Cloud Computing -- The Carbon Footprint of Datacenters

Thomas Ropars

Email: thomas.ropars@univ-grenoble-alpes.fr

Website: tropars.github.io

Warning

Studying the carbon footprint of Datacenters is complex:

- A rather recent research topic
- It is difficult to collect accurate data
- The data presented in these slides might not be 100% correct

References

The following references were used to prepare these slides:

- Reports from the Shift project: https://theshiftproject.org/en/lean-ict-2/
- L. A. Barroso, U. Hölzle, and P. Ranganathan. The datacenter as a computer: Designing warehouse-scale machines. Synthesis Lectures on Computer Architecture, 13(3):i–189, 2018.
- Other papers are cited directly in the slides

Carbon footprint

Definition

We are trying to evaluate:

- The quantity of Greenhouse Gases (GHGs) emitted into the atmosphere by some activities
  - Direct emissions (made directly during the progress of a process)
  - Embodied emissions (manufacturing, transportation, etc.)

We express it in terms of CO2 equivalent mass (CO2-e):

- There is more than CO2 in Greenhouse Gases (see the 6 Kyoto gasses)
- CO2-e is used to express the global warming potential of all these gasses
  - A common unit that allows making comparisons
  - The most important gas: 70% of the emissions in France

Pandey, Divya, Madhoolika Agrawal, and Jai Shanker Pandey. "Carbon footprint: current methods of estimation." Environmental monitoring and assessment 178 (2011): 135-160.

Impact of datacenters on the environment

Not only about the carbon footprint

- Use of different metals
    - Rare earth elements (soil pollution)
- Soil artificialization

IT Carbon footprint

Energy consumption of the IT domain

A constant growth and an increased percentage in the total energy consumption

- 6.2% of energy increase per year between 2015 and 2019
    - Doubling in 12 years
- More that 5% of the energy consumption due to IT

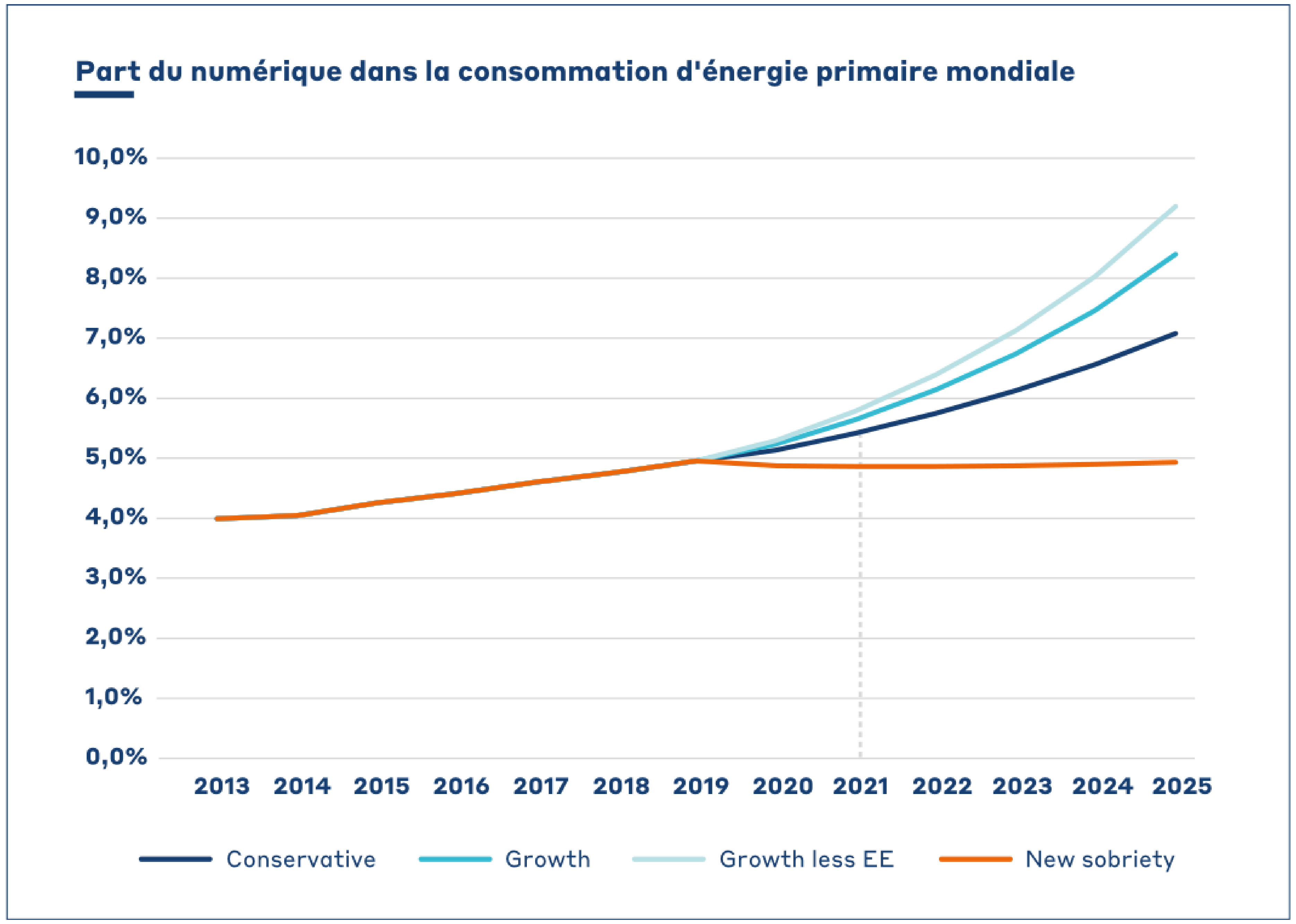


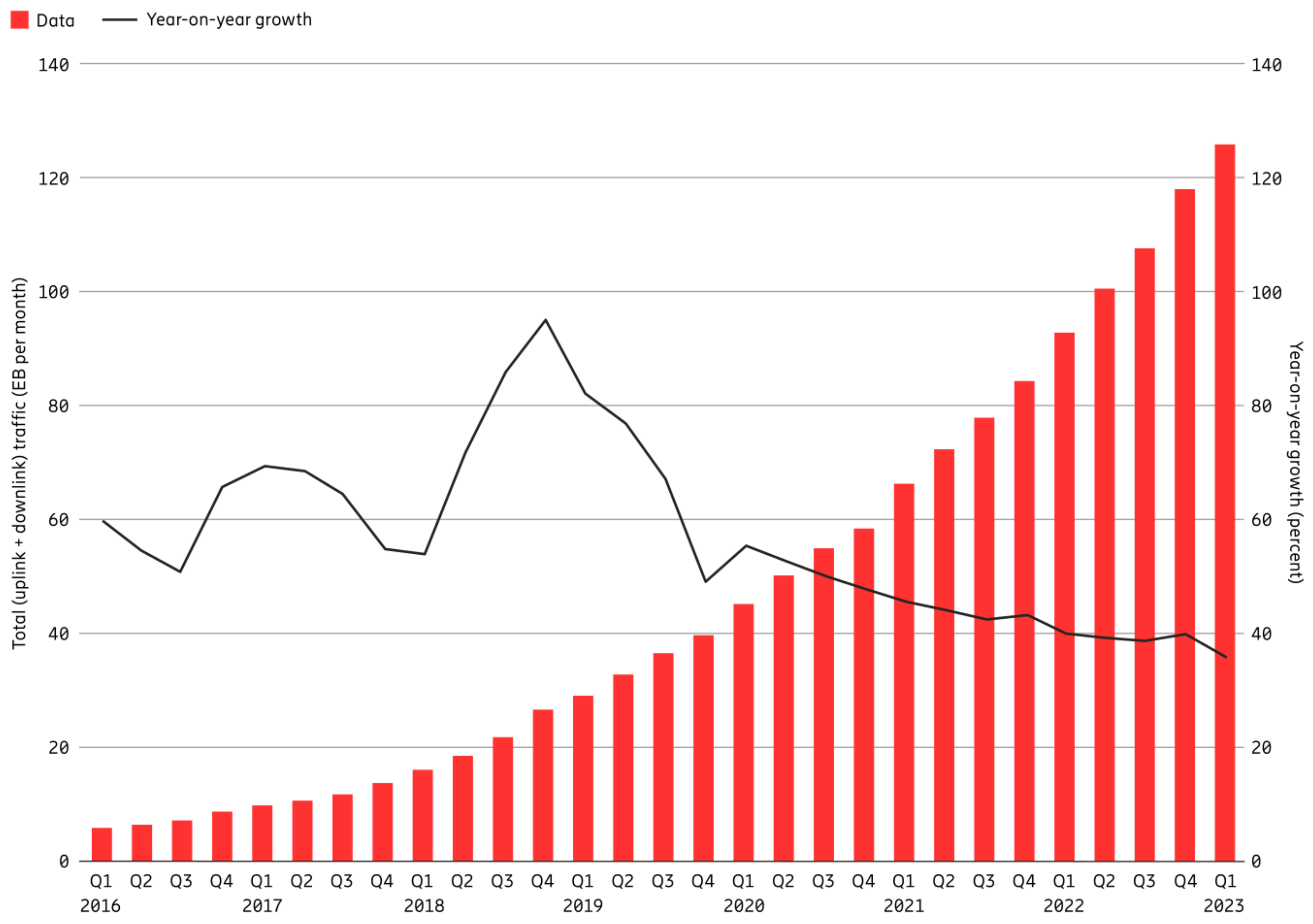
Figure 1 : Évolution 2013-2025 de la part du numérique dans la consommation d'énergie primaire mondiale
(The Shift Project - Forecast Model 2021)

Reasons for this huge growth

Huge traffic increase

- Mobile traffic has increased by 36% in 2022

Figure 18: Global mobile network data traffic and year-on-year growth (EB per month)

