Hes·so
Haute Ecole Spécialisée
de Suisse occidentale
Fachhochschule Westschweiz
University of Applied Sciences and Arts
Western Switzerland

Impact d’une taxe carbone sur le coût total du cycle de vie de serveurs

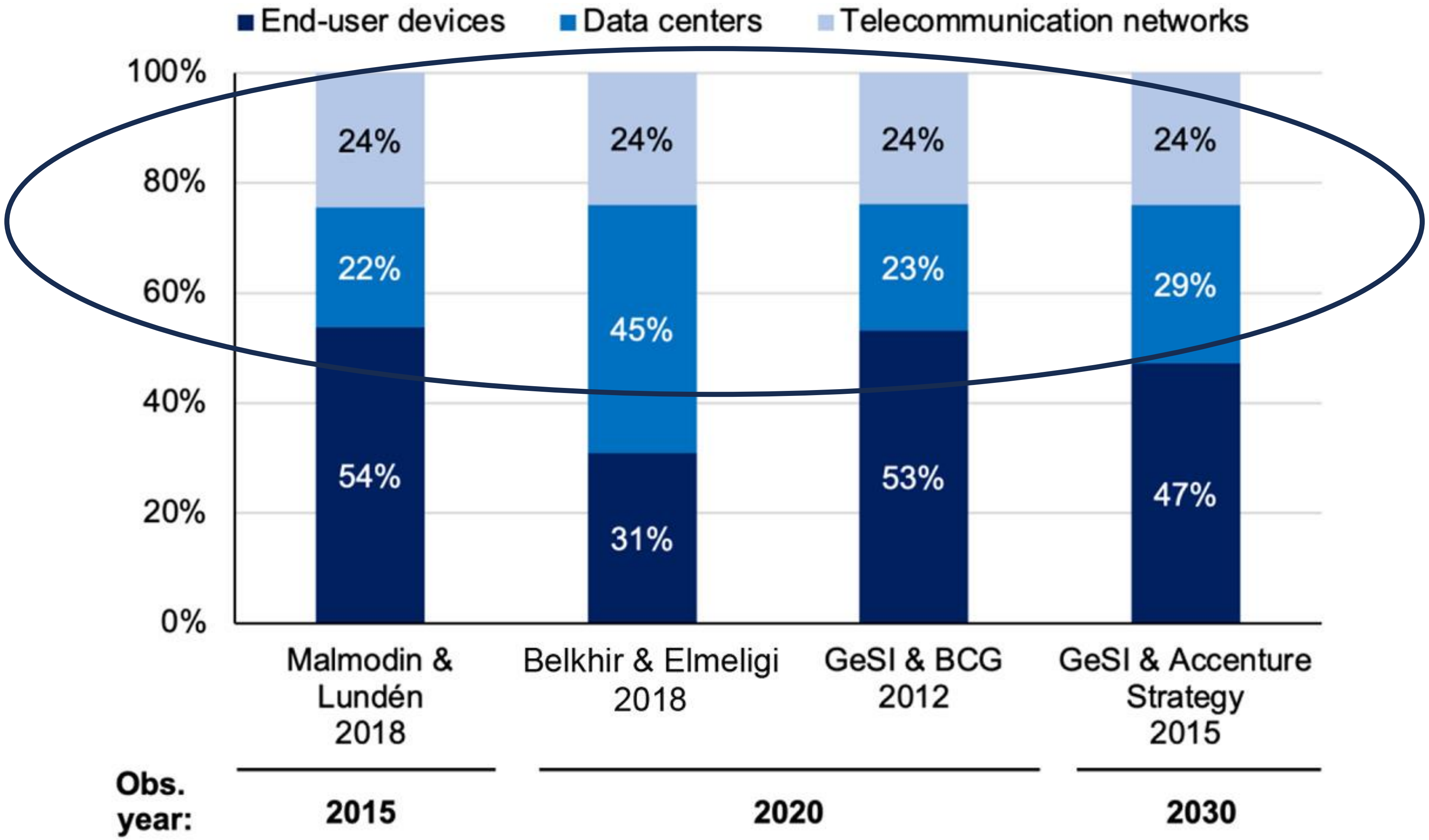
Haute école d’ingénierie et d’architecture Fribourg
Hochschule für Technik und Architektur Freiburg

esos
Électronique Soutenable
Ouverte & Souveraine

Loïc Guibert, Olivier Weppe, **Sébastien Rumley**

INSA
INSTITUT NATIONAL
DES SCIENCES
APPLIQUÉES
RENNES

Préambule



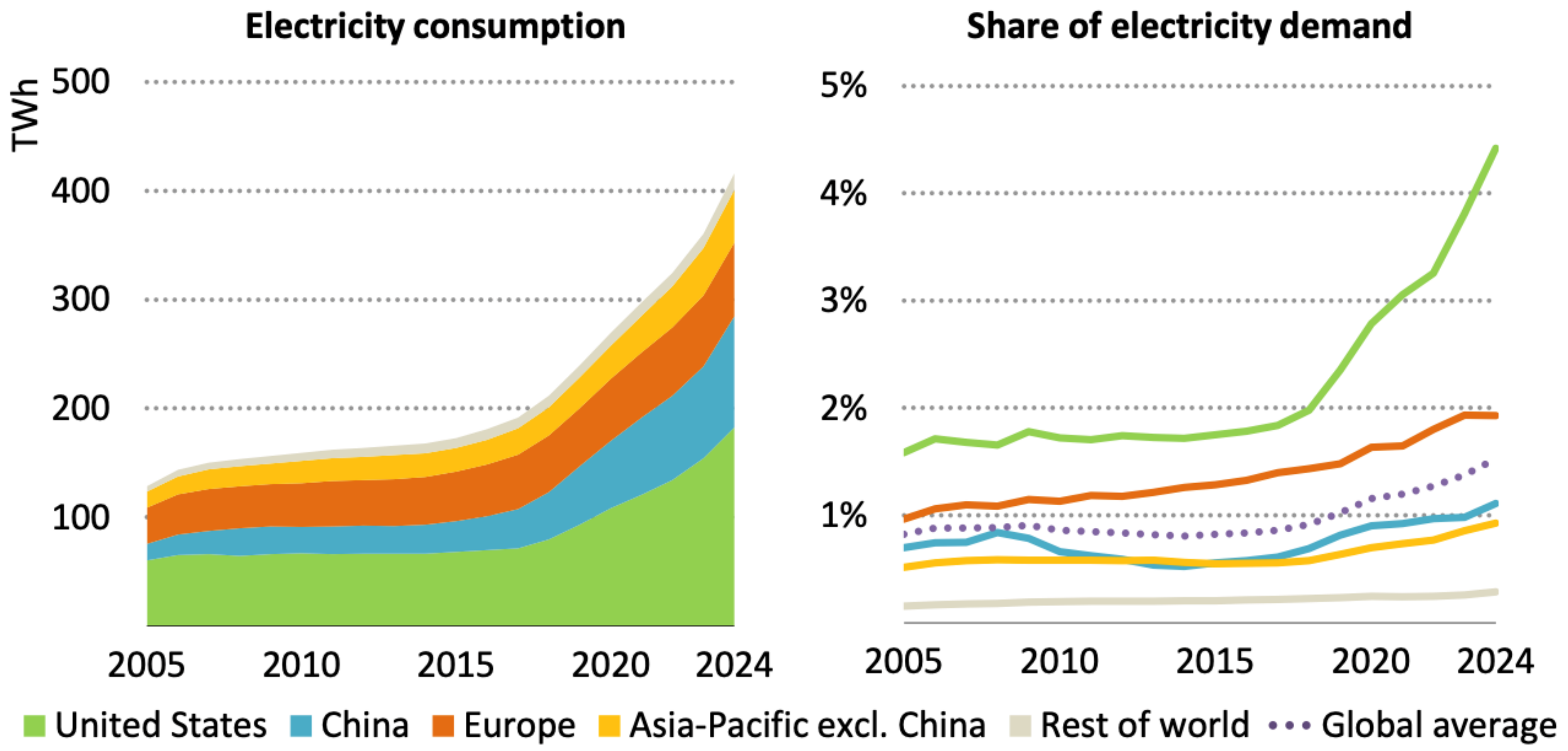
Le cloud…

Répartition des d'émission de GES liées au secteur des technologies de l'information

Jan C.T. Bieser, Ralph Hintemann, Lorenz M. Hilty, Severin Beucker, A review of assessments of the greenhouse gas footprint and abatement potential of information and communication technology, Environmental Impact Assessment Review, Volume 99, 2023,

Préambule

Figure 2.6 ▹ Electricity consumption of data centres by region, 2005-2024



The acceleration in data centre electricity consumption observed in 2017 was mainly driven by the United States and, to a lesser extent, by China

Comment réduire les impacts du “cloud”?

