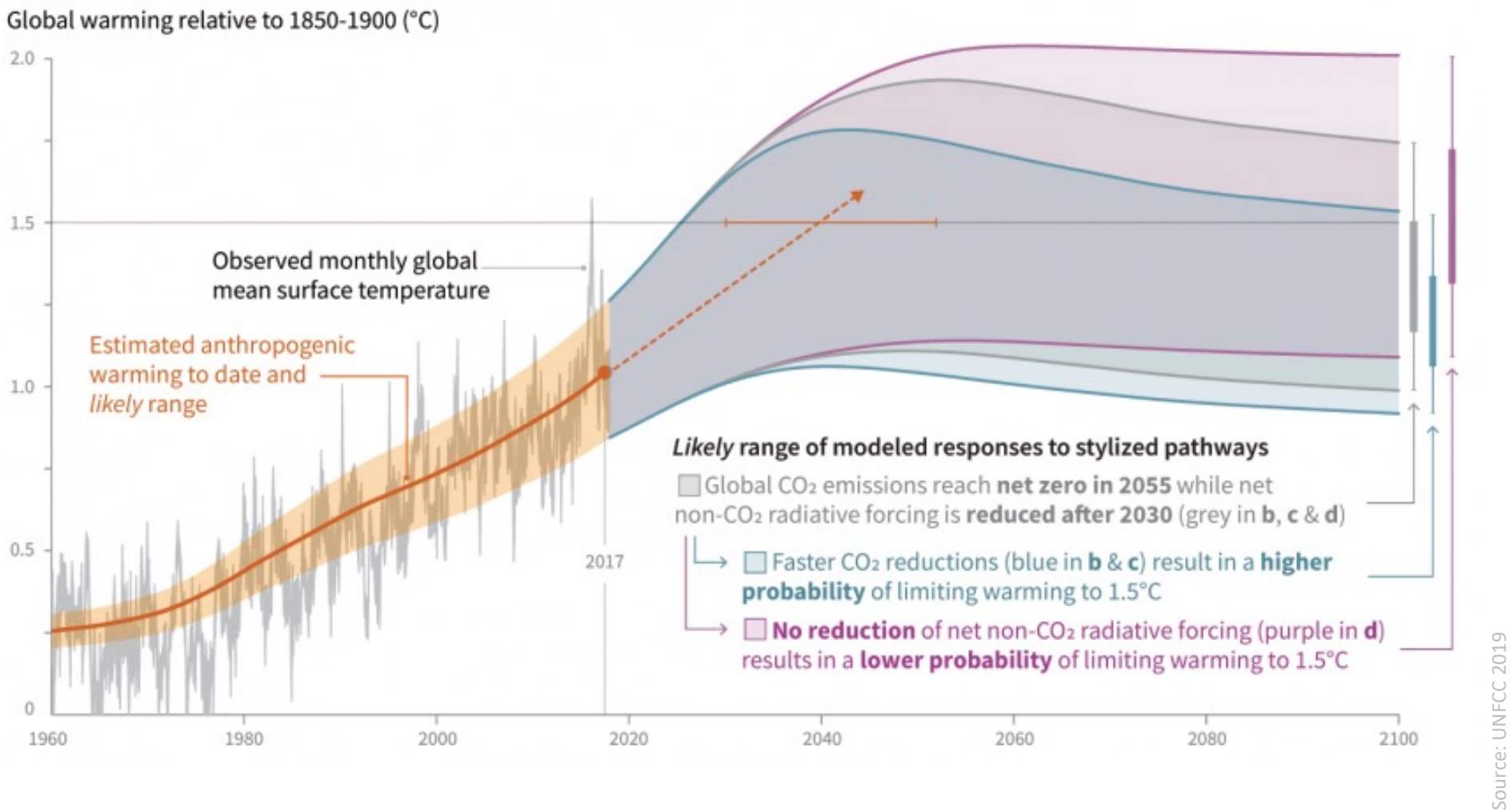
L'informatique verte … moins branchée ?

Anne-Cécile Orgerie

Espace des sciences, Rennes
8 février 2022

Contexte : 1,5°C max



Objectif 2019 : réduire les émissions de GES de 8% par an sur 2020-2030

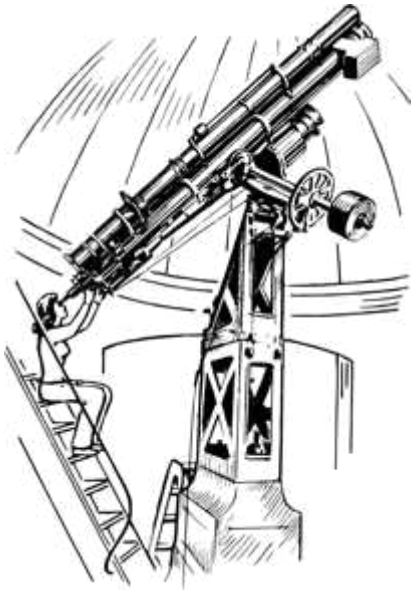
[Source : https://unfccc.int/news/cut-global-emissions-by-76-percent-every-year-for-next-decade-to-meet-15degc-paris-target-un-report]

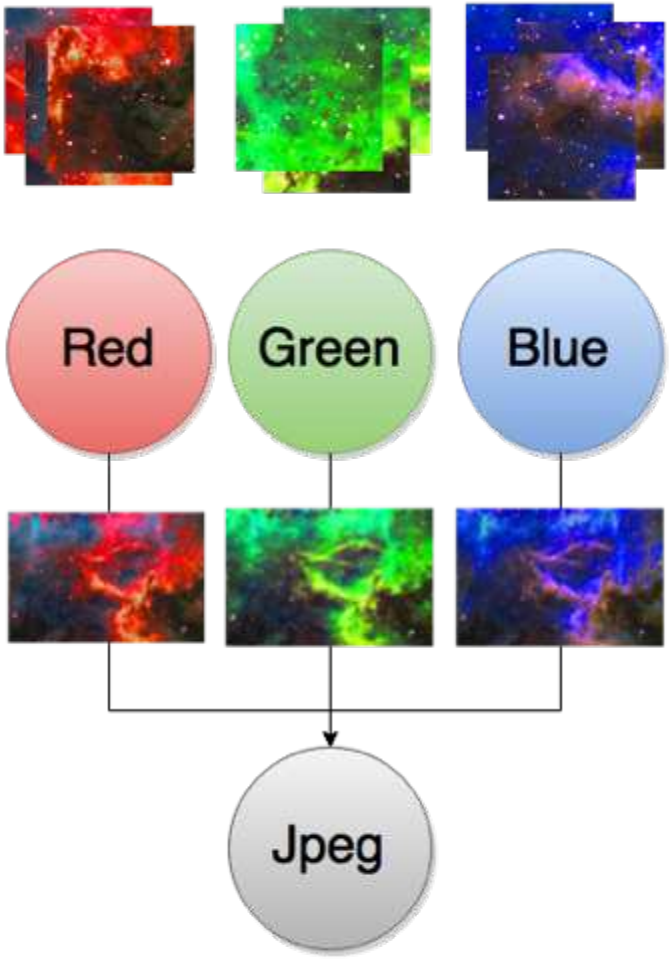
Mon contexte scientifique

- **Comprendre**
- Consommation énergétique
- Systèmes distribués à large échelle
- Parties calcul et communication
- En phase usage