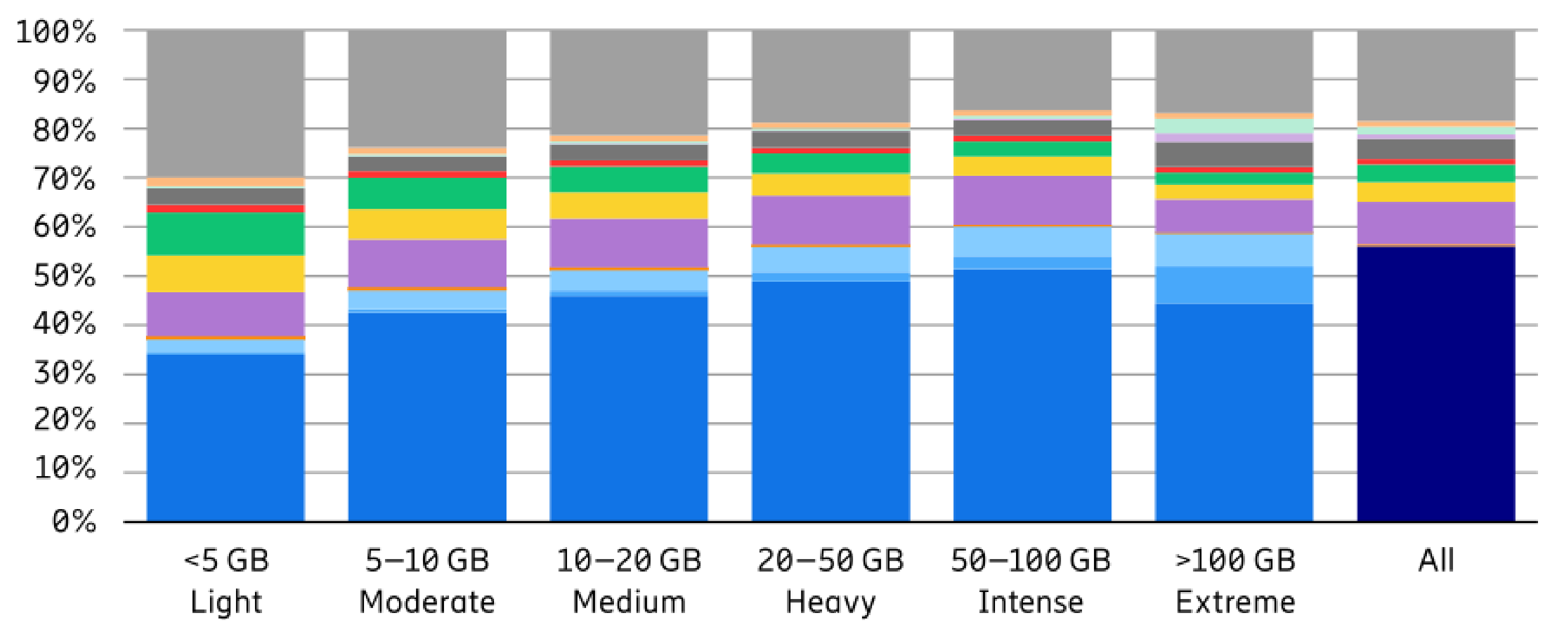


Source: Ericsson Mobility Report, June 2023

Reasons for this huge growth

Huge traffic increase

- The main data are videos (60% of the traffic)



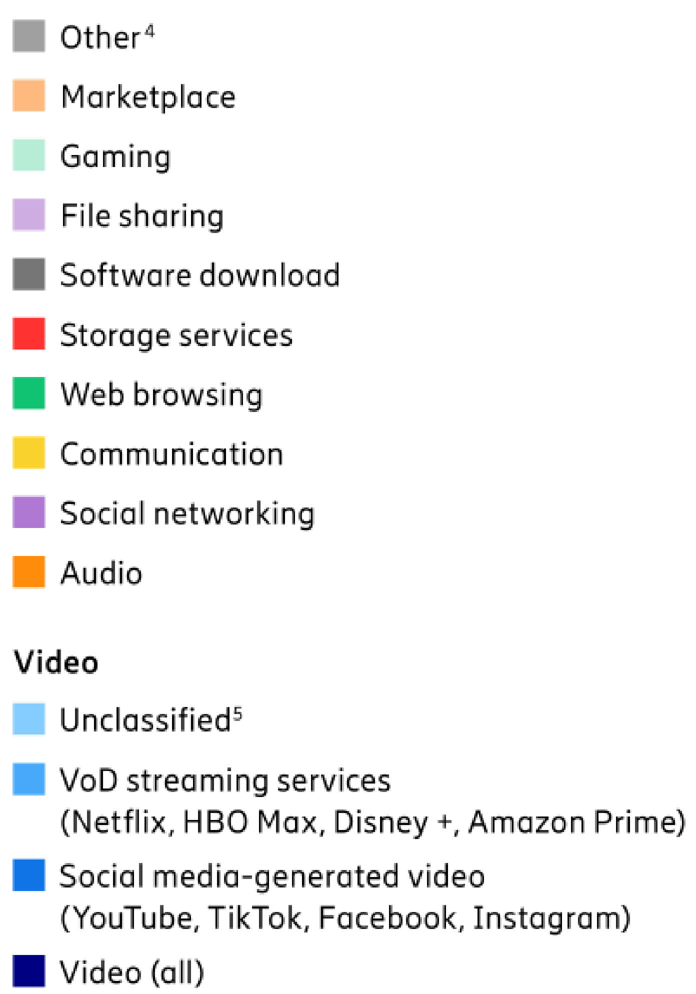
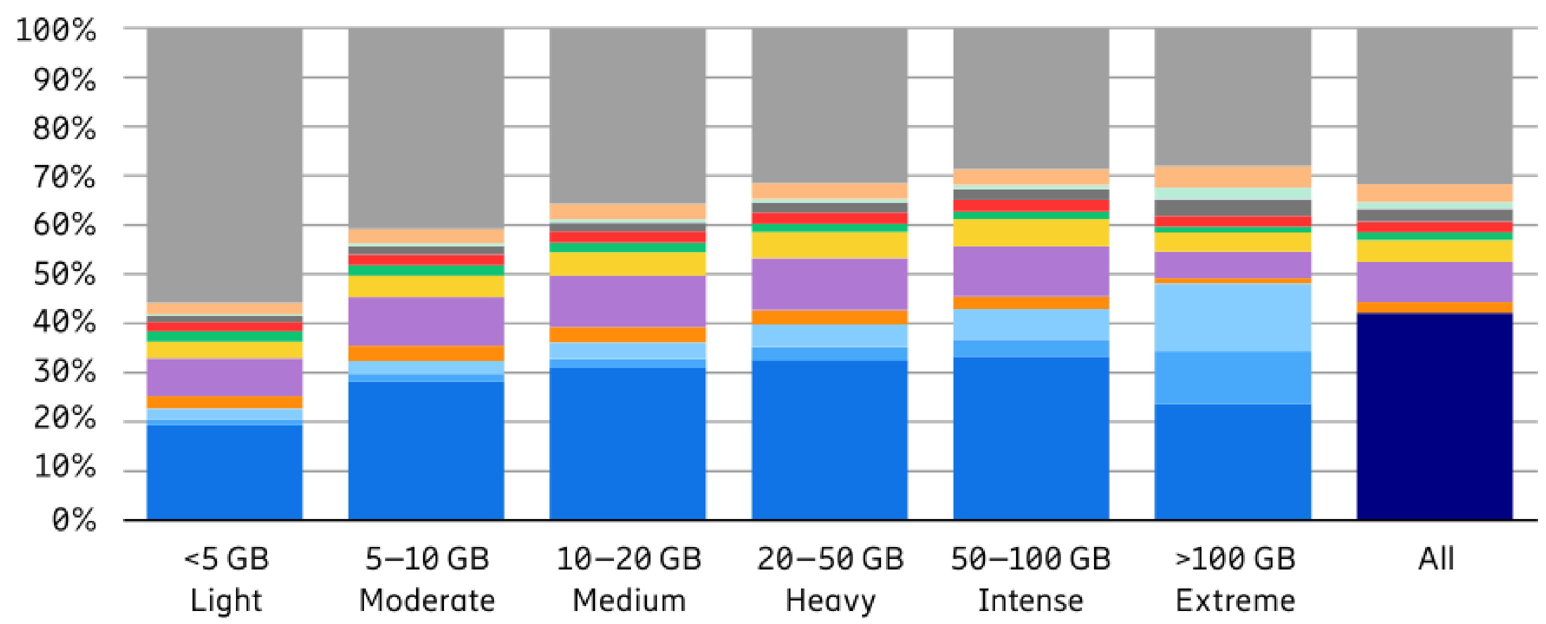


Figure 17: North American service provider: Traffic volume per application type of different subscriber clusters



Source: Ericsson Mobility Report, June 2023

Huge traffic increase

More information

- For some providers: 80% of the traffic comes from the GAFAMs

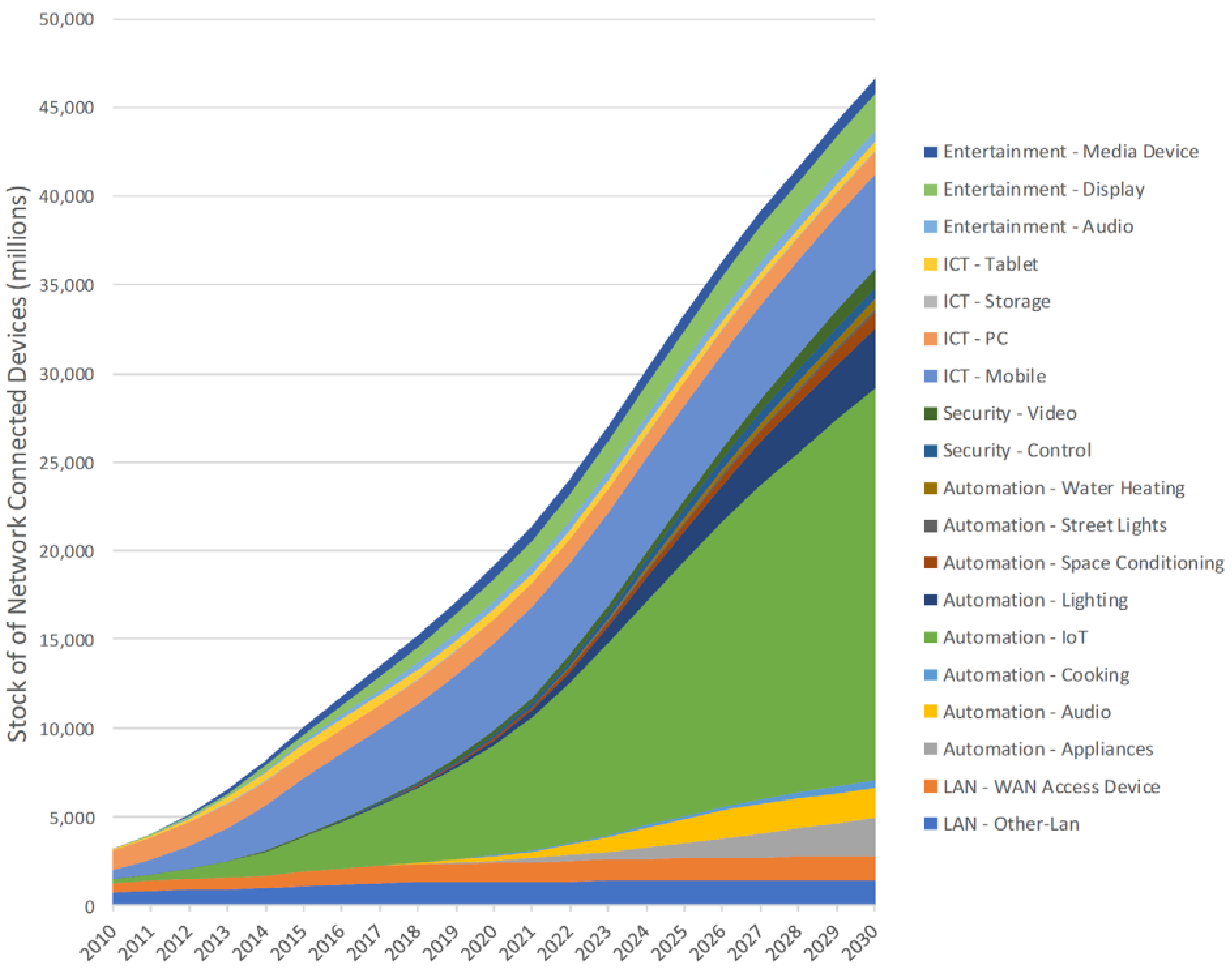
Impact on the Cloud

- Increase of total amount of data stored by data centers: +40% per year
- Data stored in datacenters represent 20% of the total data