- Moins de tout → oui.. mais c’est une question sociologique
  - Je suis avant tout informaticien…

Il y a aussi et surtout un besoin pour des mesures “primaires”, pour des expériences “bottom-up”

- Les études haut-niveau (top-down) doivent s’intégrer, s’appuyer sur ces expériences bas-niveau

- Les études haut-niveau négligent trop, à notre avis, ces questions
  - Pour **diminuer** les **impacts écologiques** du cloud, il faut (entre autres)
    1. Comprendre comment ils se rapportent à chaque demande, **à un flux de demandes**
    2. Comment ils se comparent aux coûts réels

Mesure de l’impact d’un flux de demandes

TRAFFIC: Testbed foR Assessing energy eFFiciency In throughput Computing

Loïc Guibert
HES-SO
Fribourg, Switzerland
loic.guibert@hefr.ch

Sébastien Reynaud
HES-SO
Fribourg, Switzerland
sebastienreynaud@proton.me

Olivier Weppe
INSA Rennes
Rennes, France
olivier.weppe@insa-rennes.fr

Kristjon Ciko
University of Oslo
Oslo, Norway
kristjoc@ifi.uio.no

Michael Welzl
University of Oslo
Oslo, Norway
michawe@ifi.uio.no

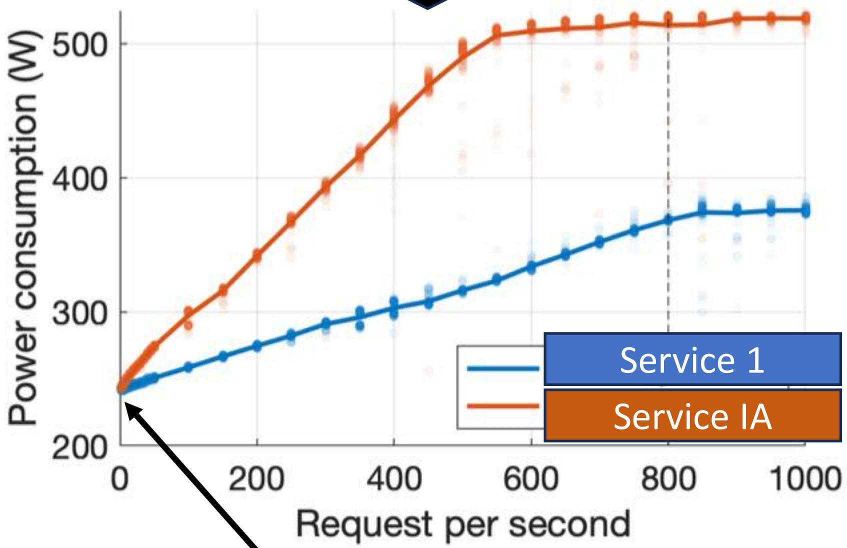
Sébastien Rumley
HES-SO
Fribourg, Switzerland
sebastien.rumley@hefr.ch

Objectif (naïf) 1 : mesurer l’impact énergétique

1 requête à un service “BERT” (ancêtre des LLMs) = environ 1-4 Joules (si machine bien chargée)

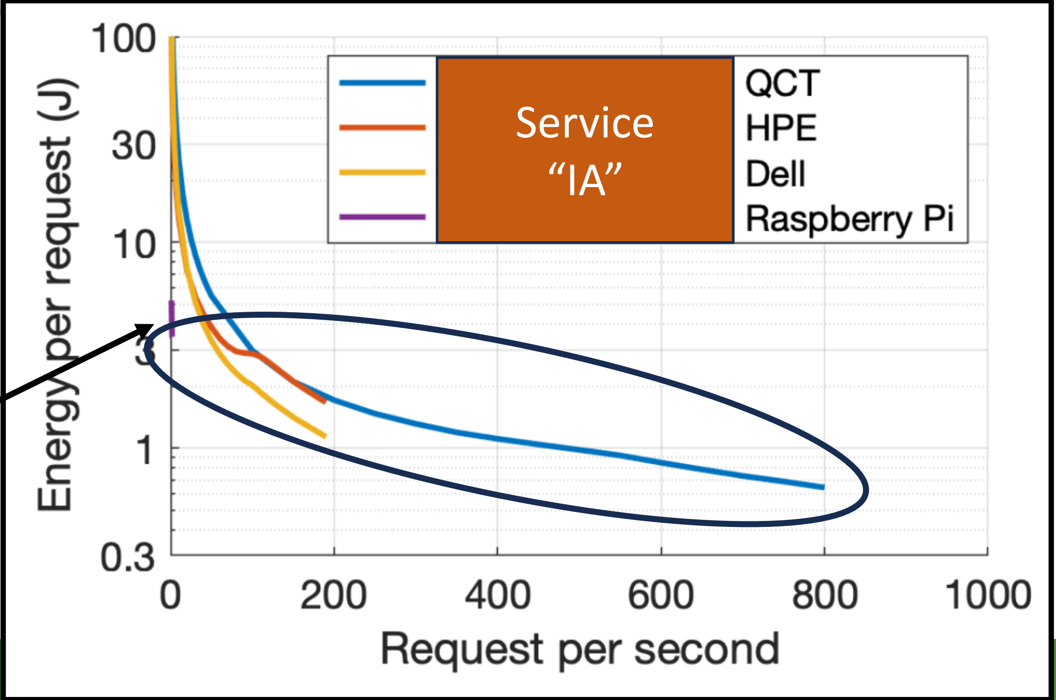
QuantaGrid S74G-2U (QCT)	PowerEdge T430 (Dell)	ProLiant DL360 Gen9 (HPE)	Raspberry Pi 4B
2024	2015	2015	2019
ARM64	x86_64	x86_64	ARM64

QCT

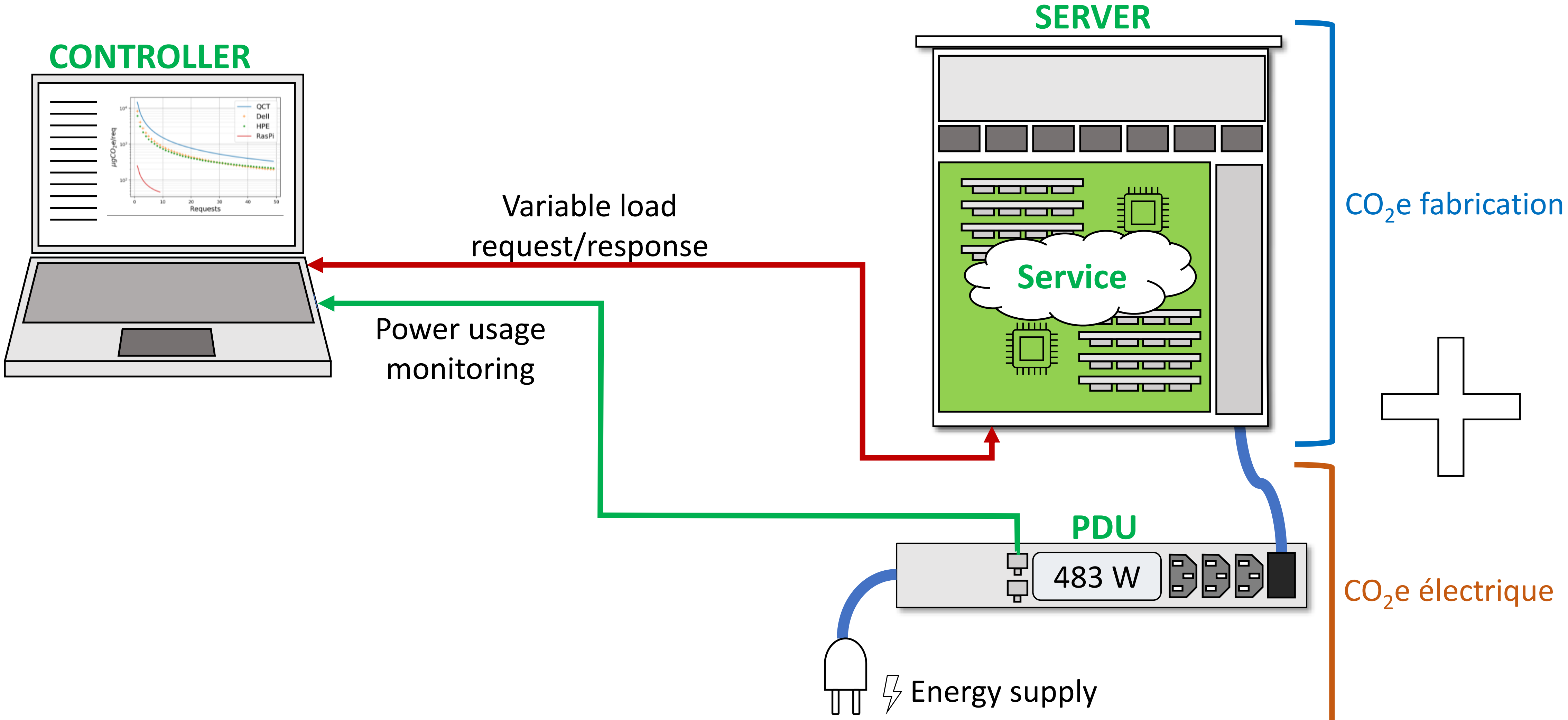


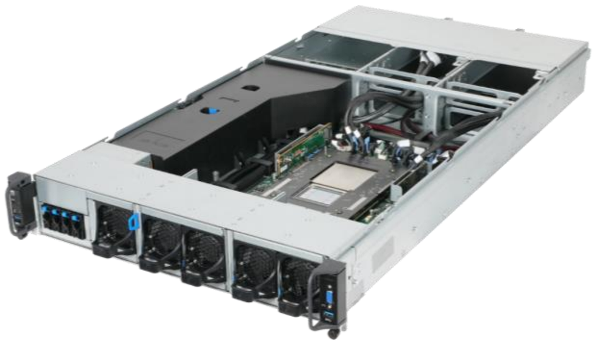
Puissance idle : 240W

Le Raspberry PI est moins efficace à pleine charge (mais efficace pour de faible charges)

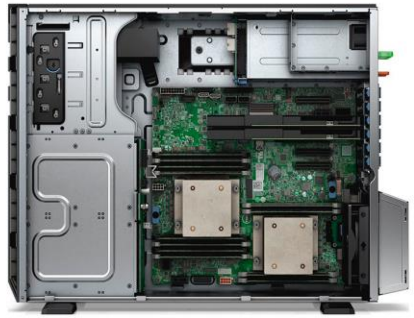


Objectif 2 : mesurer l'impact CO_2e





QuantaGrid S74G-2U (QCT)
2024
ARM64
1023 $kgCO_2e$



PowerEdge T430 (Dell)
2015
x86_64
514 $kgCO_2e$

ProLiant DL360 Gen9 (HPE)
2015
x86_64
345 $kgCO_2e$

Raspberry Pi 4B
2019
ARM64
14 $kgCO_2e$

Reste à additionner le CO_2e correspondant à la consommation électrique

- A une certaine charge
- Pendant 5 ans

Exemple :