Commencé il y a quelques années avec le *Grid computing*…

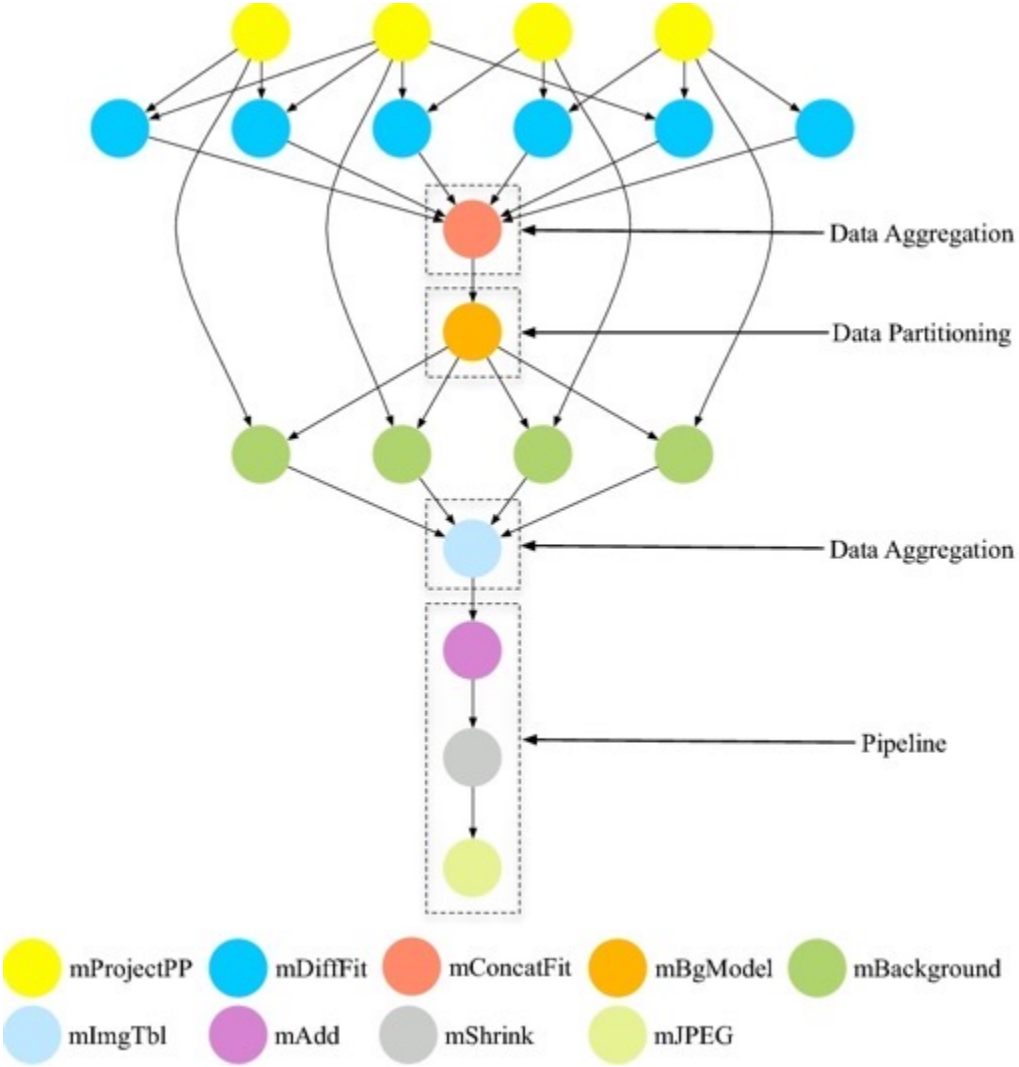
Des applications scientifiques





Les Pléiades

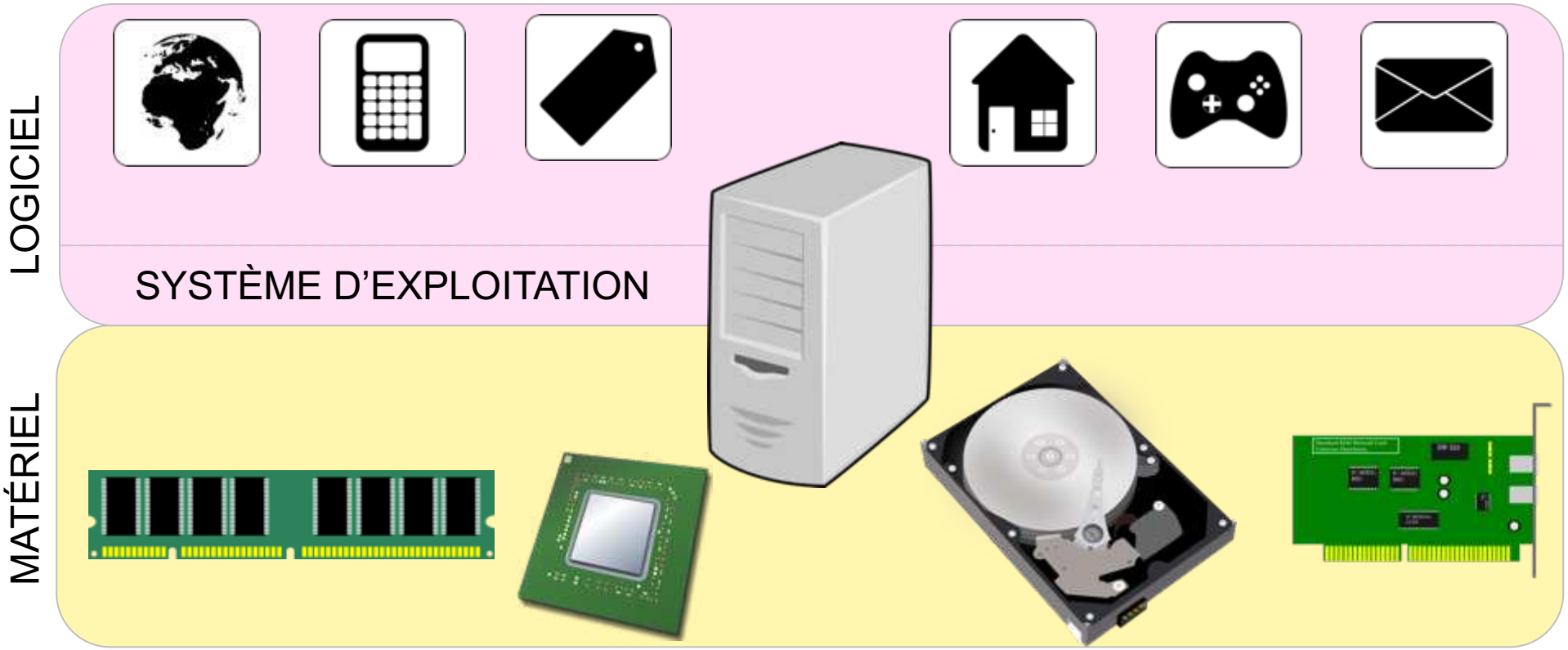
Des graphes de tâches




```python
def main():

    parser = argparse.ArgumentParser()

    parser.add_argument("--work-dir", action = "store", dest = "work_dir",
                        help = "Work directory to chdir to")
    parser.add_argument("--center", action = "store", dest = "center",
                        help = "Center of the output, for example M17 or 56.5 23.75")
    parser.add_argument("--degrees", action = "store", dest = "degrees",
                        help = "Number of degrees of side of the output")
    parser.add_argument("--band", action = "append", dest = "bands",
                        help = "Band definition. Example: dss:DSS2B:red")
    parser.add_argument("--tc-target", action = "store", dest = "tc_target",
                        help = "Transformation catalog: regular or container")
    args = parser.parse_args()

    if args.work_dir:
        os.chdir(args.work_dir)

    if args.tc_target is None:
        args.tc_target = "regular"

    if os.path.exists("data"):
        print("data/ directory already exists")
        sys.exit(1)
    os.mkdir("data")

    dax = AutoADAG("montage")
```