Source: The Shift project -- Environmental impacts of digital technology : 5-year trends and 5G governance (2021)

Reasons for this huge growth (2)

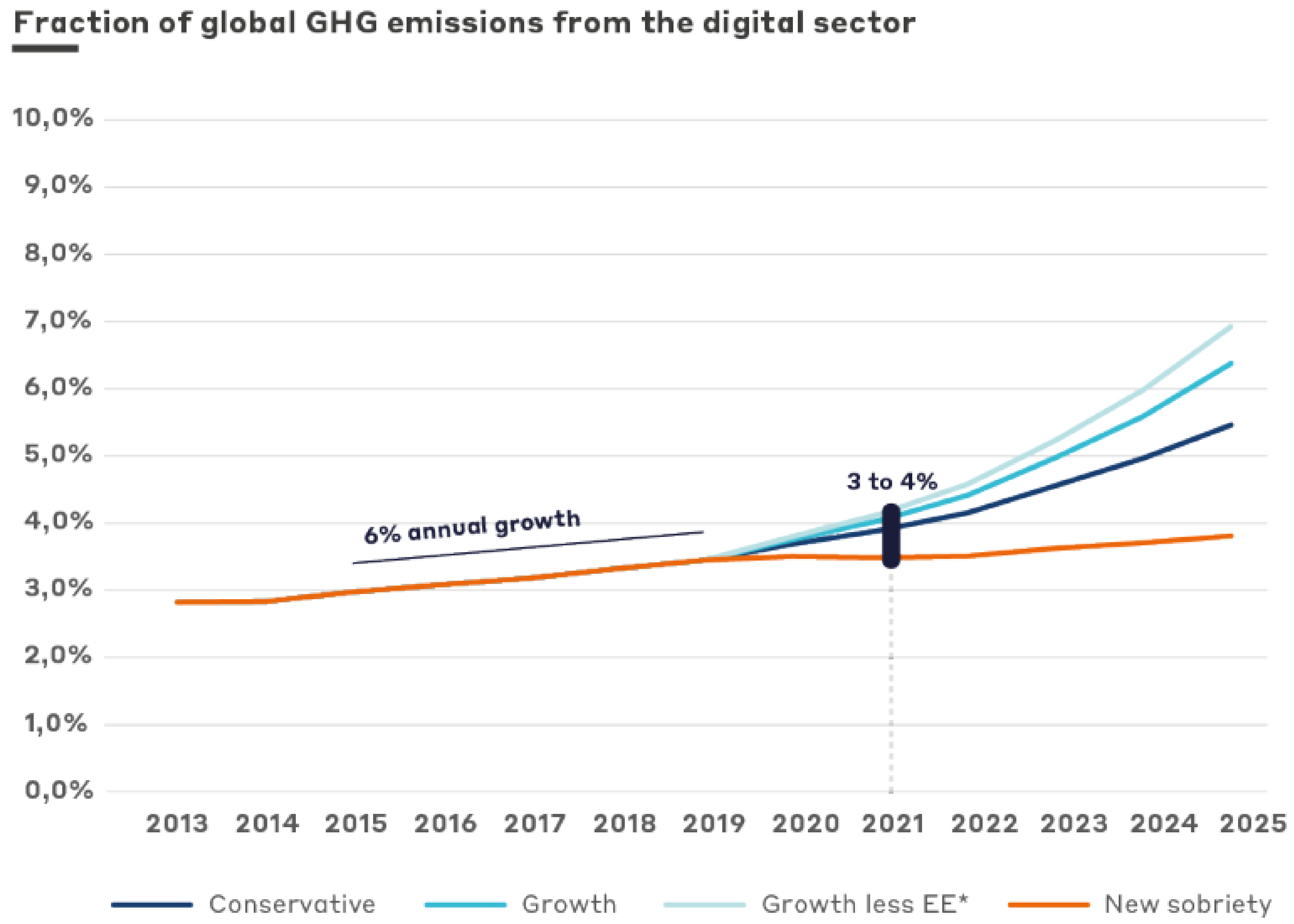
Huge increase in the number of connected devices



Source: The Shift project -- Environmental impacts of digital technology : 5-year trends and 5G governance (2021)

Carbon footprint of the IT domain

- Between 3% and 4%
  - Note that energy is often used as a proxy for evaluating the carbon footprint

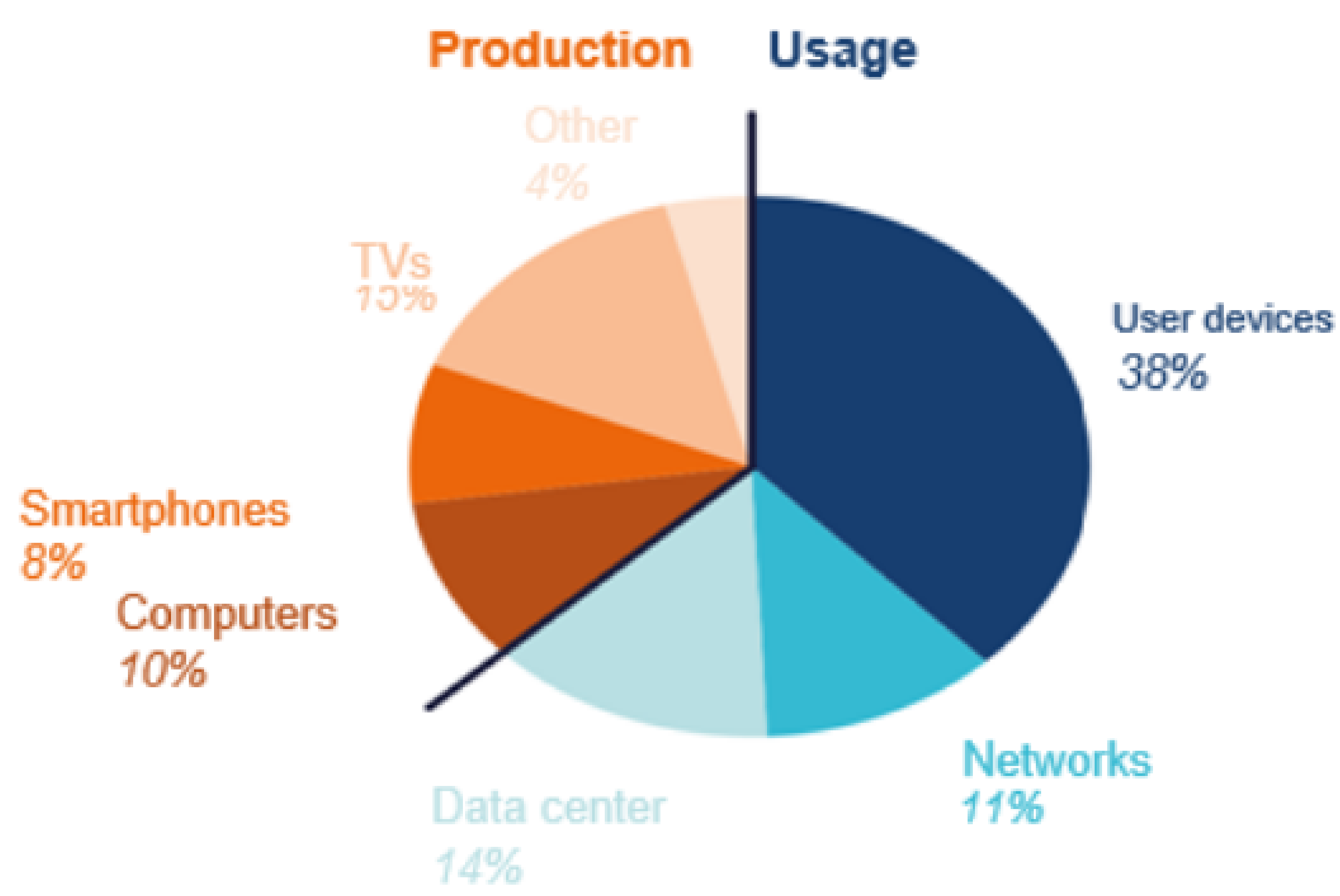


Source: The Shift project -- Environmental impacts of digital technology : 5-year trends and 5G governance (2021)

Carbon footprint: Production vs utilization

The figure presents an estimation of the Distribution of the global carbon footprint from digital technology by hardware unit in 2019

- 40% comes for production
    - In France (low-carbon electricity): Up to 80% attributed to production
- For a smartphone: 80% of the carbon footprint from production



Source: The Shift project -- Environmental impacts of digital technology : 5-year trends and 5G governance (2021)