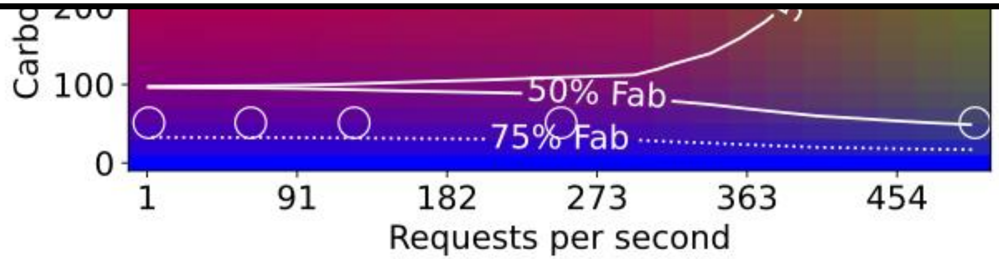
- QCT, 300 requêtes AI/s = 400W
- 400W * 5 ans = 17’520 kWh
- 17’520 kWh * 0.1 $kgCO_2e/kWh$ = 1’752 $kgCO_2e$

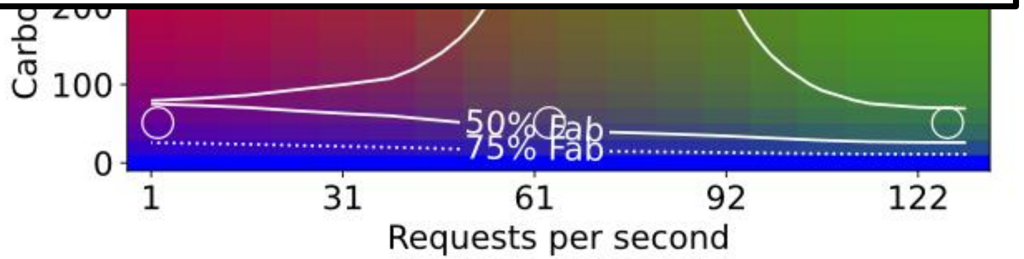
Qu'est-ce qui domine le CO_2e ? Ca dépe

HotCarbon 2025

Carbon Topography Representation: Improving Impacts of Data Center Lifecycle

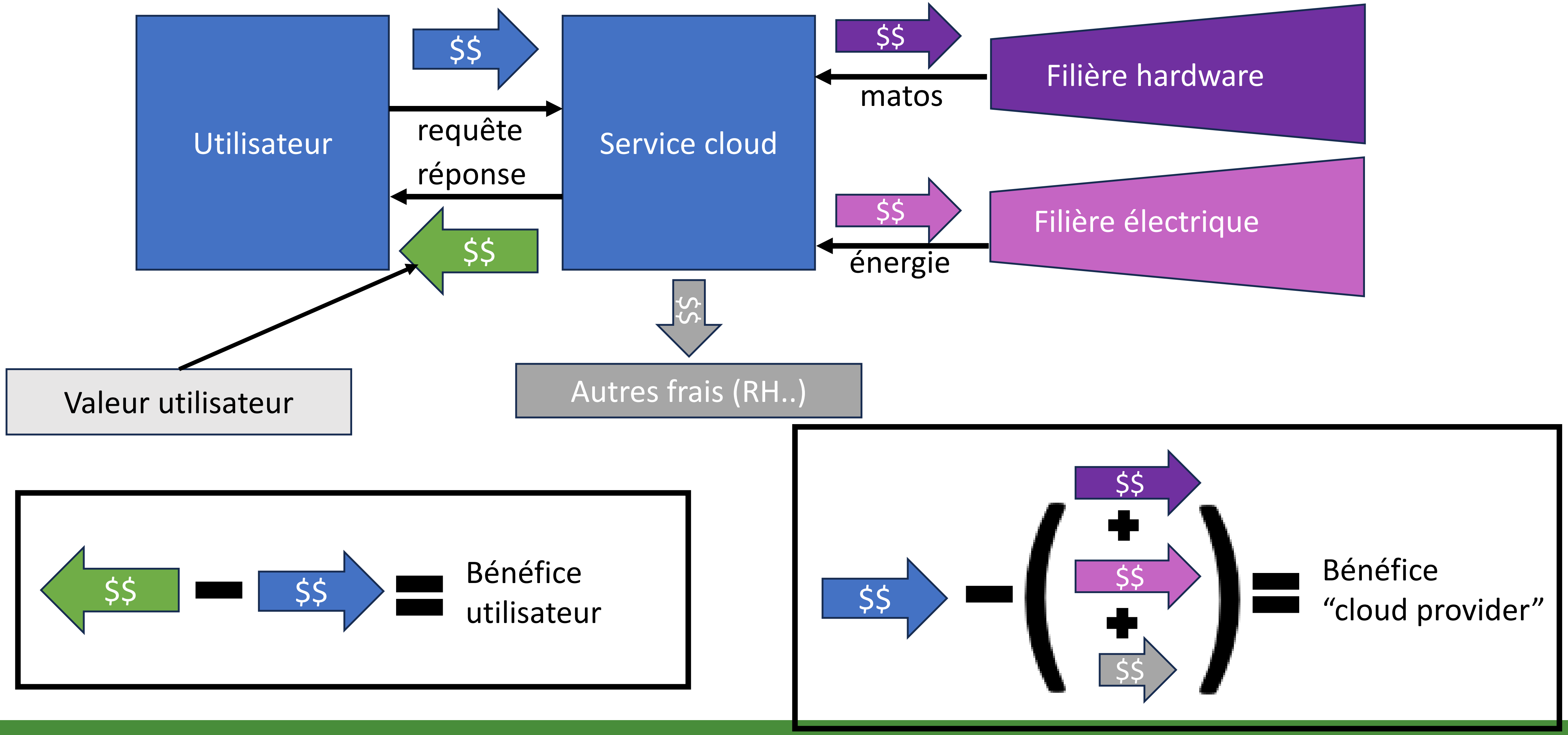
OLIVIER WEPPE, Univ. Rennes, INSA Rennes, CNRS, IETR - UMR 6164, France
DAVID BEKRI, HES-SO, Switzerland
THIBAUT MARTY, Univ. Rennes, INSA Rennes, CNRS, IETR - UMR 6164, France
LOÏC GUIBERT, HES-SO, Switzerland
LOUISE AUBET, Resilio, Switzerland
JEAN-CHRISTOPHE PRÉVOTET, Univ. Rennes, INSA Rennes, CNRS, IETR - UMR 6164, France
MAXIME PELCAT, Univ. Rennes, INSA Rennes, CNRS, IETR - UMR 6164, France
SÉBASTIEN RUMLEY, HES-SO, Switzerland