[Source : https://github.com/pegasus-isi/montage-workflow-v2]

Exécuter une application



Fugaku (N°1 TOP500 novembre 2021)

442 Petaflops, **29,9 MW**, 7 630 848 cœurs

[Source : https://spectrum.ieee.org/tech-talk/computing/hardware/japans-fugaku-supercomputer-is-first-in-the-world-to-simultaneously-top-all-high-performance-benchmarks]

Top500

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	**Supercomputer Fugaku** - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,630,848	442,010.0	537,212.0	29,899
2	**Summit** - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM DOE/SC/Oak Ridge National Laboratory United States	2,414,592	148,600.0	200,794.9	10,096
3	**Sierra** - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States	1,572,480	94,640.0	125,712.0	7,438
4	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway, NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
5	**Perlmutter** - HPE Cray EX235n, AMD EPYC 7763 64C 2.45GHz, NVIDIA A100 SXM4 40 GB, Slingshot-10, HPE DOE/SC/LBNL/NERSC United States	761,856	70,870.0	93,750.0	2,589

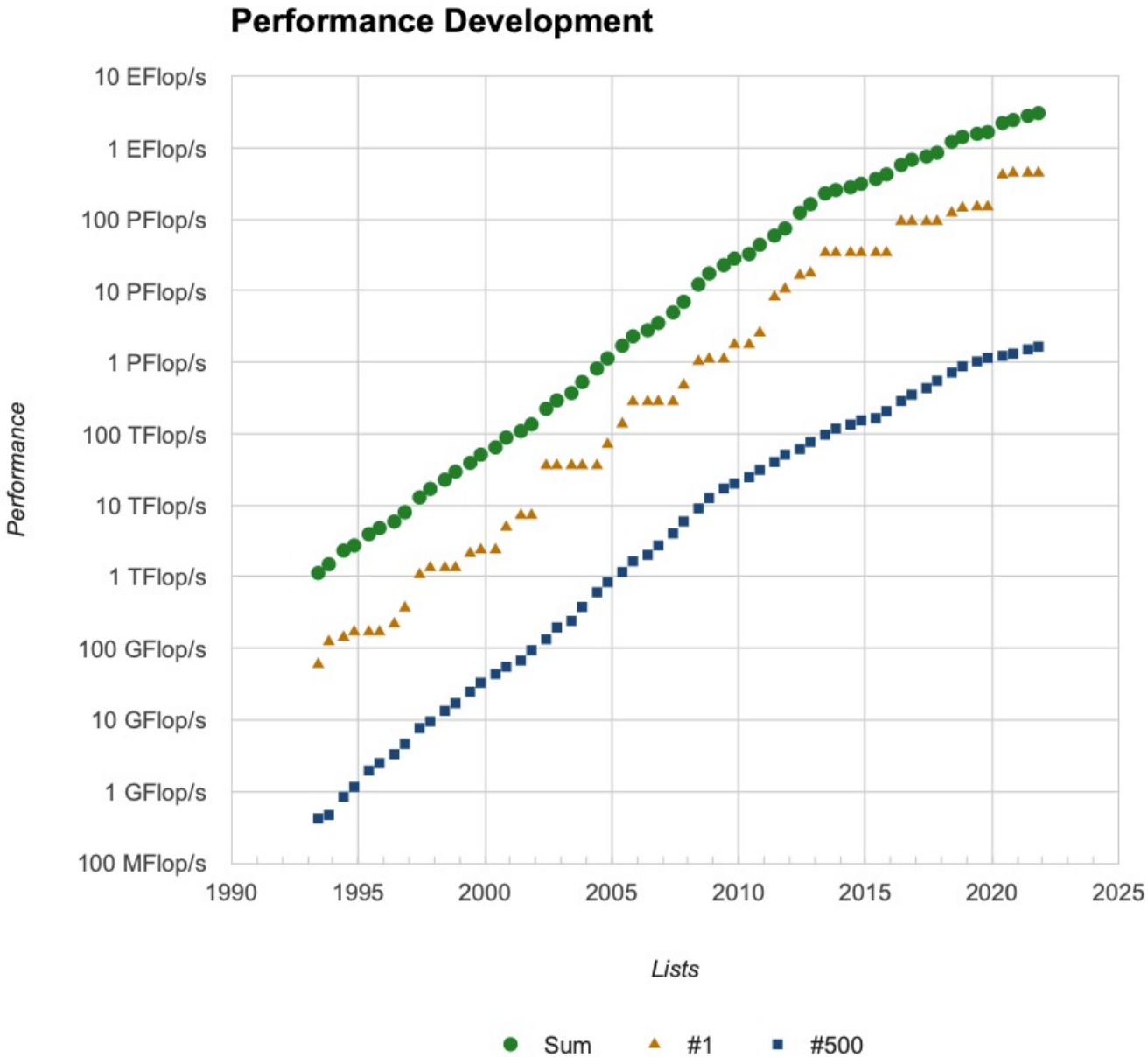
[http://www.top500.org, liste du Top500, novembre 2021]

The Green500

Rank	TOP500 Rank	System	Cores	Rmax (TFlop/s)	Power (kW)	Power Efficiency (GFlops/watts)
1	301	**MN-3** - MN-Core Server, Xeon Platinum 8260M 24C 2.4GHz, Preferred Networks MN-Core, MN-Core DirectConnect, Preferred Networks Preferred Networks Japan	1,664	2,181.2	55	39.379
2	291	**SSC-21 Scalable Module** - Apollo 6500 Gen10 plus, AMD EPYC 7543 32C 2.8GHz, NVIDIA A100 80GB, Infiniband HDR200, HPE Samsung Electronics South Korea	16,704	2,274.1	103	33.983
3	295	**Tethys** - NVIDIA DGX A100 Liquid Cooled Prototype, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100 80GB, Infiniband HDR, Nvidia NVIDIA Corporation United States	19,840	2,255.0	72	31.538
4	280	**Wilkes-3** - PowerEdge XE8545, AMD EPYC 7763 64C 2.45GHz, NVIDIA A100 80GB, Infiniband HDR200 dual rail, DELL EMC University of Cambridge United Kingdom	26,880	2,287.0	74	30.797
5	30	**HiPerGator AI** - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Infiniband HDR, Nvidia University of Florida United States	138,880	17,200.0	583	29.521

[https://www.top500.org/green500/, liste du Green500, novembre 2021]