Environmental impact of production

Not only about carbon footprint

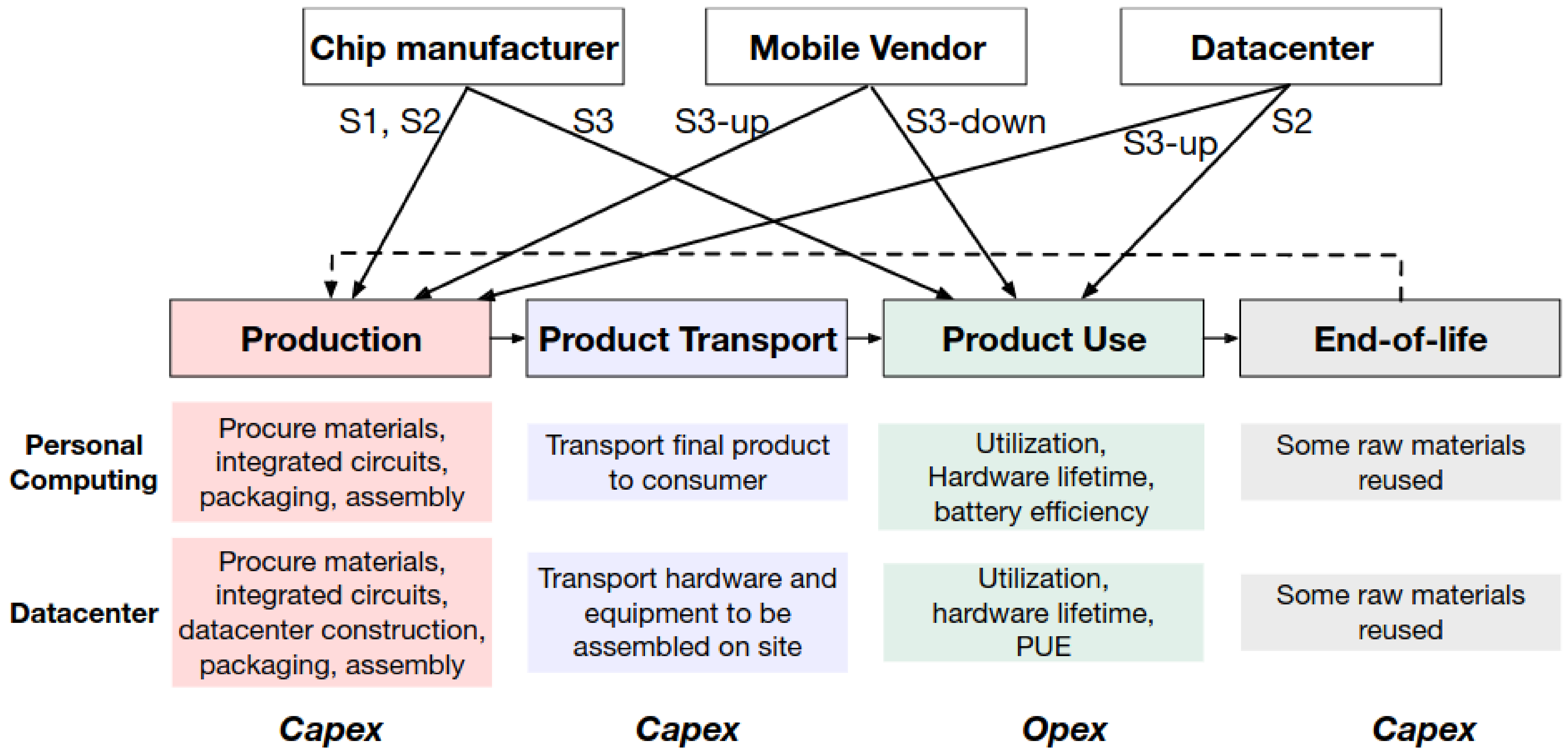
- 30L of water to produce one chip for a smartphone
    - Largest manufacturer companies are in Taiwan
    - 10% of the water of Taiwan used for producing chips

Carbon footprint of datacenters

How to look at the problem

Capex vs Opex

- Capex = Embodied carbon footprint
- Opex = Direct emissions
- They need to be evaluated through life-cycle analyses



Source: Gupta, Udit, et al. "Chasing carbon: The elusive environmental footprint of computing." HPCA 2021.

How to look at the problem

- We will look at the Opex first
    - More difficult to have an impact on the Capex
    - We will come back to it at the end

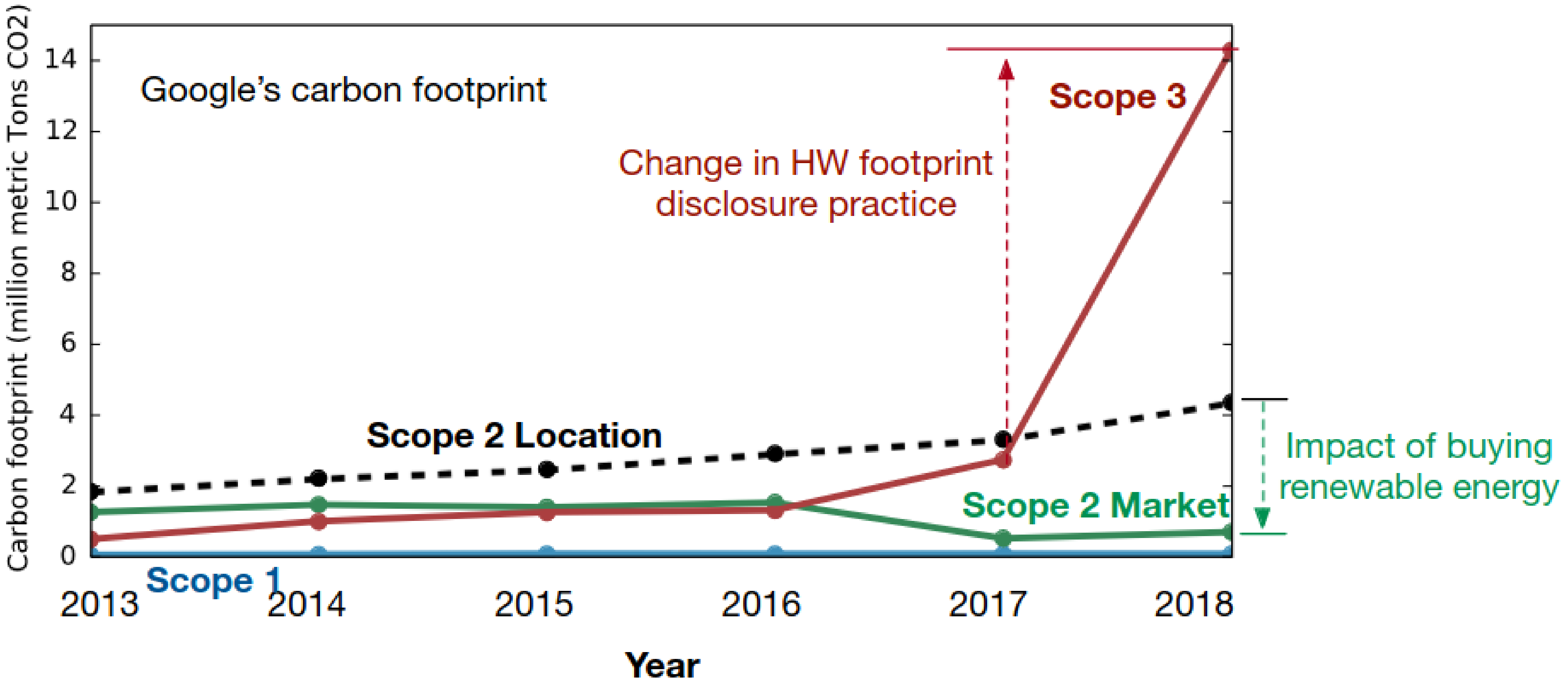
Easy way to improve Opex carbon footprint

How to look at the problem

- We will look at the Opex first
  - More difficult to have an impact on the Capex
  - We will come back to it at the end

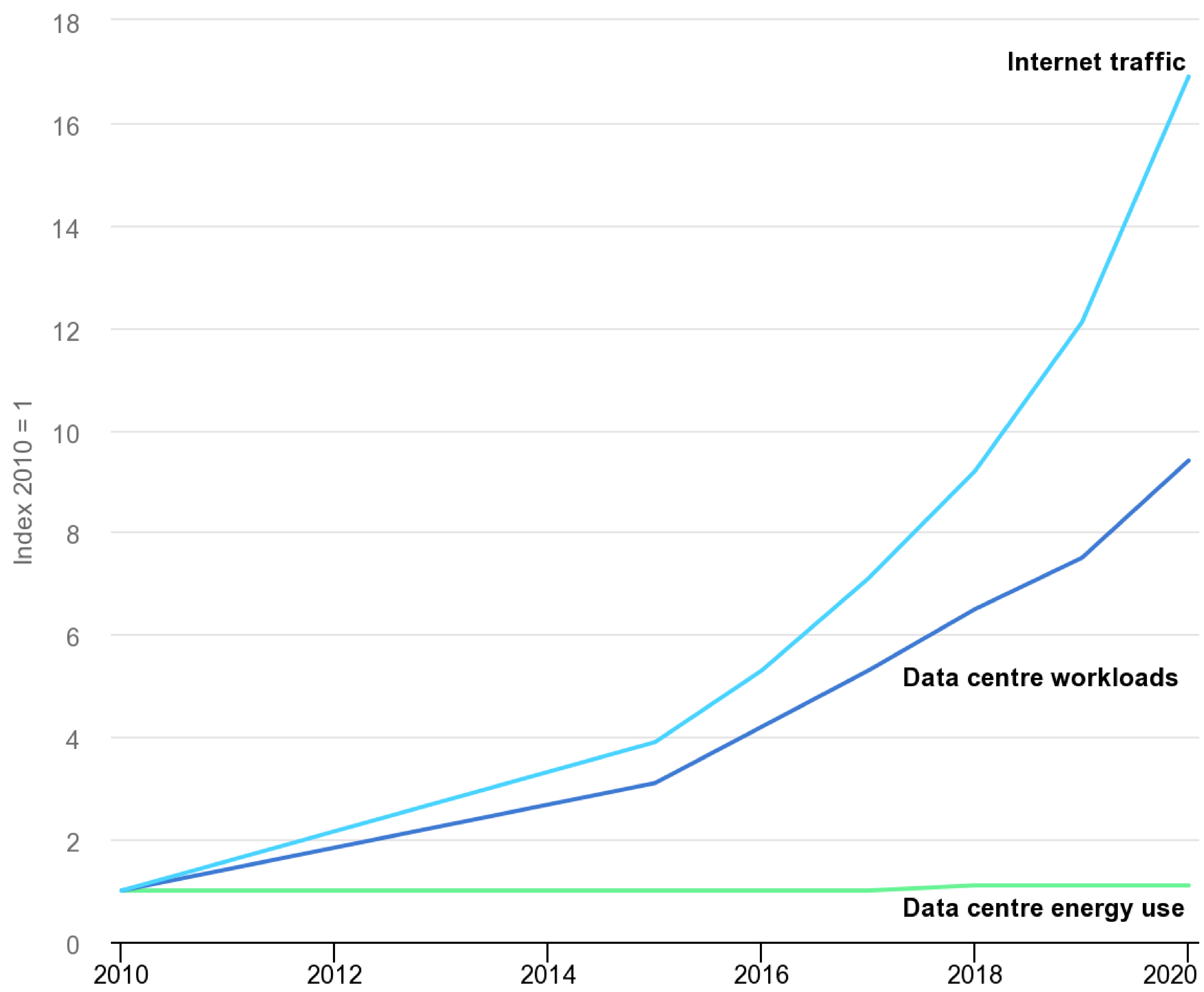
Easy way to improve Opex carbon footprint

- Change the energy source -- renewable energy
  - Example of a google data center (scope 3 = manufacturing)



Energy consumption of data centers

To improve Opex, we need to look at the energy consumption of data centers



Source: *https://www.iea.org/data-and-statistics/charts/global-trends-in-internet-traffic-data-centres-workloads-and-data-centre-energy-use-2010-2020*

Energy consumption of data centers

Good news or bad news

Energy consumption of data centers

Good news or bad news

- Positive side
  - Datacenters energy efficiency has improved dramatically
- Negative side
  - The increase in workload is so big that it nullifies efficiency improvements

The goal should be to decrease the absolute energy consumption

The rebound effect

Problem with the optimization of energy

- It almost always leads to an increase in the usage
    - If I have a more efficient car, I use it more
- Improving the energy efficiency implies that:
    - Reduction of the costs of the goods
    - Possibility to improve the service
- Does it apply to all domains of CS? (see: Woodruff, Jackson, et al. "When Does Saving Power Save the Planet?." Proceedings of the 2nd Workshop on Sustainable Computer Systems. 2023.)