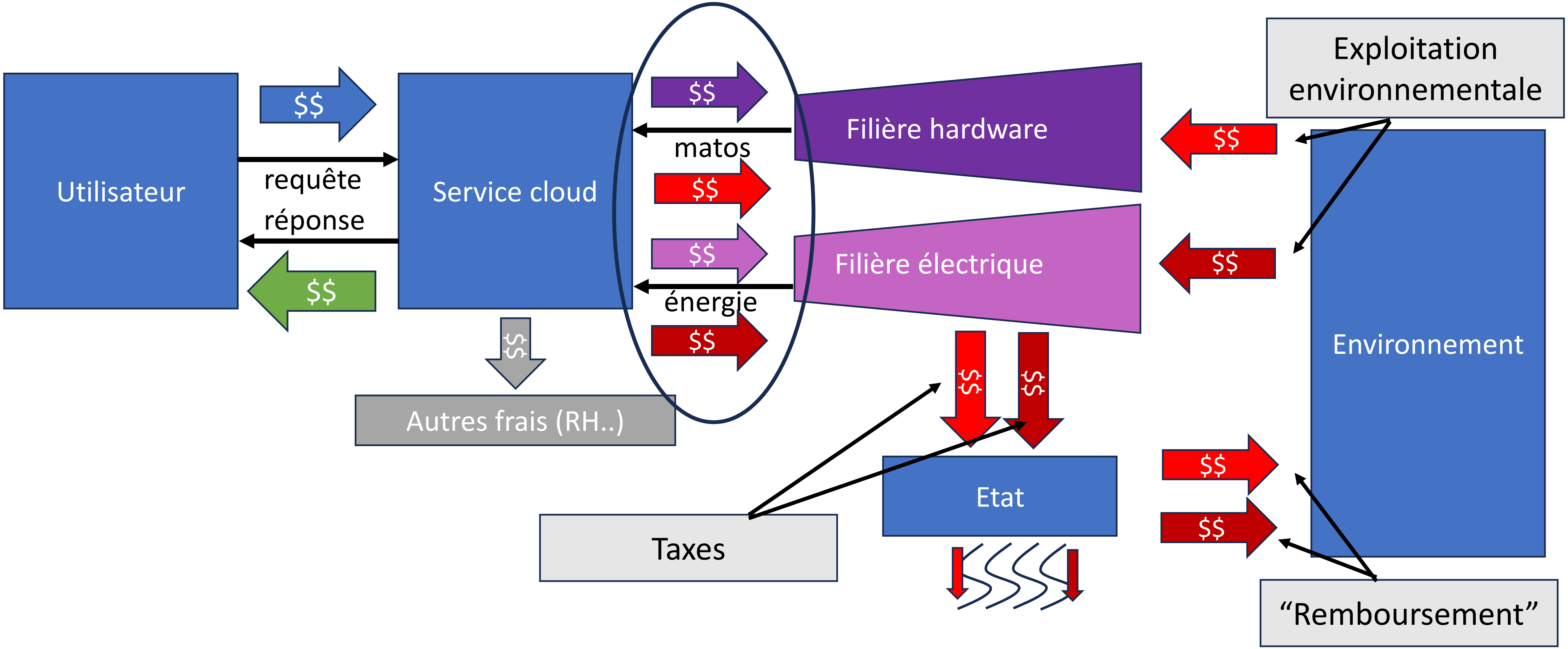


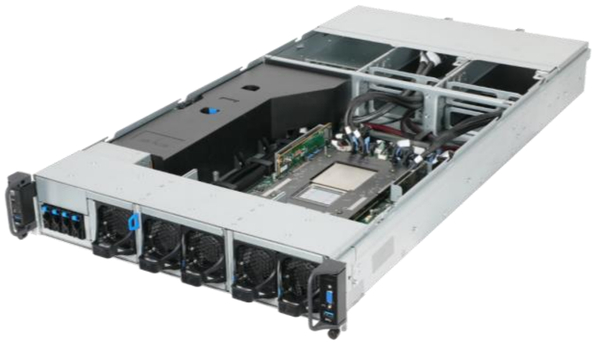
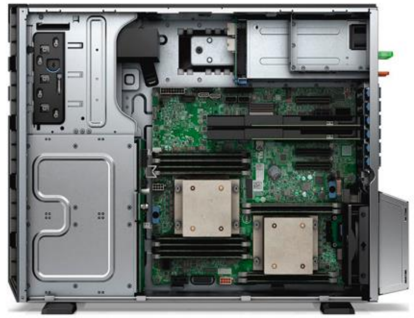
Dans des pays comme la France (<100gCO_2e/kWh), la fabrication tend à dominer les impacts