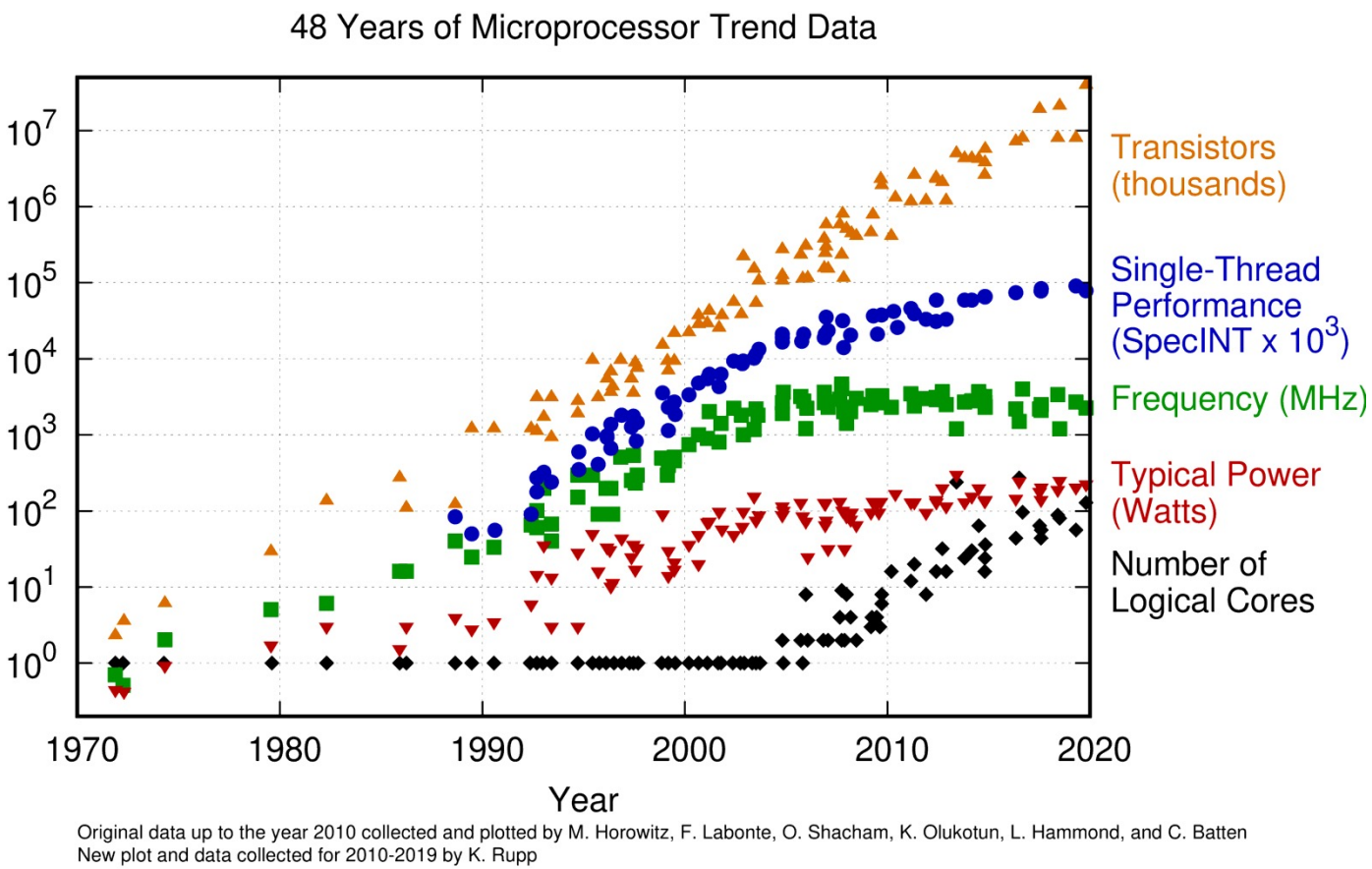
Performance development



[TOP500 https://www.top500.org/statistics/perfdevel/]

Lois empiriques

Loi de Moore : le nombre de transistors sur une puce de microprocesseur double tous les deux ans.



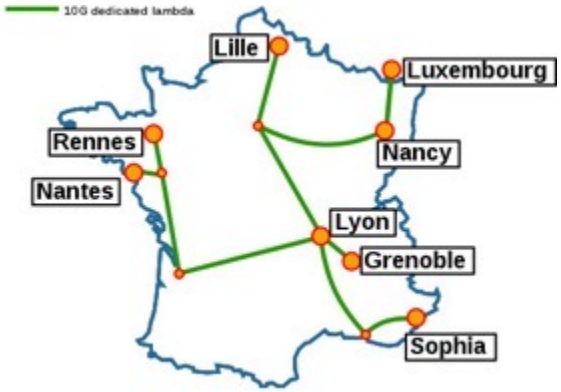
[Source : Karl Rupp, https://github.com/karlrupp/microprocessor-trend-data]

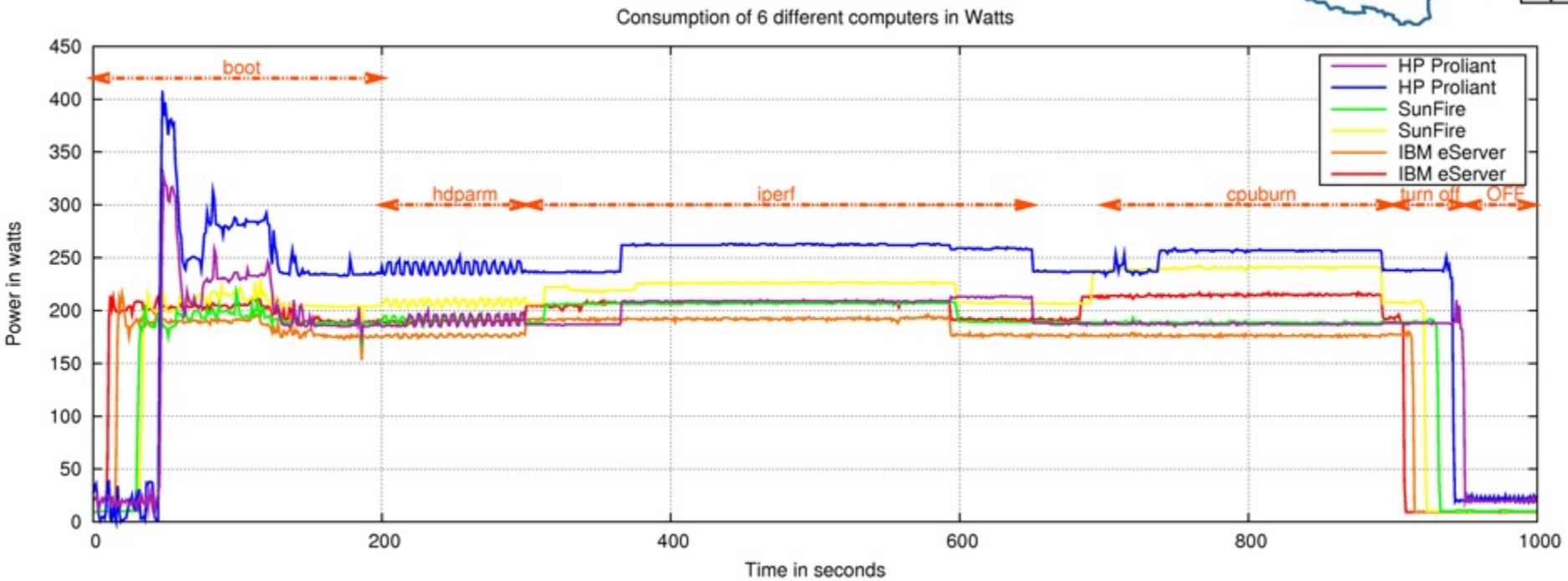
- Augmenter la fréquence du processeur
- Augmenter le nombre de cœurs par processeur
- Diminuer la finesse de gravure

Limites physiques de dissipation thermique

Combien consomment ces serveurs ?