Jevons Paradox

The rebound effect

Problem with the optimization of energy

- It almost always leads to an increase in the usage
    - If I have a more efficient car, I use it more
- Improving the energy efficiency implies that:
    - Reduction of the costs of the goods
    - Possibility to improve the service
- Does it apply to all domains of CS? (see: Woodruff, Jackson, et al. "When Does Saving Power Save the Planet?." Proceedings of the 2nd Workshop on Sustainable Computer Systems. 2023.)

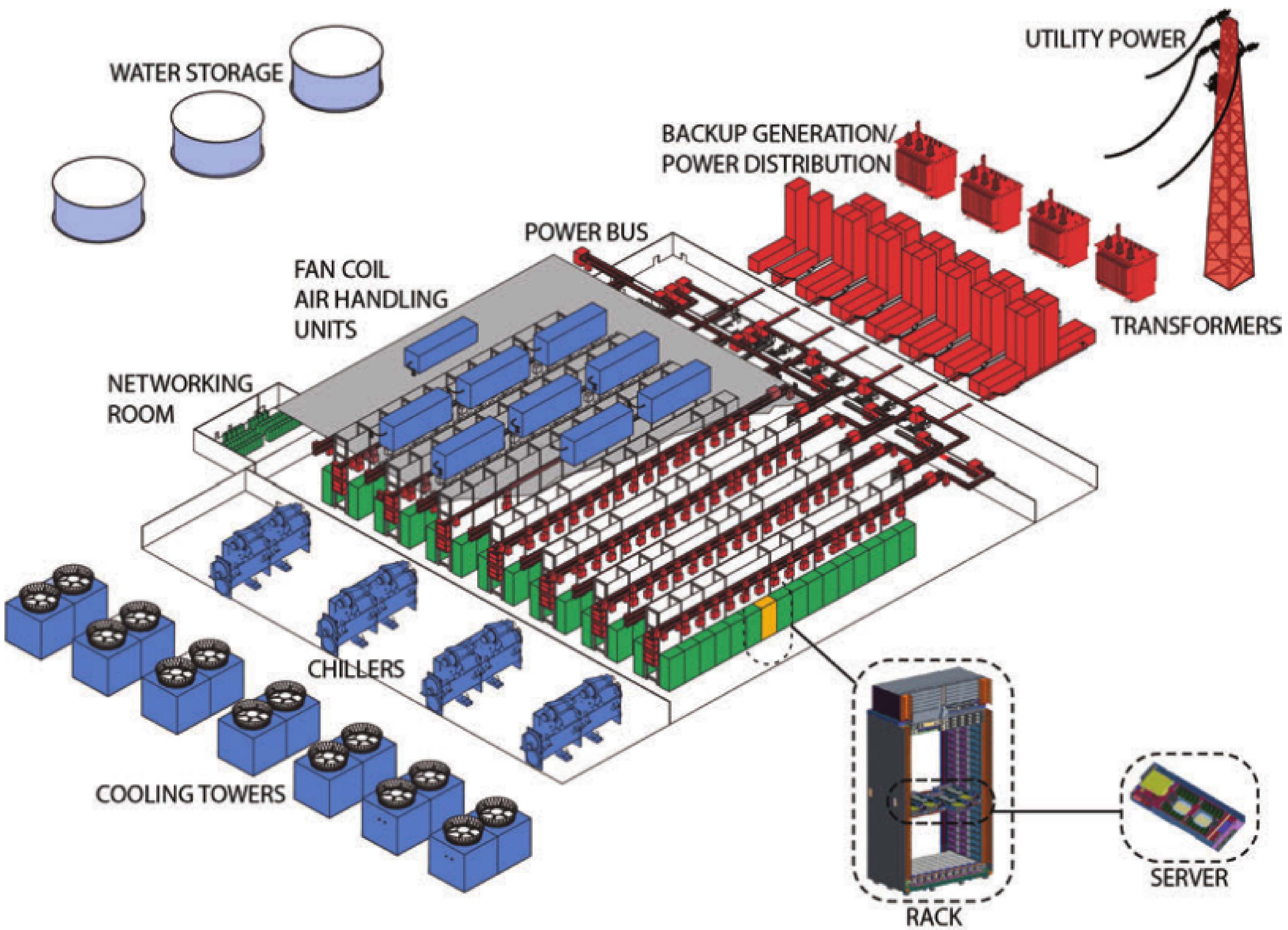
Jevons Paradox

- First discussed the the 19's century about coal

The rebound effect might be so important that it leads to an increase in energy usage

Optimizing the energy consumption of data centers

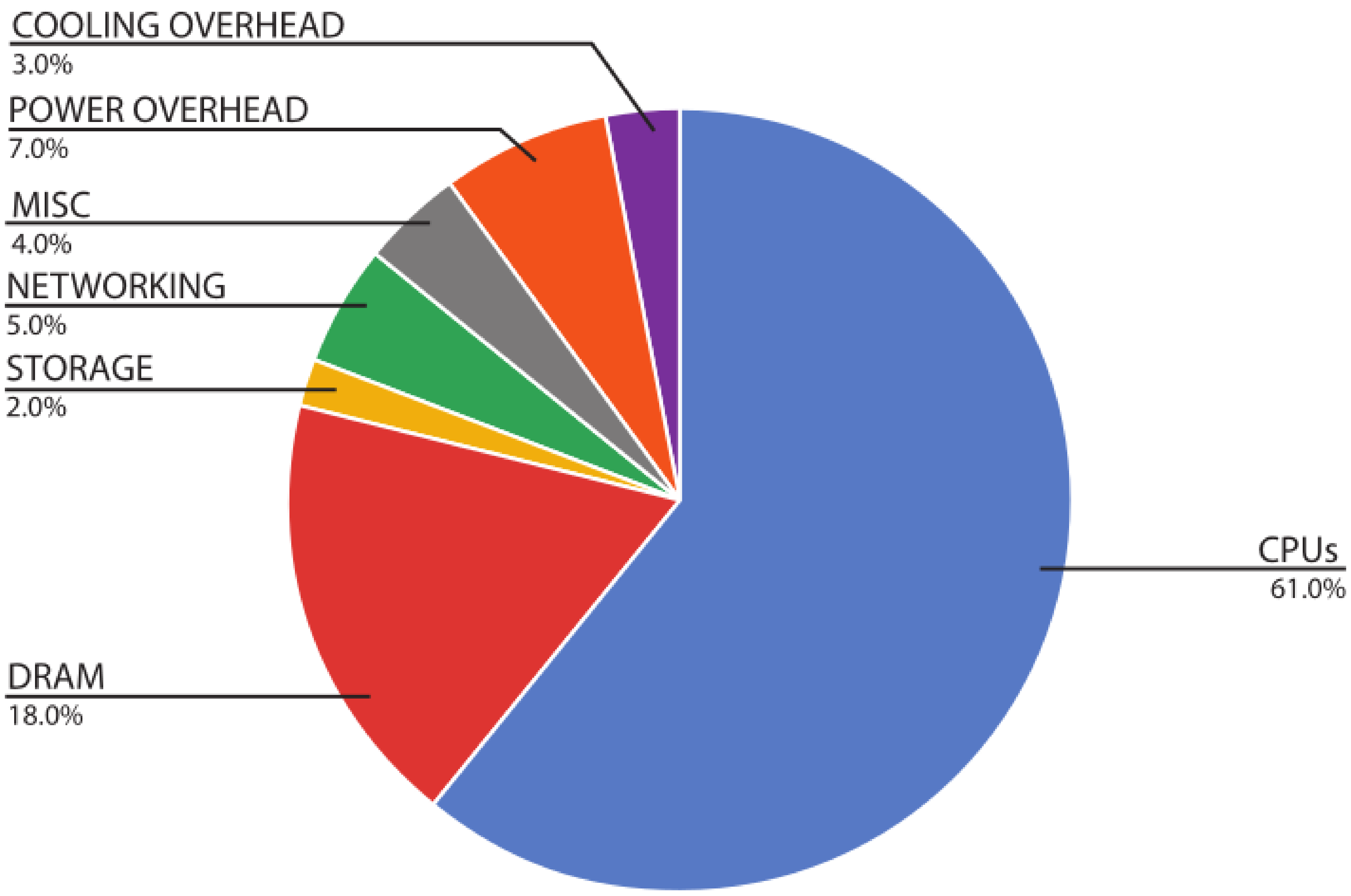
What consumes energy?



Source: The datacenter as a computer

What consumes energy?