Coûts écologiques comparés aux coûts réels

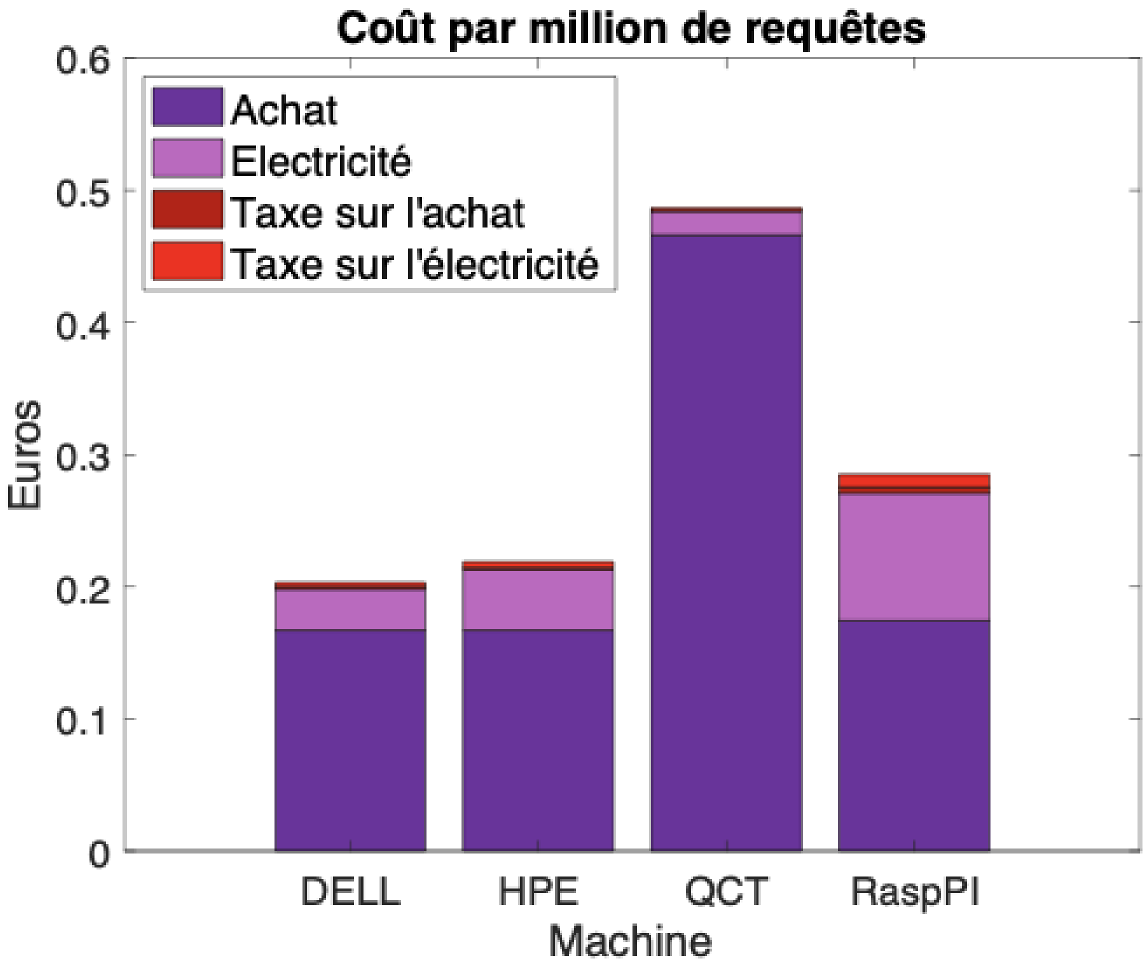


Coûts écologiques comparés aux coûts réels

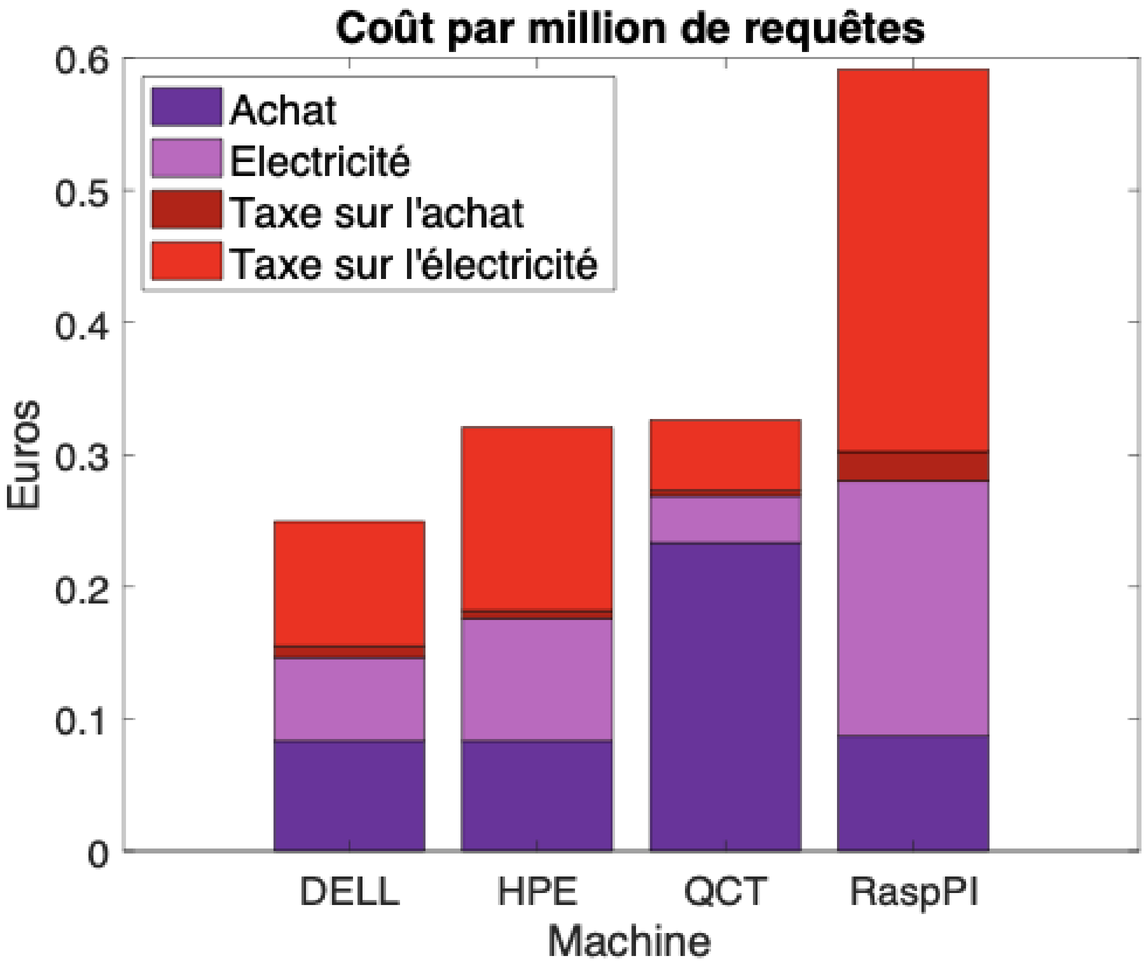


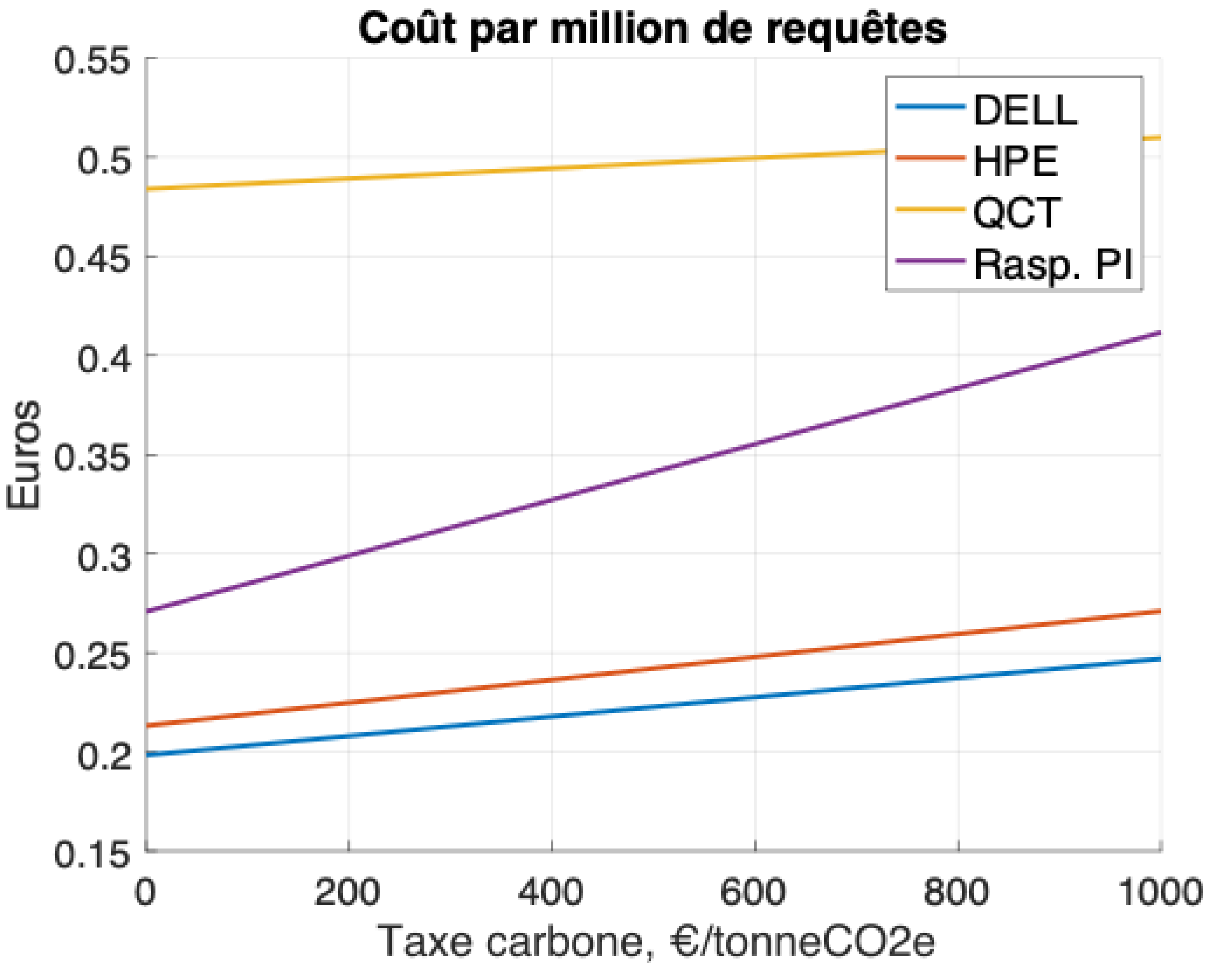
QuantaGrid S74G-2U (QCT)	PowerEdge T430 (Dell)	ProLiant DL360 Gen9 (HPE)	Raspberry Pi 4B
2024	2015	2015	2019
ARM64	x86_64	x86_64	ARM64
1023 kgCO_2e	**514 kgCO_2e**	**345 kgCO_2e**	**14 kgCO_2e**
58’800 €	**5’000 €**	**5’000 €**	**55 €**
Puissance à charge max.			
514 Watt	**215 Watts**	**317 Watts**	**7 Watts**
Flux max. supporté			
800 req/s	**190 req/s**	**190 req/s**	**2 req/s**



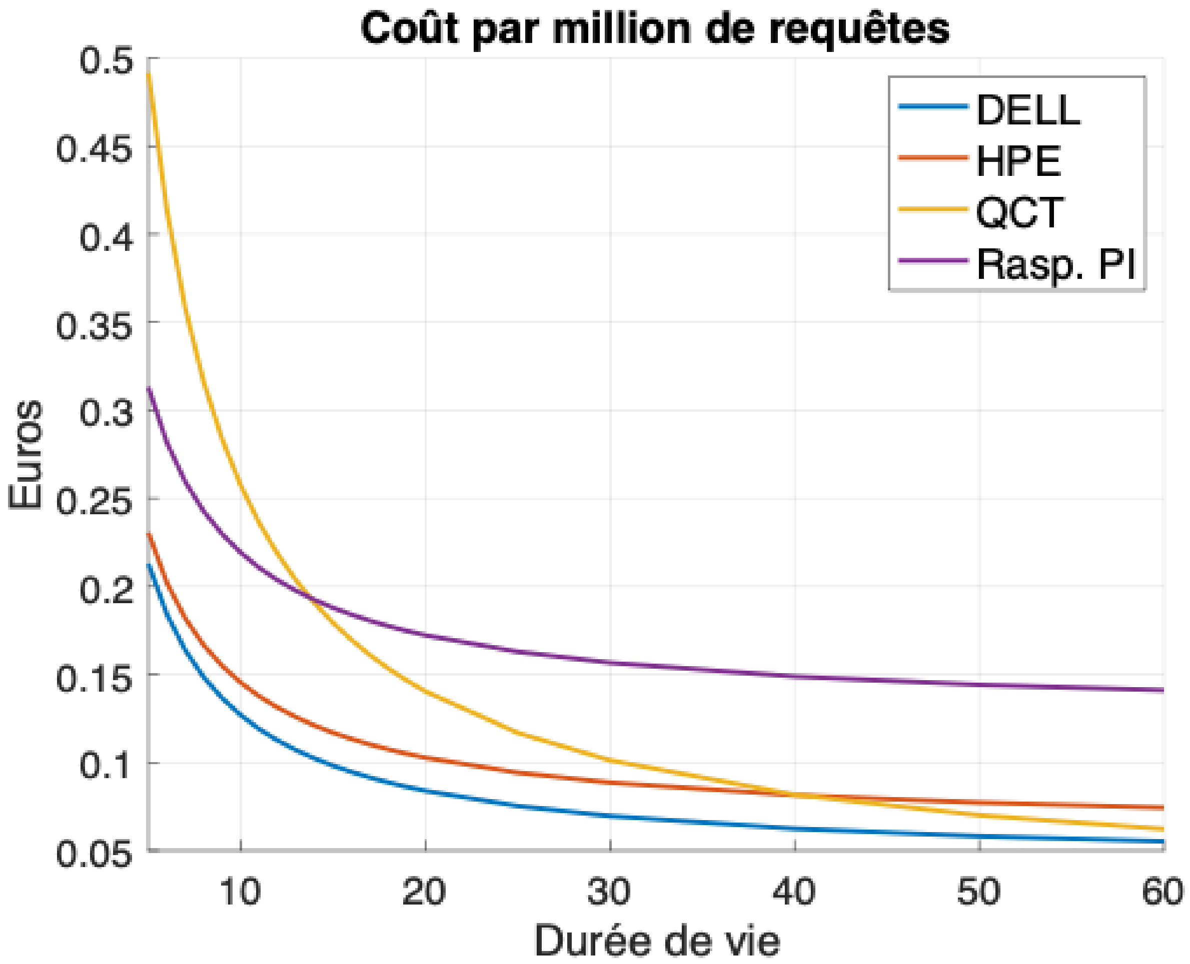
Durée de vie : 5 ans
Intensité carbone: 100gCO_2e/kWh
Prix du kWh : 0.1 €
Taxe carbone : 100 €/tonne



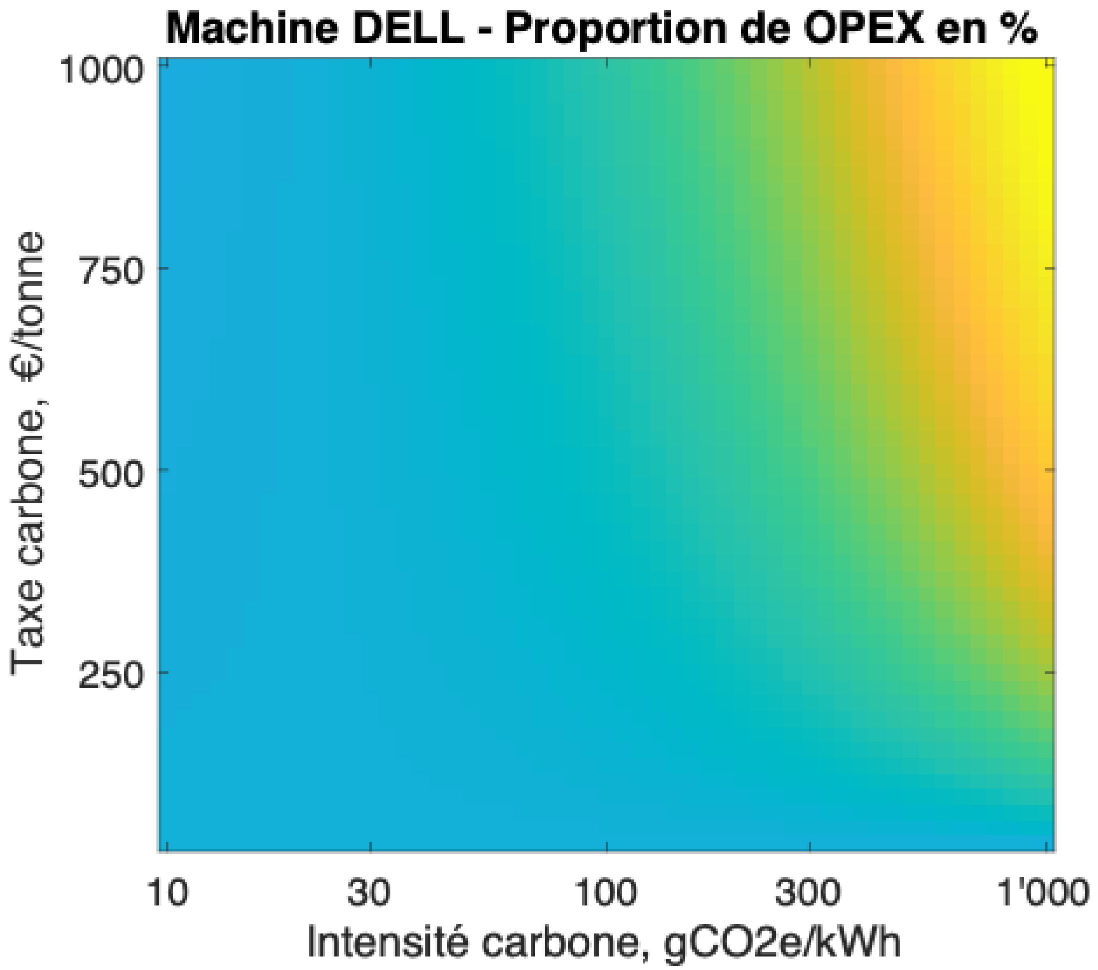
Durée de vie : **10 ans**
Intensité carbone: **300 gCO_2e/kWh**
Prix du kWh : **0.2 €**
Taxe carbone : **1000 €/tonne**