Grid'5000 : plateforme expérimentale française
8 sites, 40 clusters, 759 serveurs, 15 882 cœurs





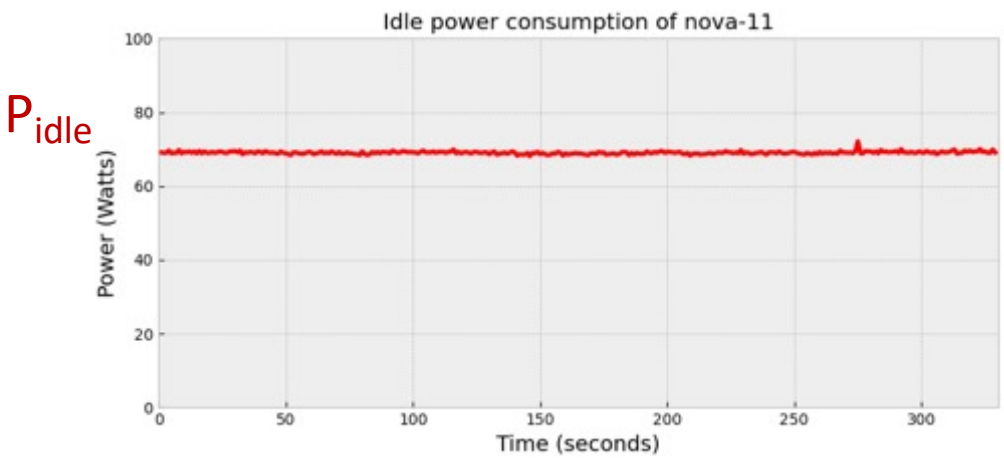
IBM eServer 325 (2005) : AMD Opteron 246 (2.0GHz, 2 CPUs), 2 Go RAM, 36 Go HDD
SunFire (2006) : AMD Opteron 250 (2.4GHz, 2 CPUs), 2 Go RAM, 74 Go HDD
HP Proliant 385 (2008) : G2 (2.2GHz, 2 CPUs dual core)

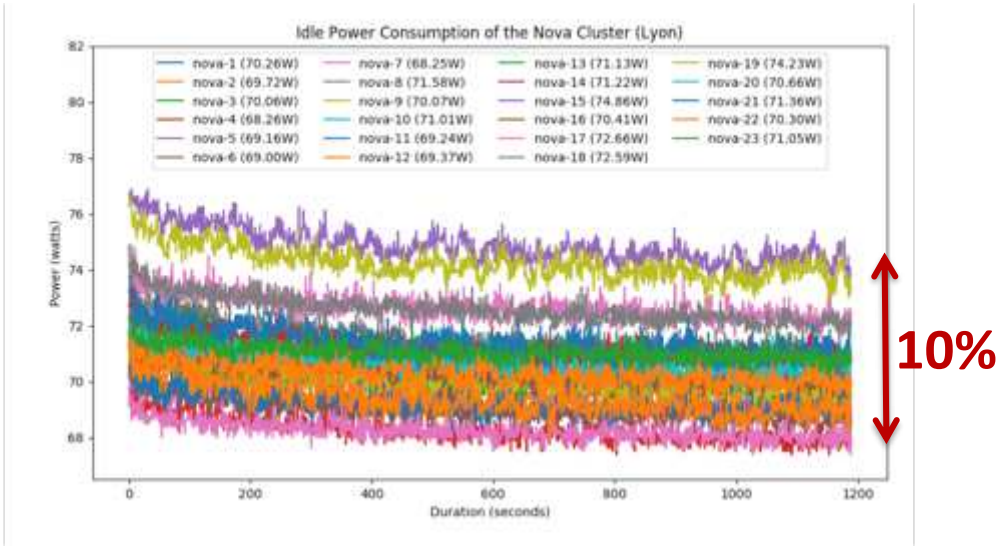
[PDCAT 2008]

Idées fausses

#1 Un serveur qui ne fait rien consomme rien ou peu.

#2 Un modèle de serveur consomme une puissance donnée.





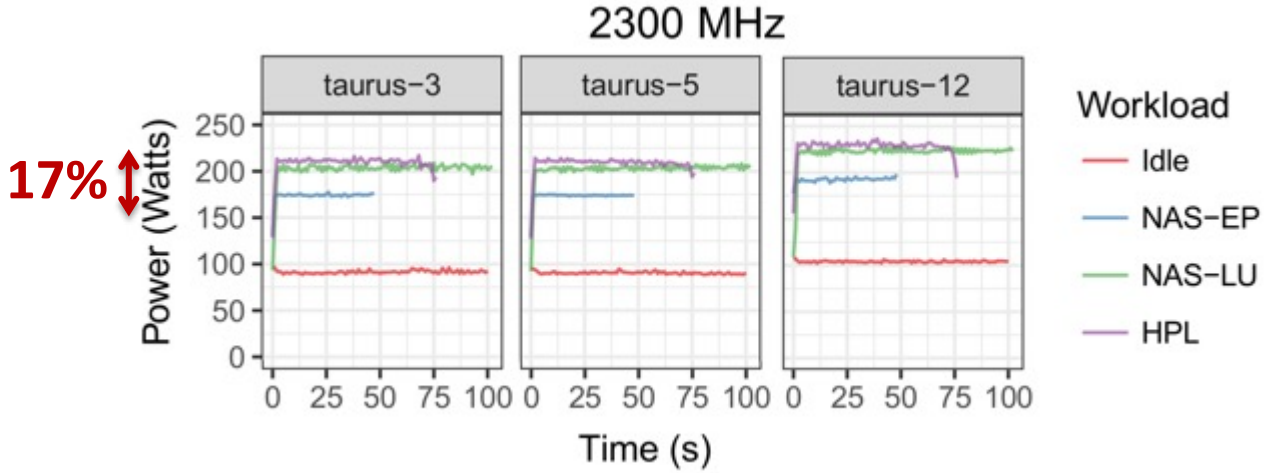
Dell PowerEdge R430 : 2 x Intel Xeon E5-2620 v4, 8 cœurs/CPU, 64 Go RAM, 598 Go HDD (2016)

70 Watts environ pendant une période d'inactivité pour ce serveur.

10% de différence de consommation en période d'inactivité et plus à plein charge.

D'autres idées fausses

#3 La consommation d'un serveur dépend du taux d'utilisation du CPU.

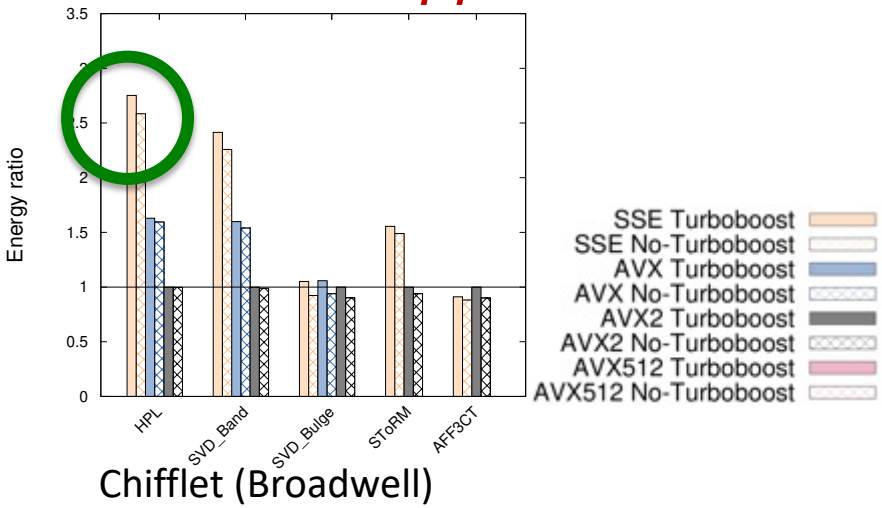


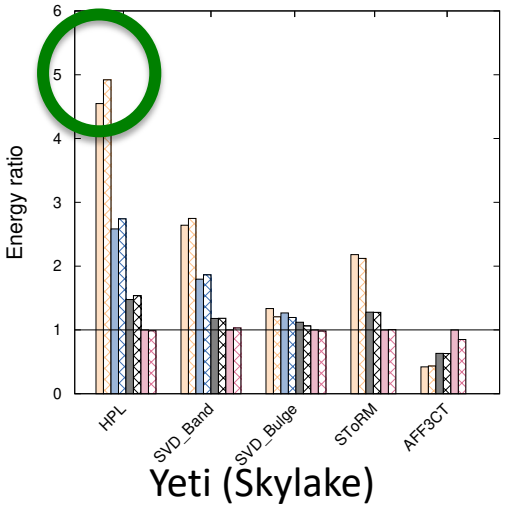
Dell PowerEdge R720 :
2 x Intel Xeon E5-2630,
6 cœurs/CPU, 32 Go RAM,
299 GB HDD (2012)

[Cluster 2017]

17% de différence de consommation pour des applications utilisant 100% du CPU.