Peak power usage for a 2-socket server at 80% of max utilization (2017)



Power Usage Efficiency (PUE)

We want to optimize energy efficiency:

$$\text{Energy Efficiency} = \frac{\text{Work}}{\text{Used energy}}$$

PUE

- Power Usage Efficiency captures the quality of the datacenter building
    - How much energy is used for something else than *computing*?
- SPUE is the PUE at the level of a server

$$\text{Energy Efficiency} = \frac{1}{\text{PUE}} \times \frac{1}{\text{SPUE}} \times \frac{\text{Work}}{\text{Energy used by electronic components}}$$

Power Usage Efficiency

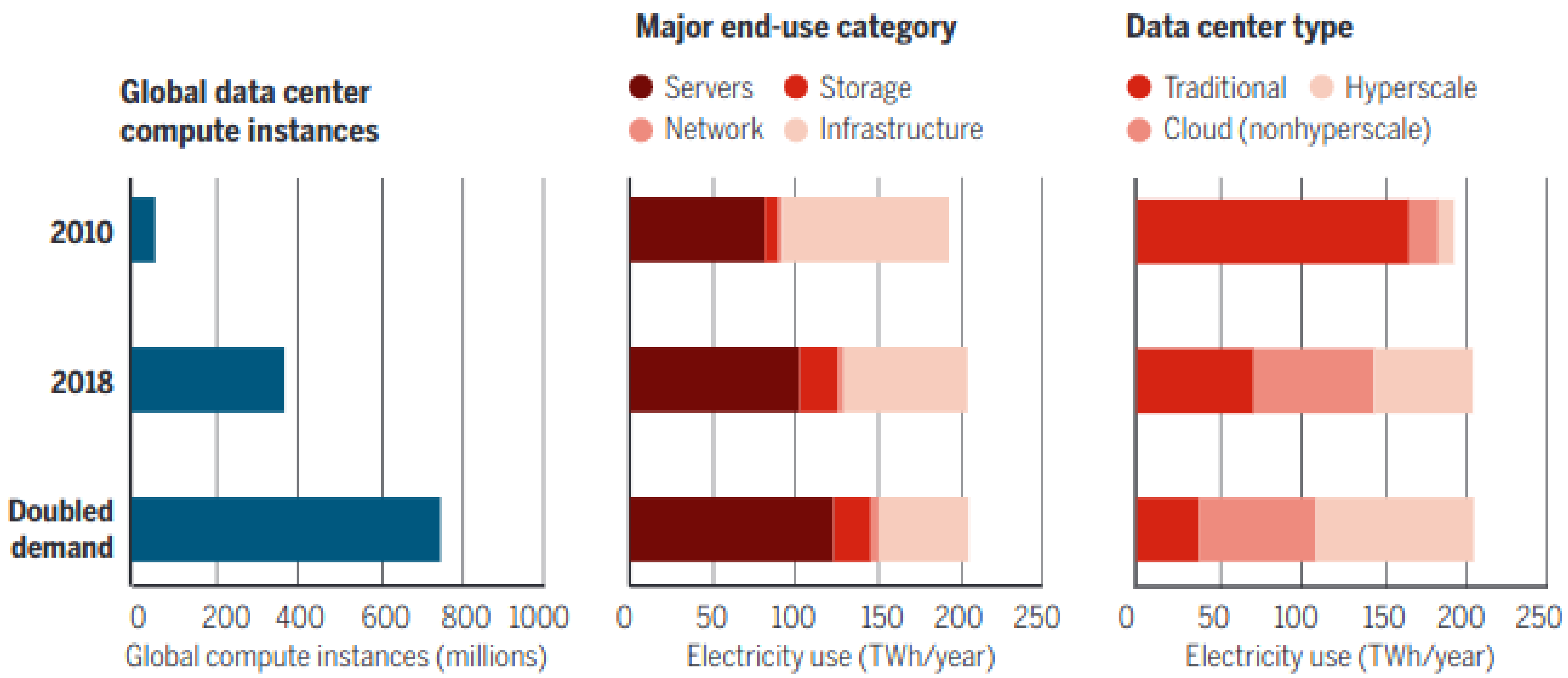
Evolution of the PUE

- Before 2006, the PUE of most datacenters was above 3

- Today:
    - Traditional DC have a PUE between 1.6 and 2.5
    - Hyperscalers have a PUE below 1.2
        - Google DC average PUE is below 1.1

About hyperscalers

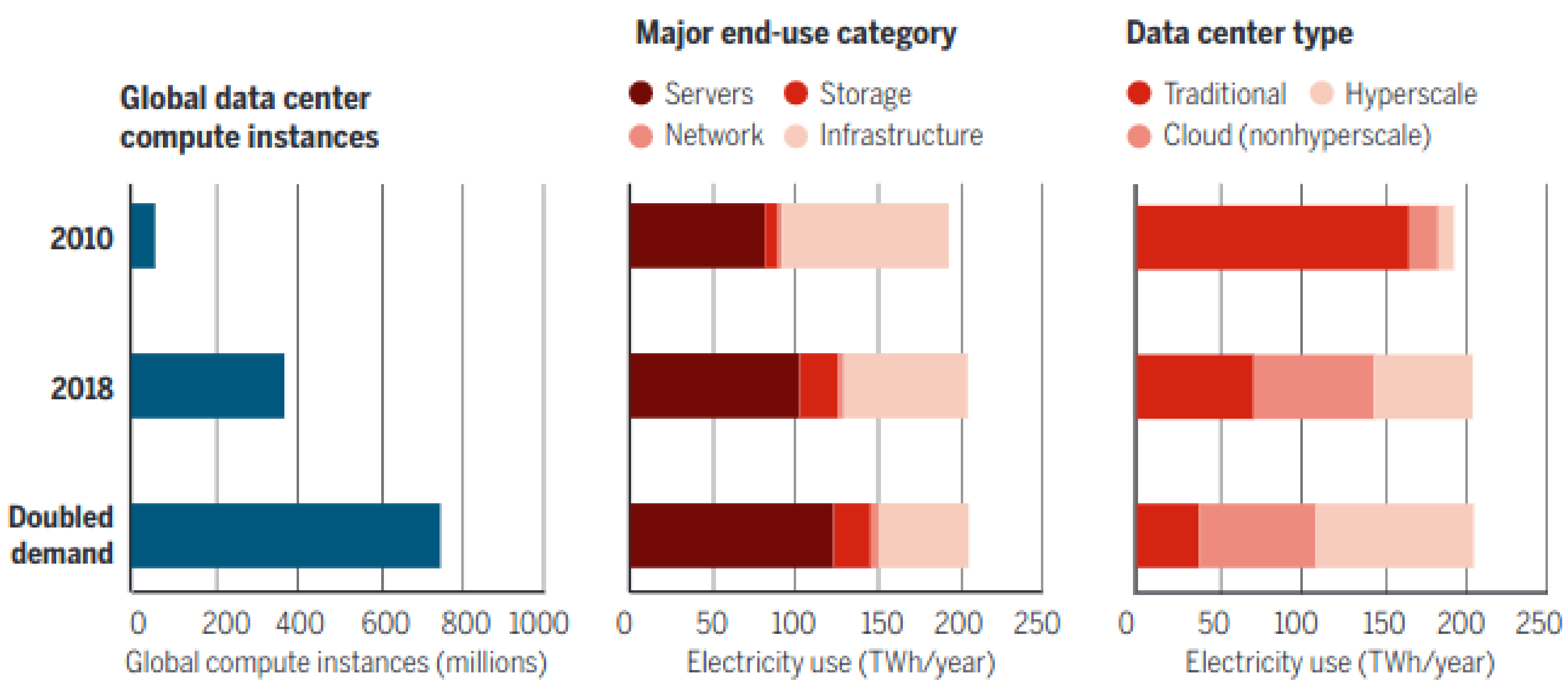
Trend: More and more applications are hosted by hyperscalers



Good news?

About hyperscalers

Trend: More and more applications are hosted by hyperscalers



Good news?

- Yes because better PUE
- No because more data movements

It is a complex question

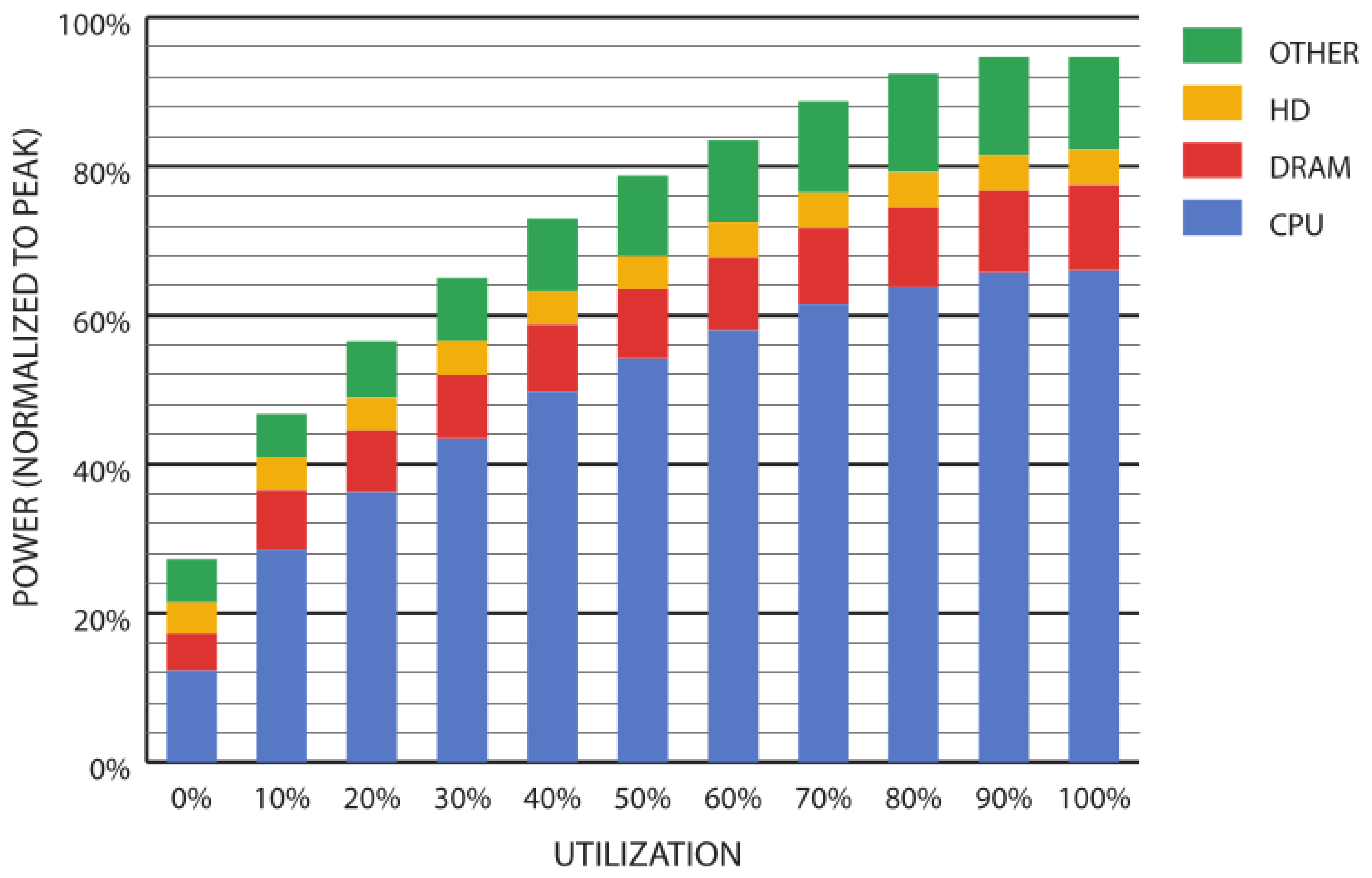
Energy proportionality

Definitions

- **An energy proportional system is one in which the energy consumed by the system is directly proportional to the activity**
- Energy consumed by a device = static energy + dynamic energy
    - Static energy: Energy consumed when the activity is null
    - Dynamic energy: Energy variations depending on the activity