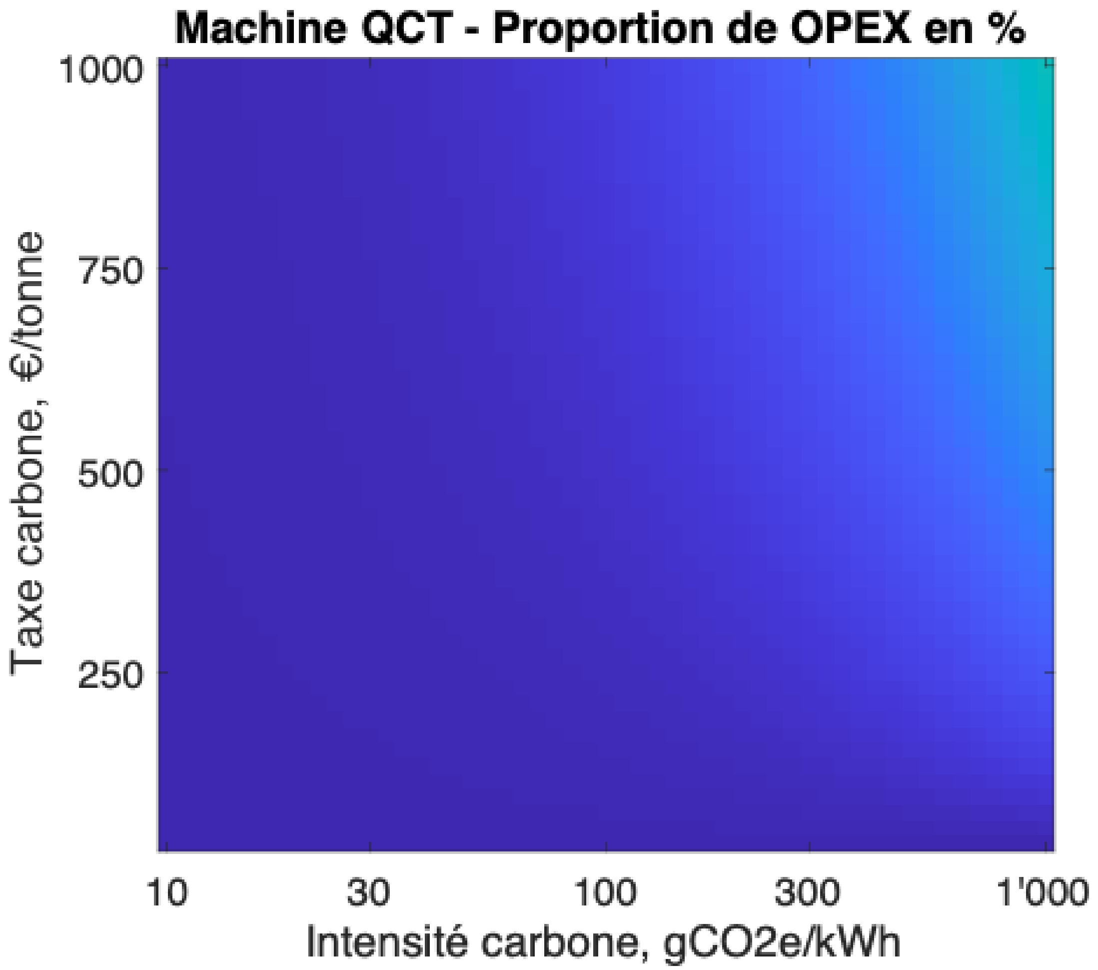


Durée de vie : 5 ans
Intensité carbone: 100 gCO_2e/kWh
Prix du kWh : 0.1 €



Intensité carbone: 100 gCO_2e/kWh
Prix du kWh : 0.1 €
Taxe carbone : 300 €/tonne





Durée de vie : **8 ans**
Prix du kWh : **0.2 €**

Conclusions

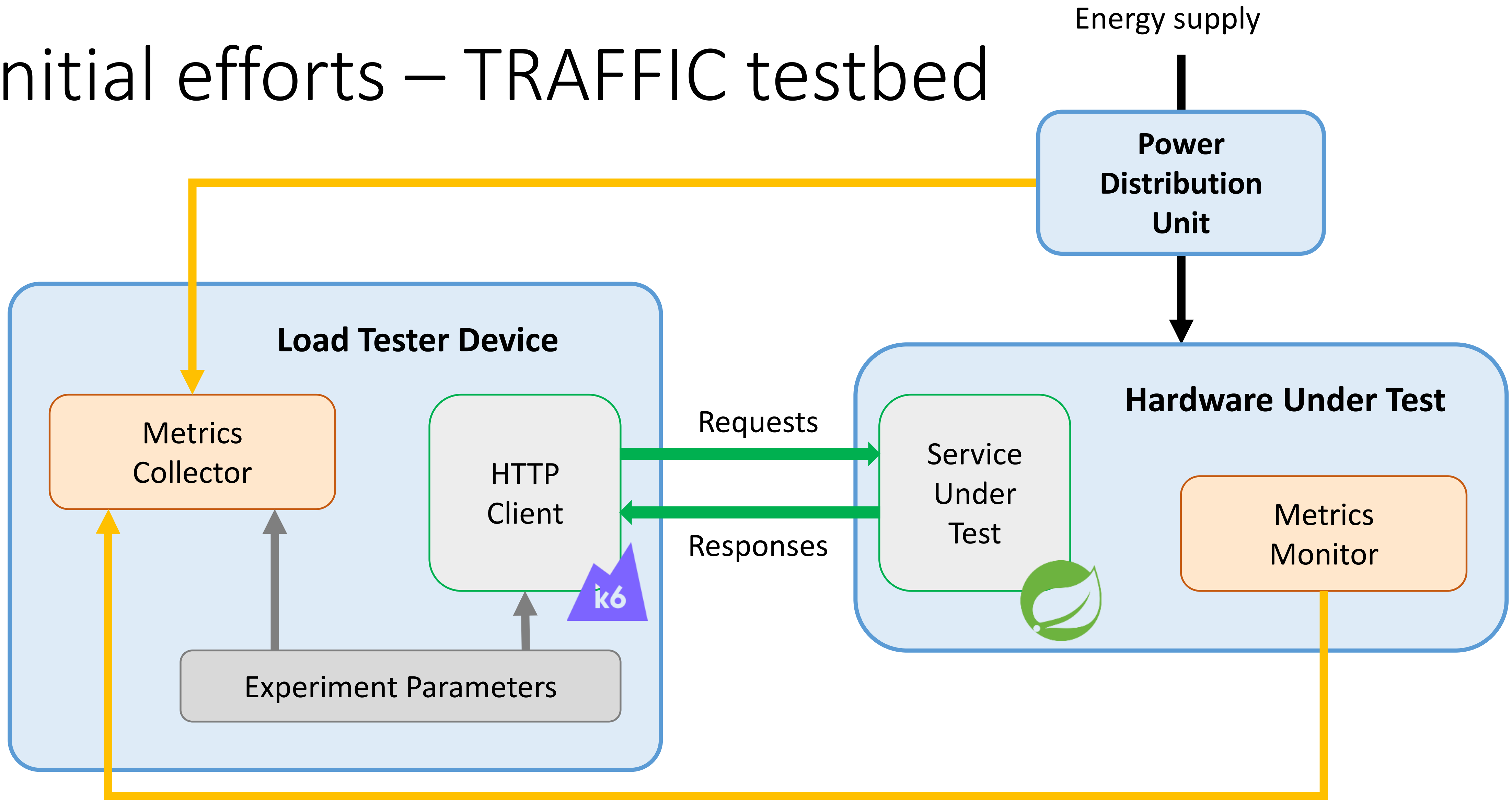
- En Joule/requête, la machine récente à base de ARM est la plus efficace
  - … pour autant que la charge soit suffisante
- En CO_2/requête, la machine récente est aussi la plus efficace
  - … mais là encore, pour autant que le demande soit suffisante
- En €/requête, la machine DELL donne les meilleurs résultats
  - Moindre consommation que la machine HPE
  - Le prix élevé de la machine récente ne peut pas être amorti

Lié à la non-utilisation du GPU

- Seules des taxes CO_2 très hautes affectent le rapport OPEX/CAPEX
- Toute cette analyse néglige les “autres frais”
  - … qui devraient marginaliser un peu plus les coûts écologiques.