#4 Pour une application donnée, il y a une configuration optimale.

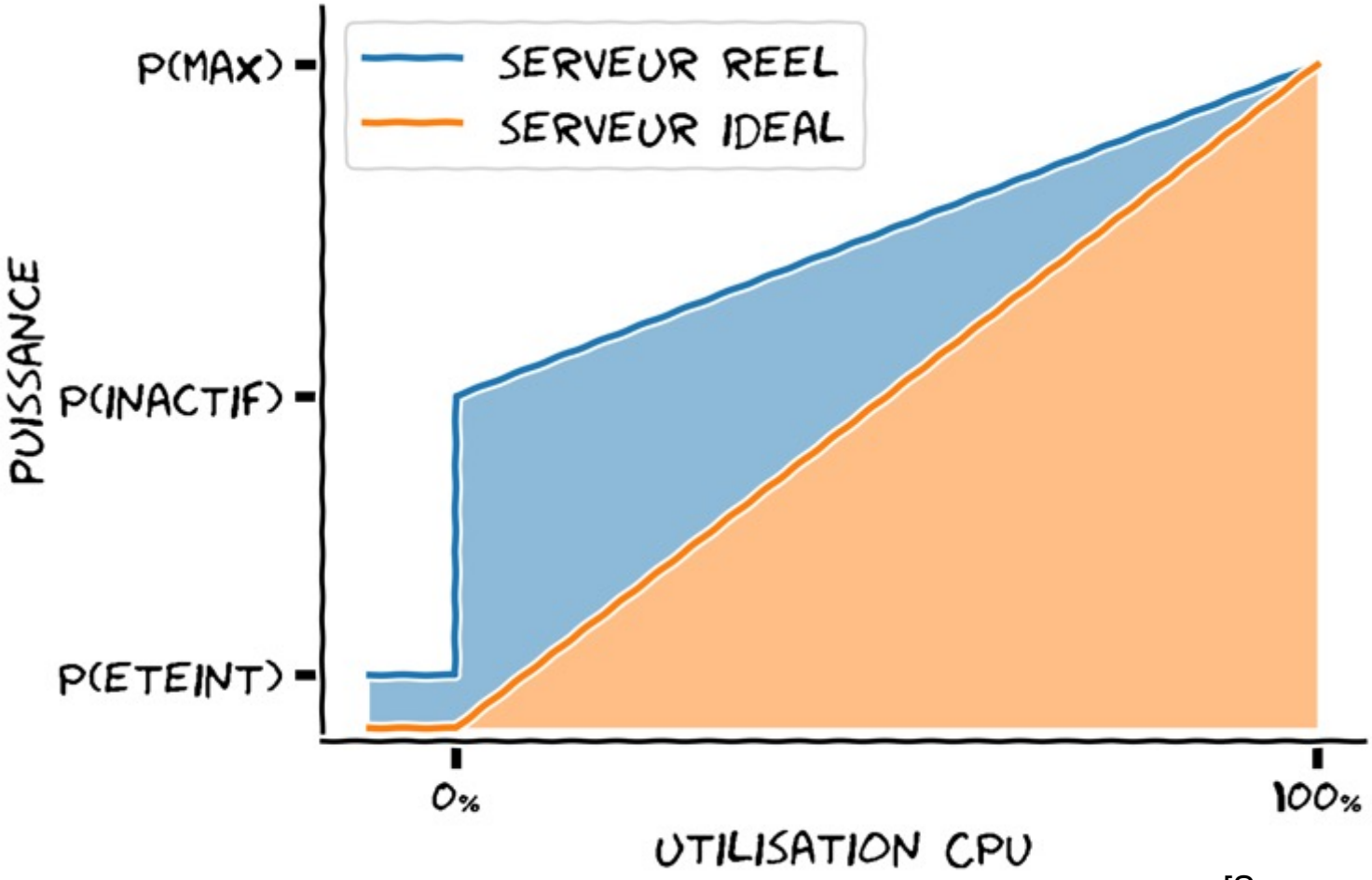




Plus rapide avec le Turboboost dans les deux cas, mais consomme plus d'énergie pour Broadwell et moins pour Skylake.

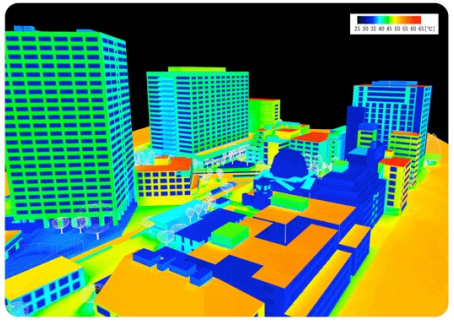
[CCPE 2021]

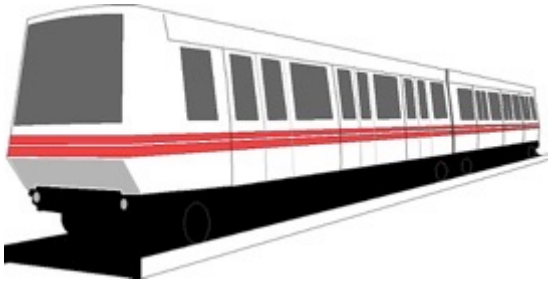
Serveur idéal vs. Serveur réel

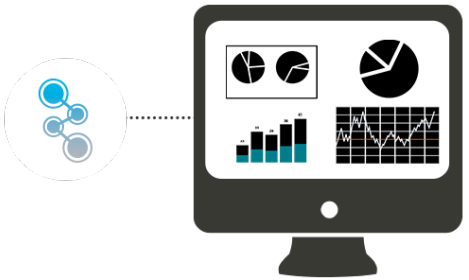
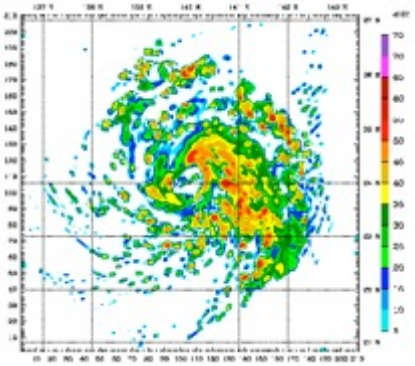


[Source : David Guyon]

Calculer au 21ème siècle?