Energy proportionality

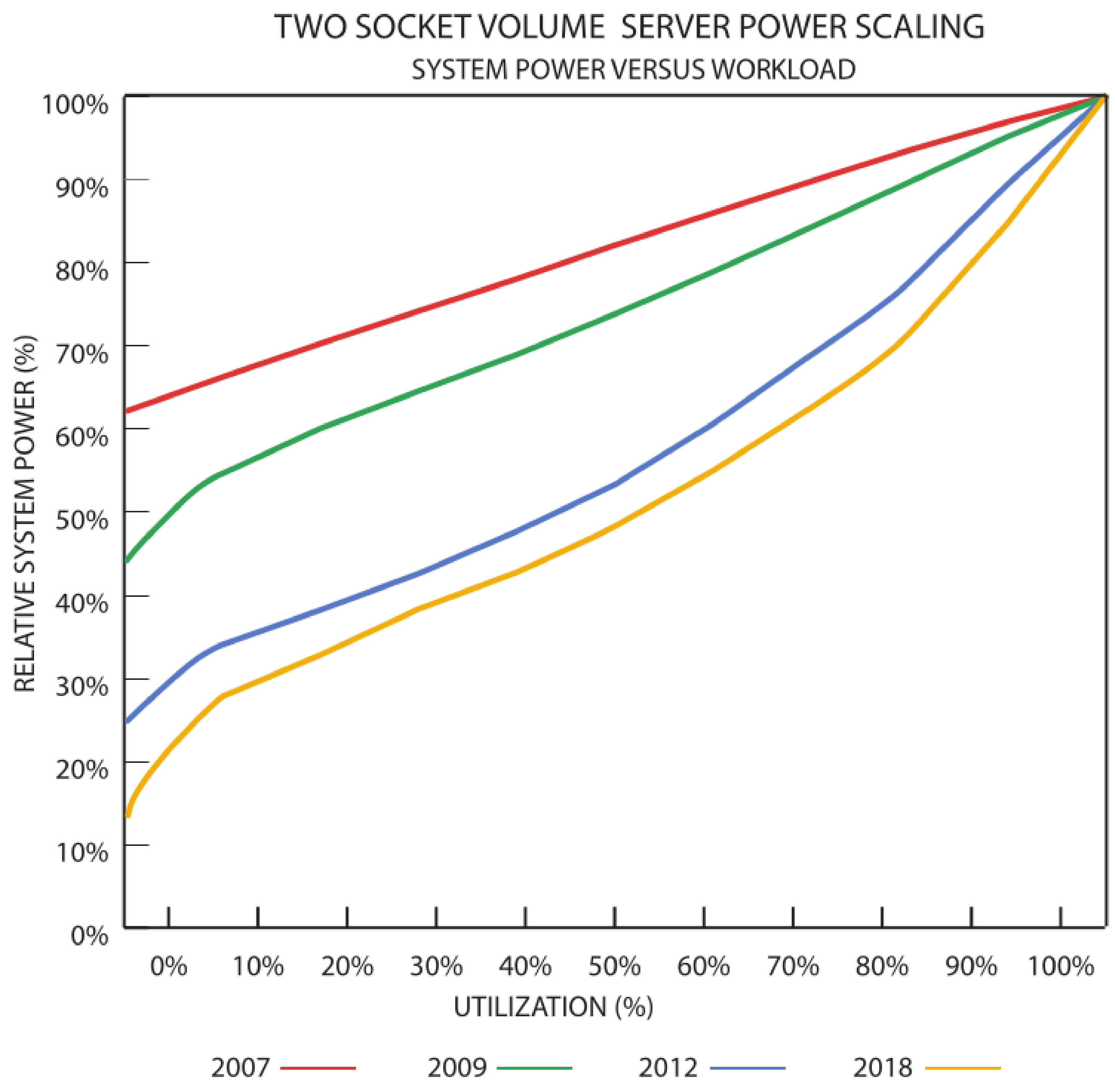
Energy proportionality of a x86 server



- The CPU is not the only thing to focus on
- The network is an example of device with a high static energy consumption

Energy proportionality

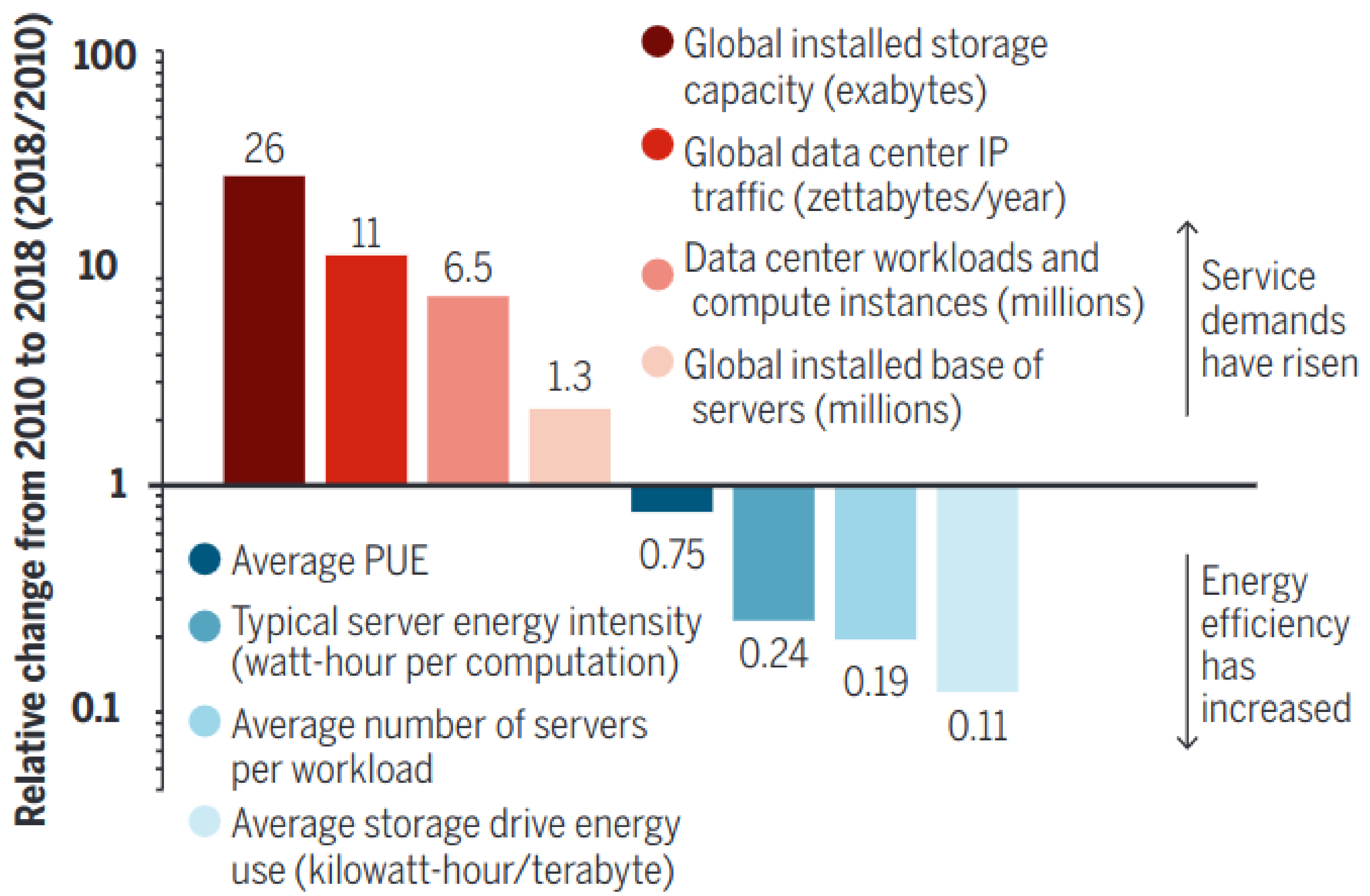
At the level of the CPU



Huge improvements have been made:

- Dynamic Voltage and Frequency Scaling (DVFS)
- Sleep states

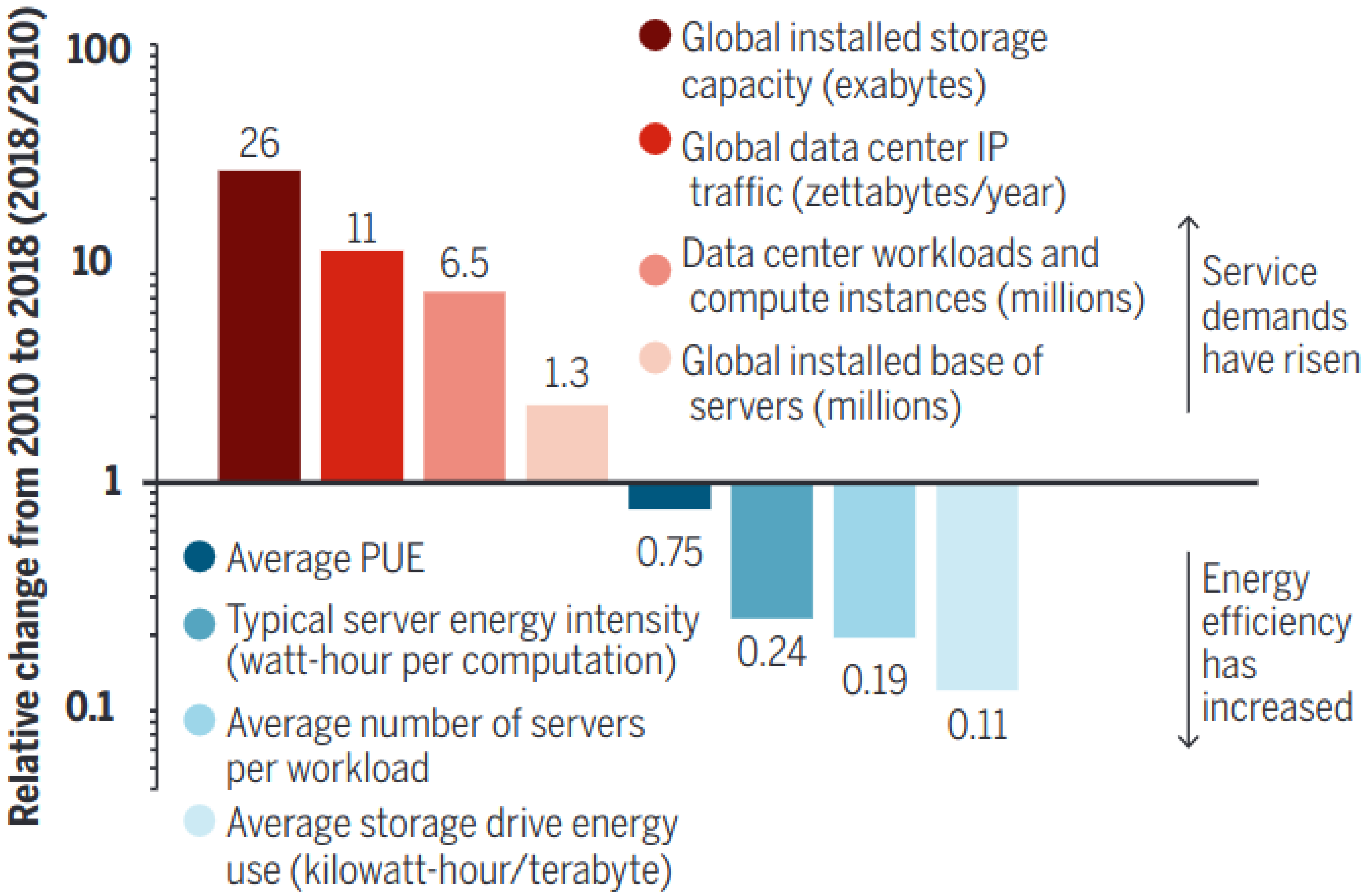
A summary



Is it a good result?