Slides additionnelles

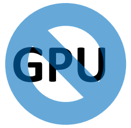
Initial efforts – TRAFFIC testbed

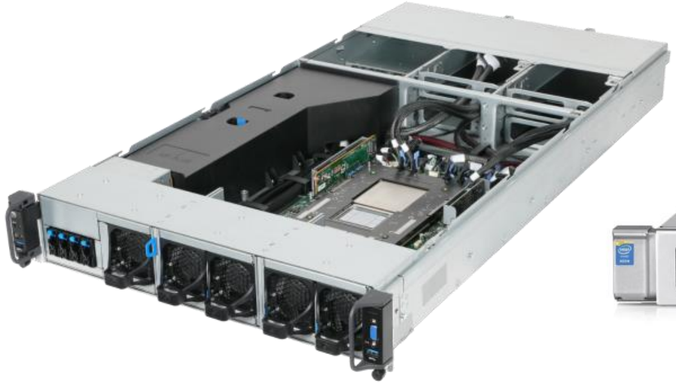


[3] Loïc Guibert, Sébastien Reynaud, Olivier Weppe, Kristjon Ciko, Michael Welzl, and Sébastien Rumley. 2025. TRAFFIC: Testbed foR Assessing energy eFFiciency In throughput Computing. In Proceedings of the 22nd ACM International Conference on Computing Frontiers: Workshops and Special Sessions (CF '25 Companion).

Hardware under test

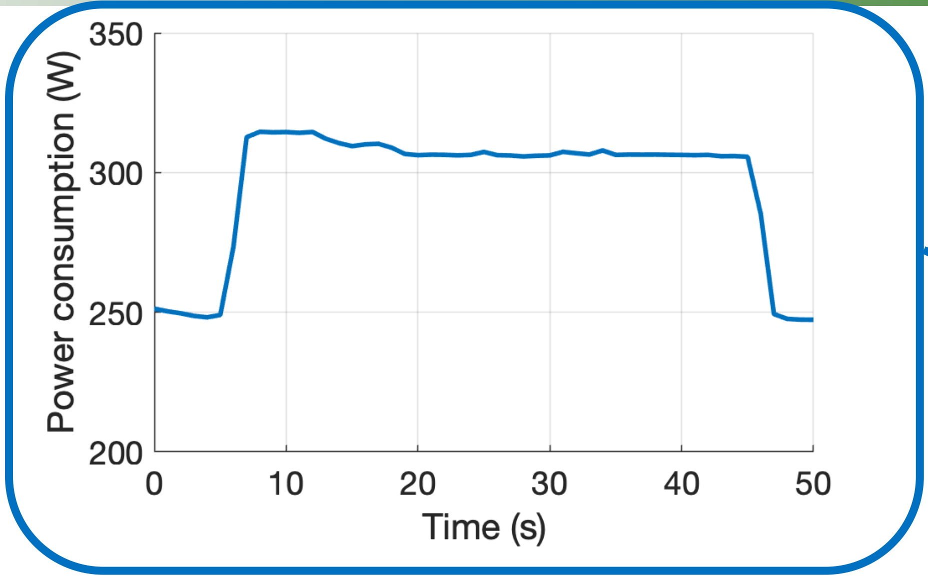
Machine	CPU	Frequency	Cores/Threads	Year
QCT QuantaGrid S74G-2U (sold by **E4**)	Nvidia Grace Hopper Superchip (ARM Neoverse V2)	3.3	72 / 72	2024
HPE ProLiant DL360 Gen9	2 x Intel(R) Xeon(R) E5-2640 v3	2.6	16 / 32	2015
Dell PowerEdge T430	2 x Intel(R) Xeon(R) E5-2620 v4	2.1	16 / 32	2015
Raspberry Pi 4 Model B	Broadcom BCM2711C0 (ARM Cortex-A72)	1.8	4 / 4	2021

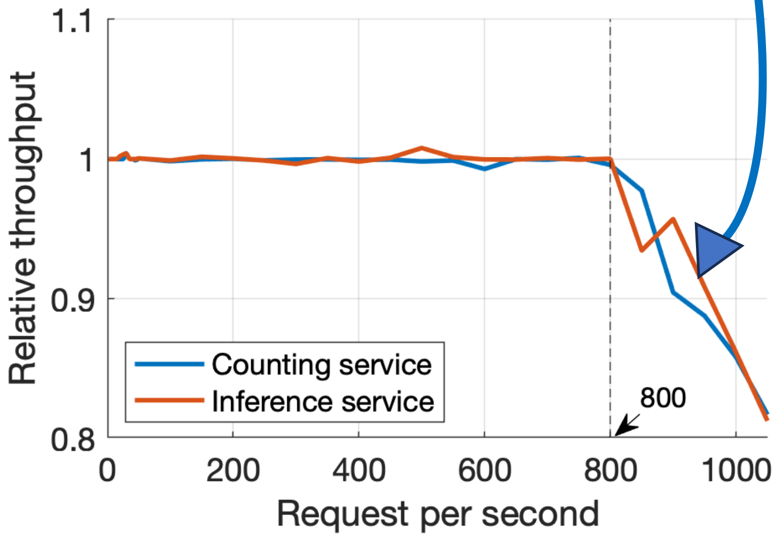




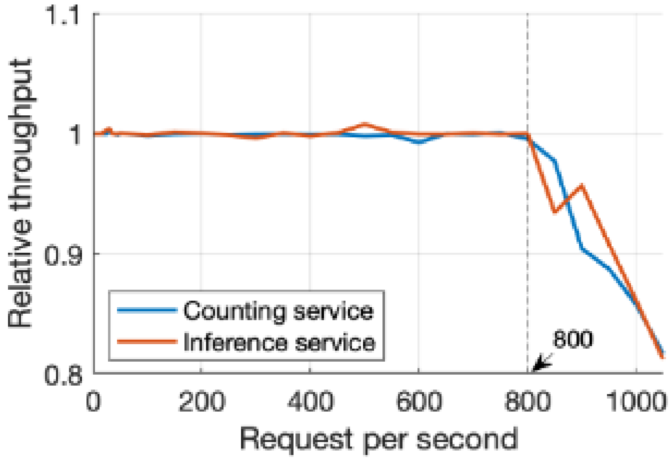
One experiment

- Inputs
  - … one hardware
  - … one service
  - … one parameter to the service
    - For example, count until 280 millions
  - … one load parameter
    - For example, 300 requests per second
- Outputs
  - (relative) throughput
  - Average power
  - CPU utilization
  - CPU frequency
  - …