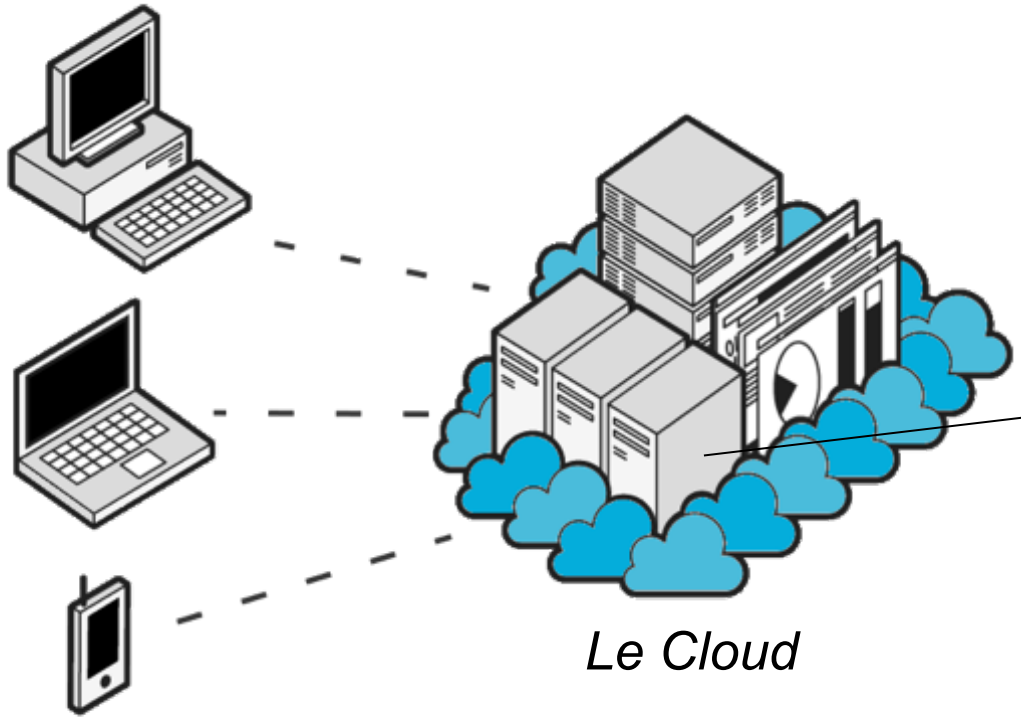




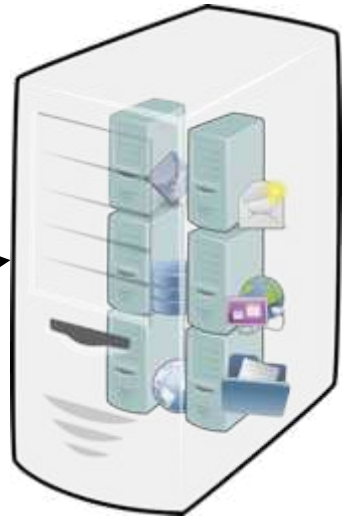



STIC : Sciences et Technologies de l'Information et de la Communication

Le Cloud : des ressources virtuelles



Le Cloud



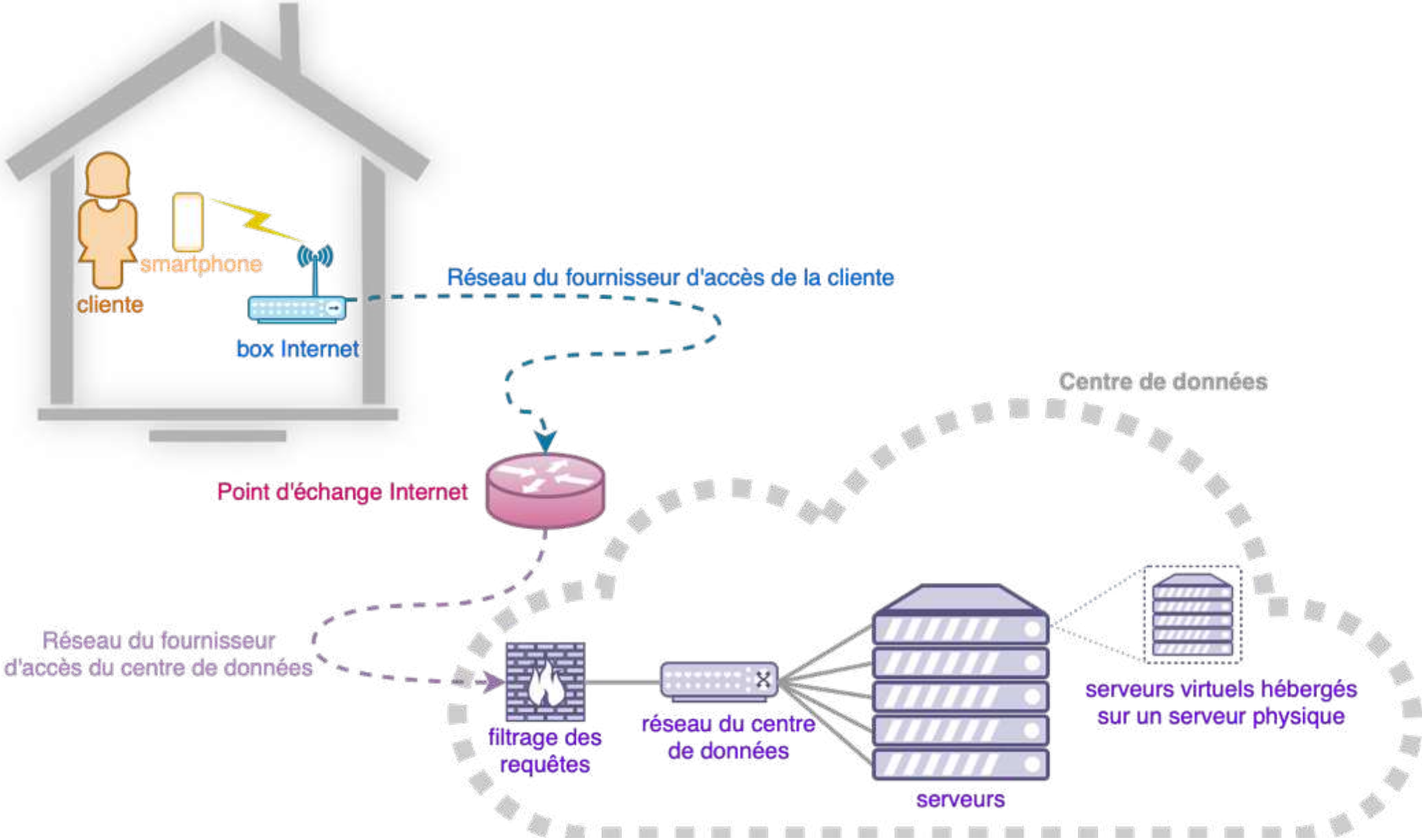
Machines virtuelles hébergées sur un serveur

Les utilisateurs

Salesforces en 2009 : 54 000 compagnies hébergées (et leurs 1,5 millions d'employés) avec 1 000 serveurs.

[Source : https://techcrunch.com/2009/03/23/the-efficient-cloud-all-of-salesforce-runs-on-only-1000-servers]

Quand utilise-t-on le cloud ?



Un Cloud

[Source : https://www.google.fr/about/datacenters]