Masanet, Eric, et al. "Recalibrating global data center energy-use estimates." Science 367.6481 (2020): 984-986.

A summary



Is it a good result?

Not really! Energy consumption did not decrease

Masanet, Eric, et al. "Recalibrating global data center energy-use estimates." Science 367.6481 (2020): 984-986.

Improving the embodied carbon footprint of datacenters

A difficult problem

- Not so many factors we can have an impact on

A difficult problem

- Not so many factors we can have an impact on
- Main directions
    - Increase the lifetime of servers
    - Select the hardware carefully

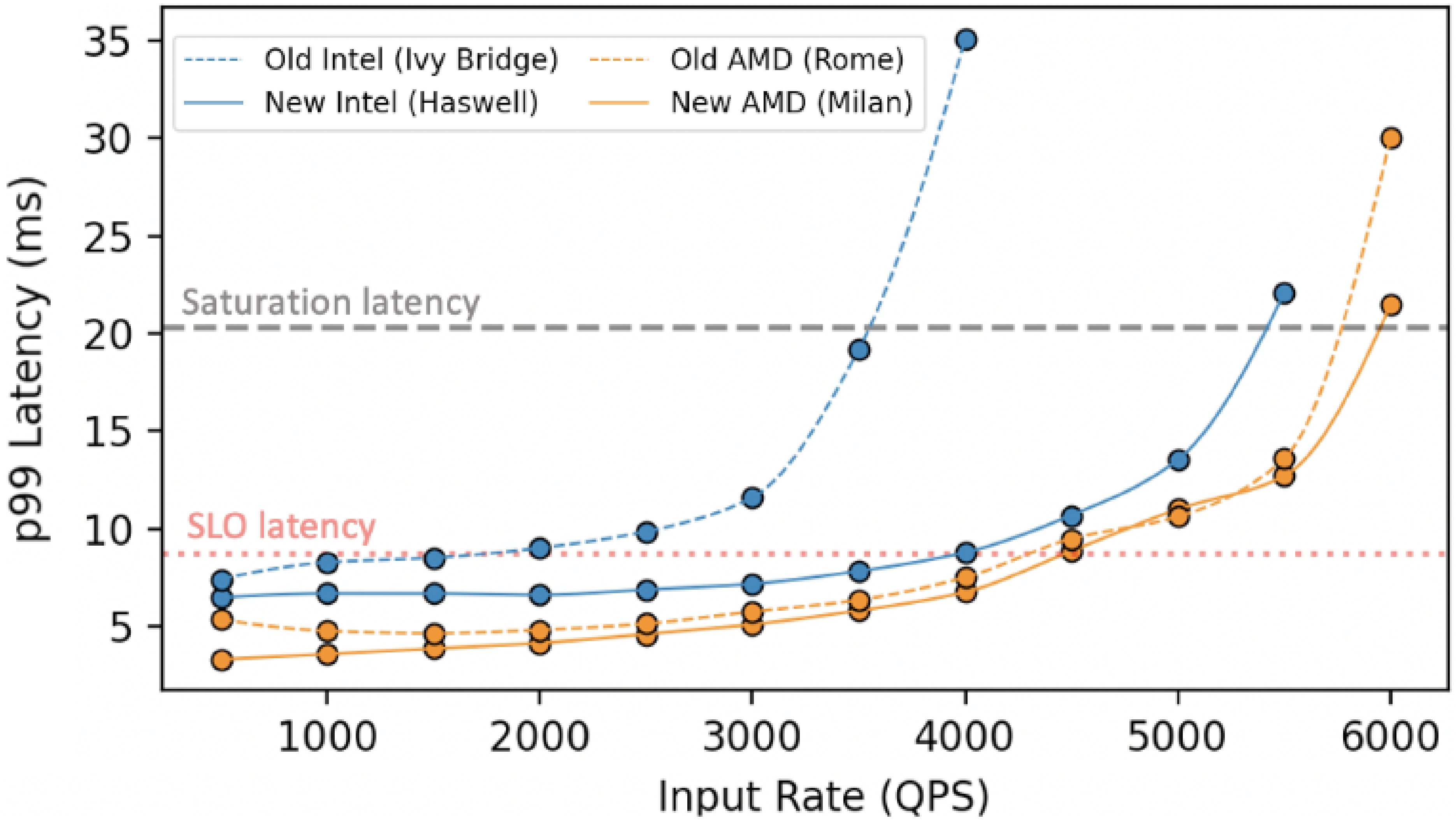
Problem: Evaluating/Collecting information about the embodied carbon footprint can be difficult

Increasing the lifespan of servers

- Using the servers for longer period allows amortizing the embodied carbon footprint
    - What about the reliability of the hardware?
    - What about performance?

Increasing the lifespan of servers

Performance impact



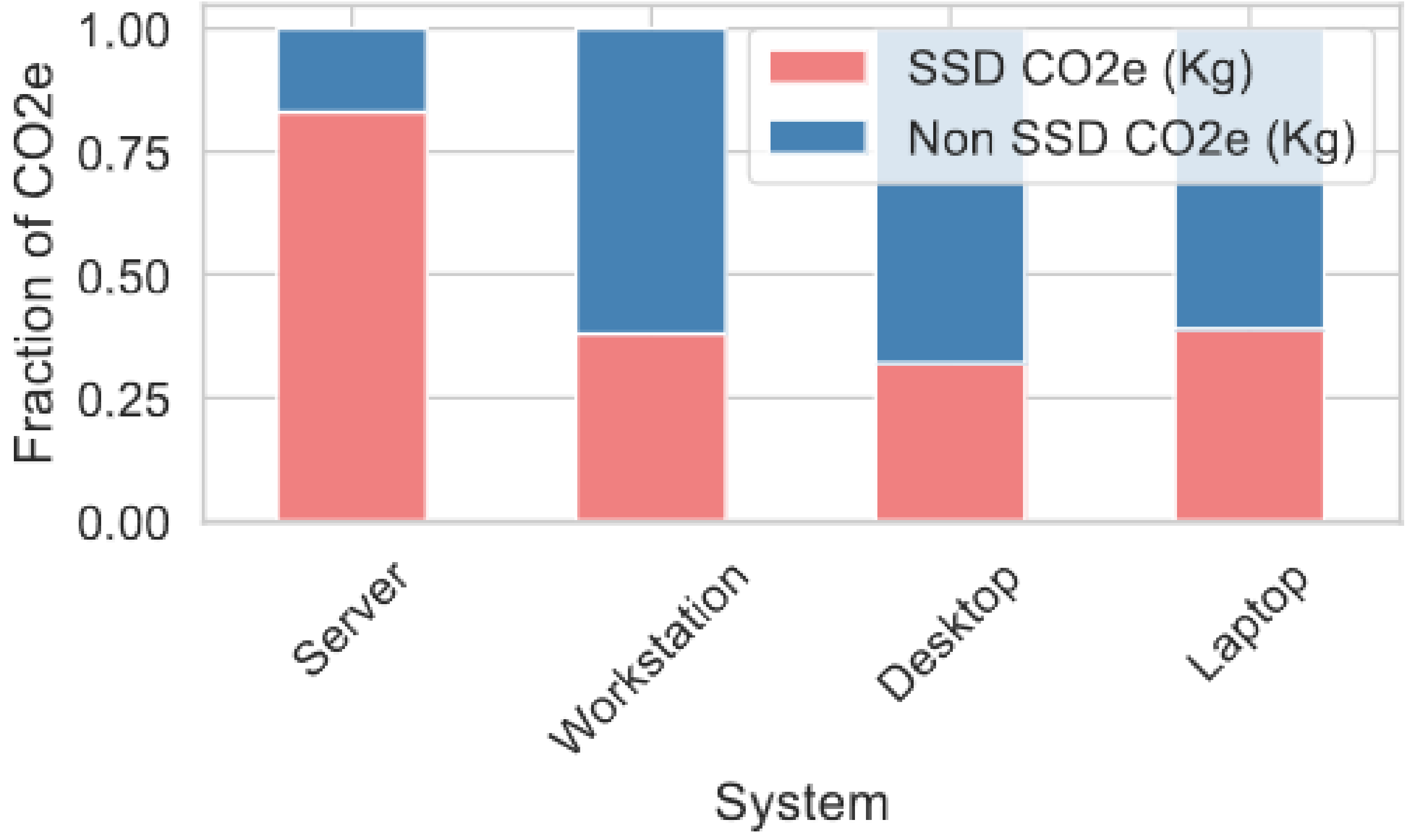
- DeathStarBench application deployed over 15 servers of a given type
  - Intel servers (2012, 2013)
  - AMD servers (2019, 2021)
- SLO = 75% of saturation for the best performing server
  - For low load, old servers can work

See: Wang, Jaylen, Udit Gupta, and Akshitha Sriraman. "Peeling Back the Carbon Curtain: Carbon Optimization Challenges in Cloud Computing." Proceedings of the 2nd Workshop on Sustainable Computer Systems. 2023.

Performance improvement vs Embodied Carbon footprint

A difficult question to address

- SSDs are much more efficient than hard drives but:
  - The embodied Carbon Footprint of SDDs is 8x higher



See: Tannu, Swamit, and Prashant J. Nair. "The dirty secret of ssds: Embodied carbon." ACM SIGENERGY Energy Informatics Review 3.3 (2023): 4-9.