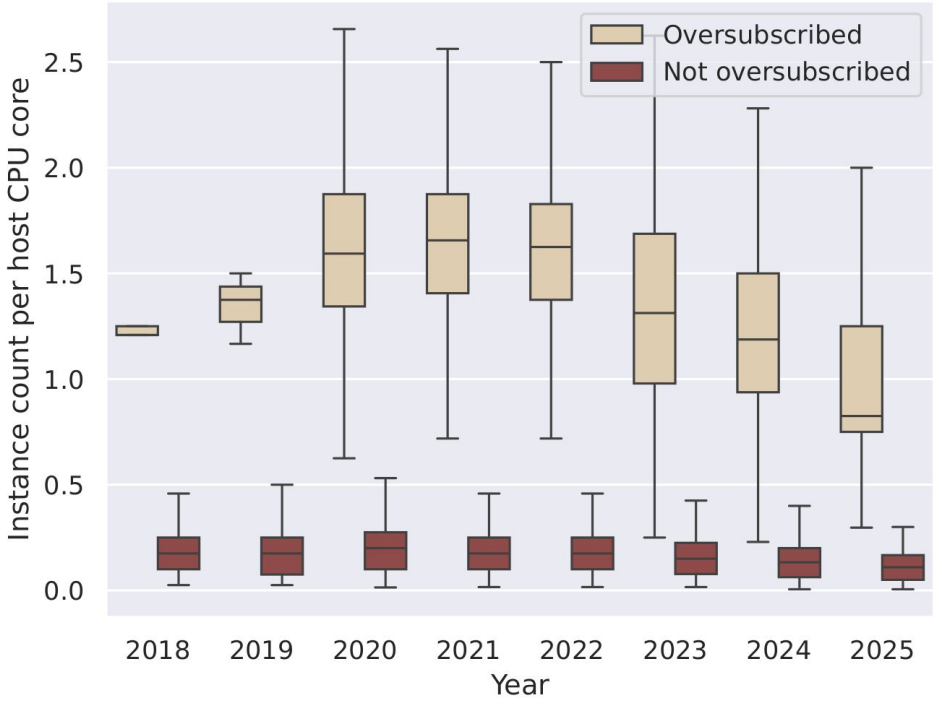




User side: choice of smallest instances is declining

Provider side: small memory options are declining

VM instances evolution



Orchestration impact:
relative packing efficiency declines

Reference:

- P. Jacquet et al, Larger Cloud Servers, Fewer Hosts? On the Evolution of VM Sizes in IaaS Platforms, CCGrid 2026

Take-aways

- Operational **usage feedback** is essential for user/provider cooperation
  - Cloud usage metrics must be embedded in DevOps pipelines
  - Providers must expose more energy/impact metrics!

- **DevOps philosophy** applies to cloud providers too
  - Architectural choices have **measurable adoption consequences**: Providers must close the loop!