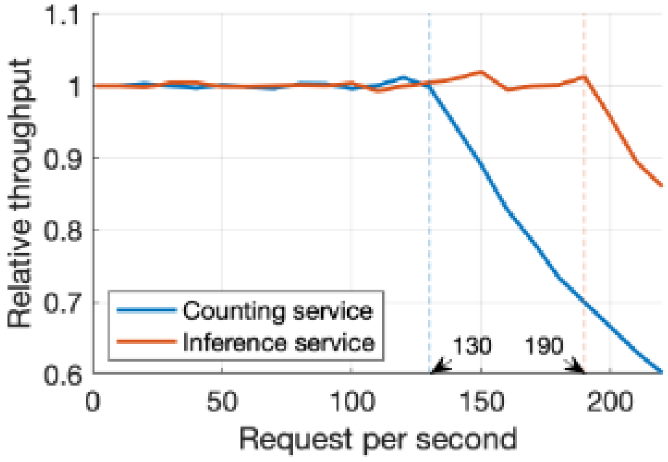




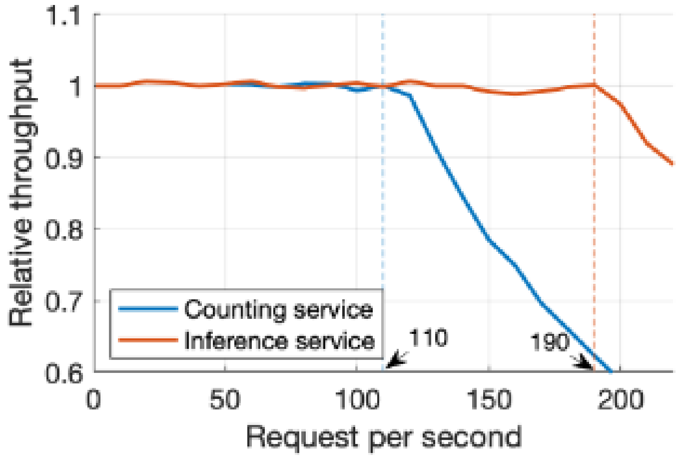
Throughput



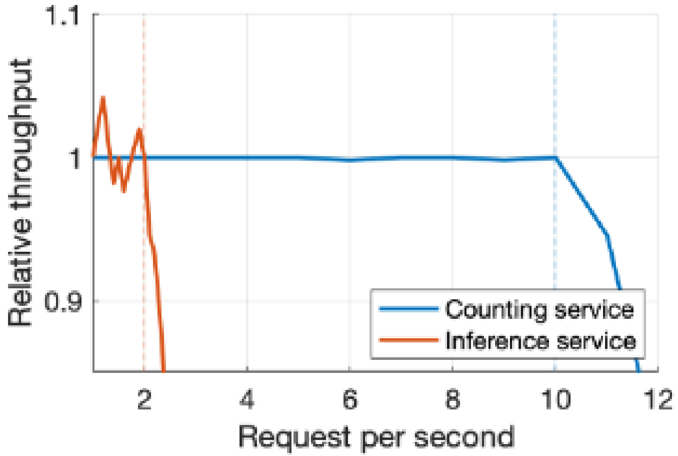
(a) QCT



(b) HPE



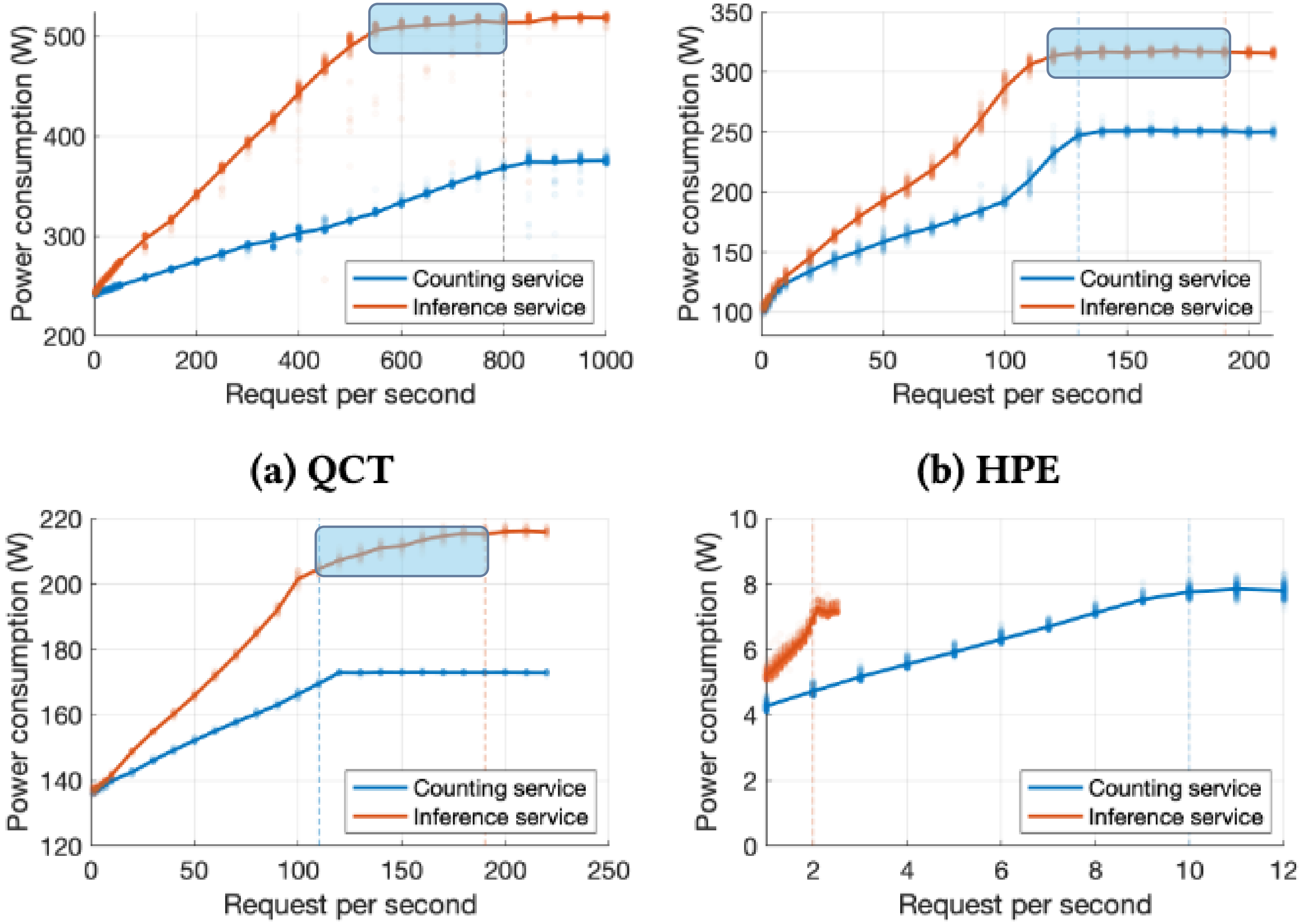
(c) Dell



(d) Raspberry Pi

	Counting	Inference
QCT	800	800
HPE	130	**190**
Dell	110	**190**
RasPi	**10**	2

Power consumption



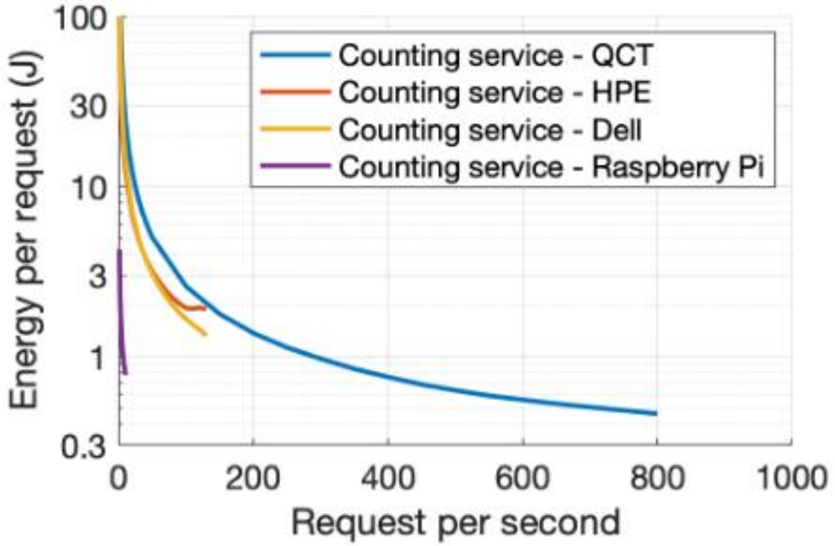
(a) QCT

(b) HPE

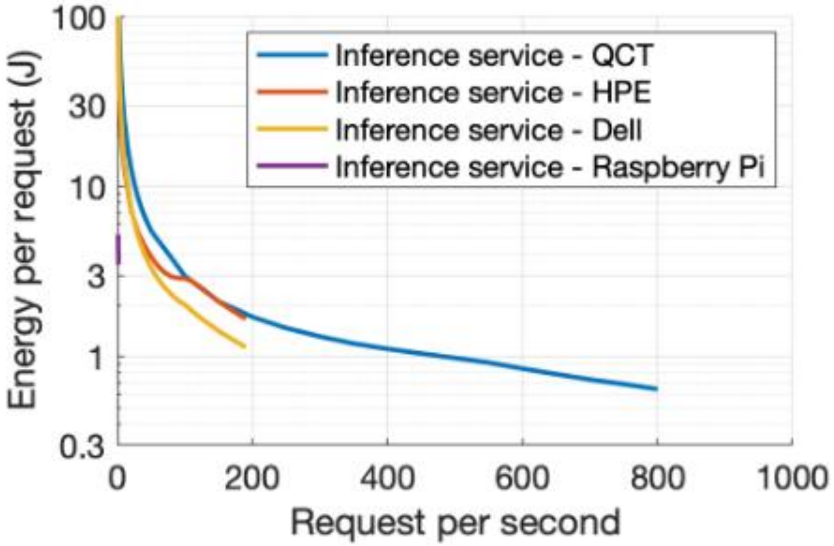
(c) Dell

(d) Raspberry Pi

Energy usage comparisons



(a) Counting service



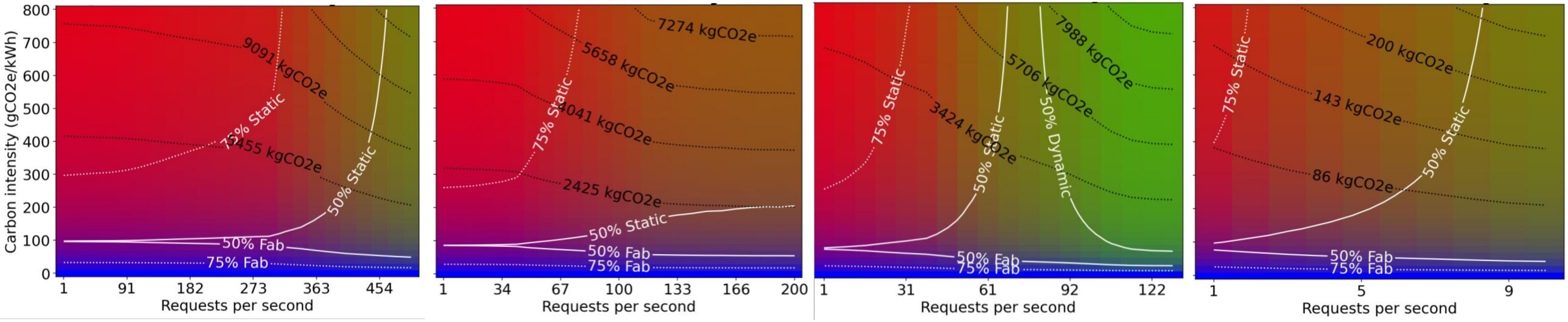
(b) Inference service

@ 2 req/s	Counting	Inference
QCT	115 J	116 J
HPE	100 J	101 j
Dell	103 J	103 J
RasPi	**5 J**	**7 J**

Hes·so

Carbon Topography Representation

QuantaGrid S74G-2U (QCT) | PowerEdge T430 (Dell) | ProLiant DL360 Gen9 (HPE) | Raspberry Pi 4B



[4] Olivier Weppe, David Bekri, Thibaut Marty, Loïc Guibert, Louise Aubet, Jean-Christophe Prévotet, Maxime Pelcat, and Sébastien Rumley. 2025. Carbon Topography Representation: Improving Impacts of Data Center Lifecycle. SIGENERGY Energy Inform. Rev. 5, 2 (July 2025), 41–47.