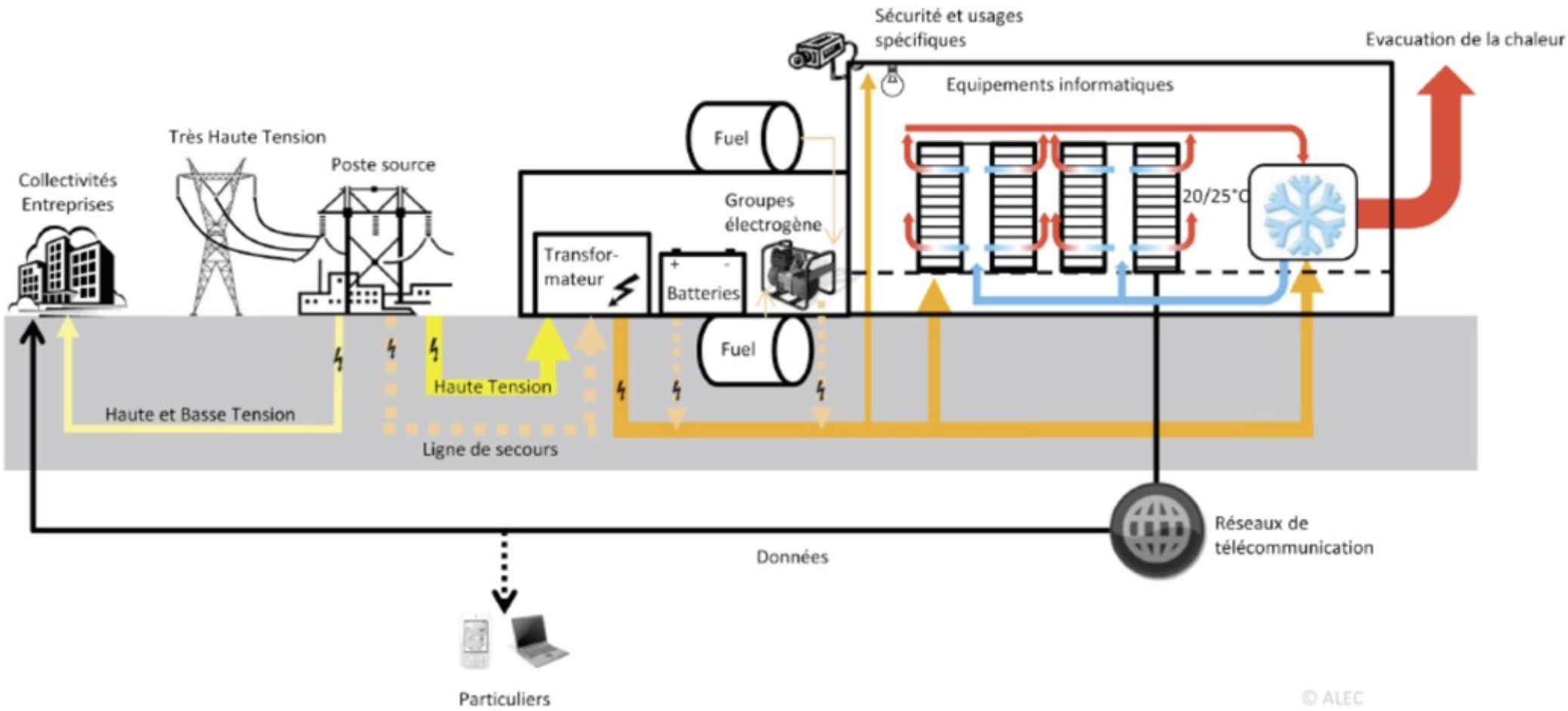
Centre de calcul de Google

Dalles, Oregon

11,5 terrains de football
100 MWatts
100 000 serveurs

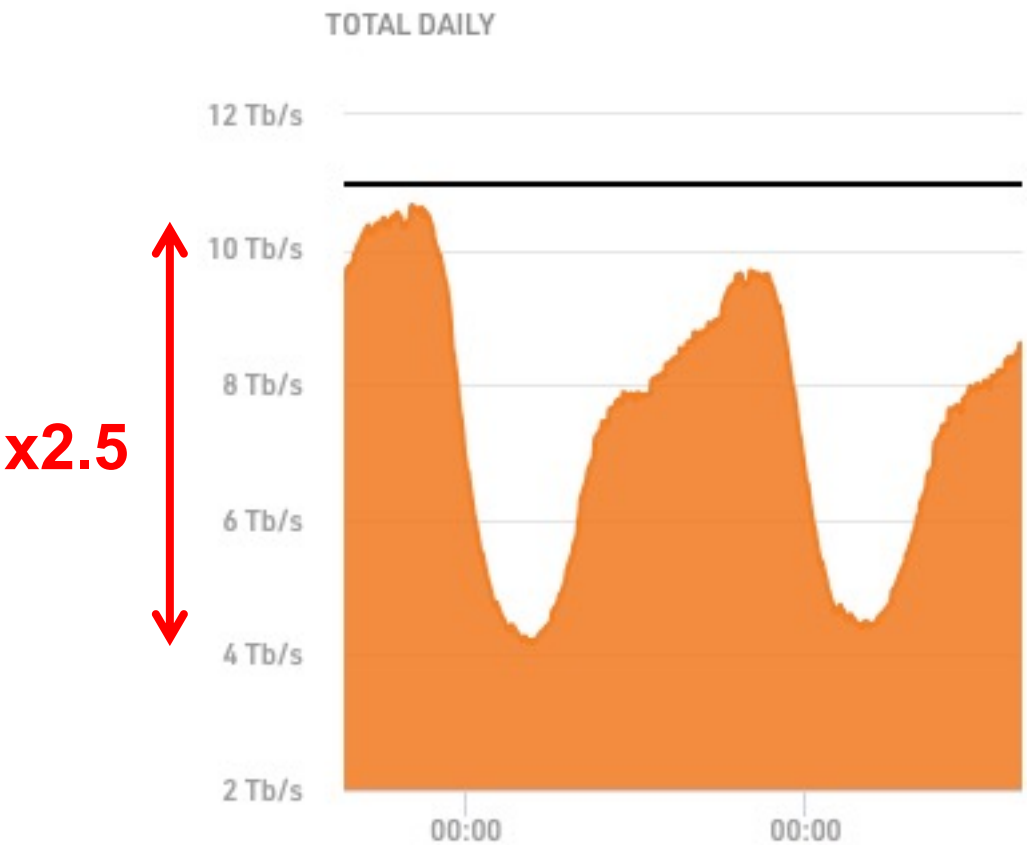
[Source : https://www.google.com/about/datacenters/inside/locations/the-dalles/]

Aperçu d'un centre de calcul



[Source : http://www.alec-plaineco.org/wp-content/uploads/2013/10/ALEC-Plaine-Commune-2013-Les-data-centers-sur-Plaine-Commune.pdf]

Dimensionnement pour les pics d'usage

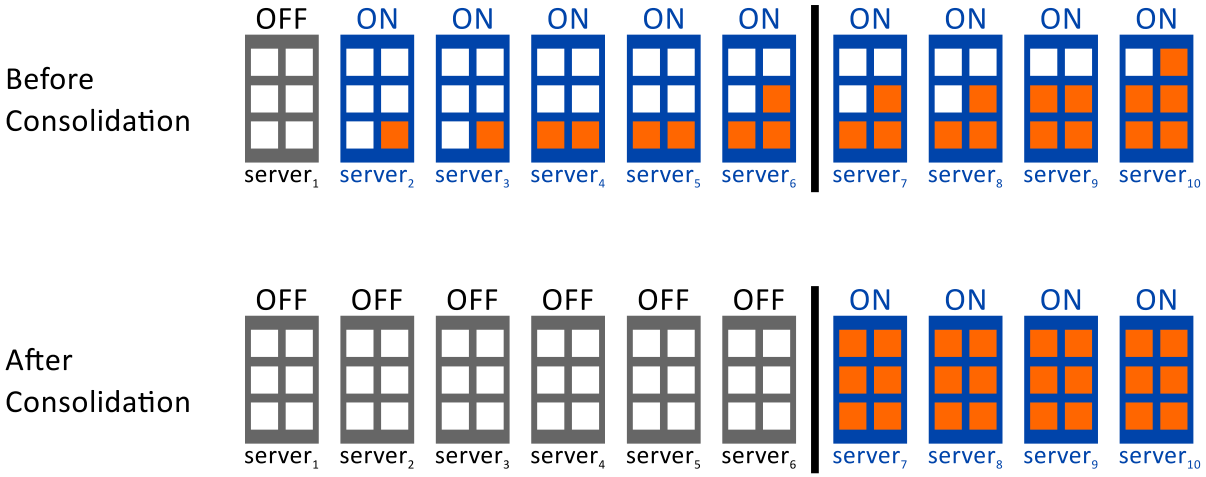


PEAK IN	PEAK OUT
10.66 Tb/s	10.679 Tb/s
AVERAGE IN	**AVERAGE OUT**
7.41 Tb/s	7.428 Tb/s
CURRENT IN	**CURRENT OUT**
8.614 Tb/s	8.638 Tb/s

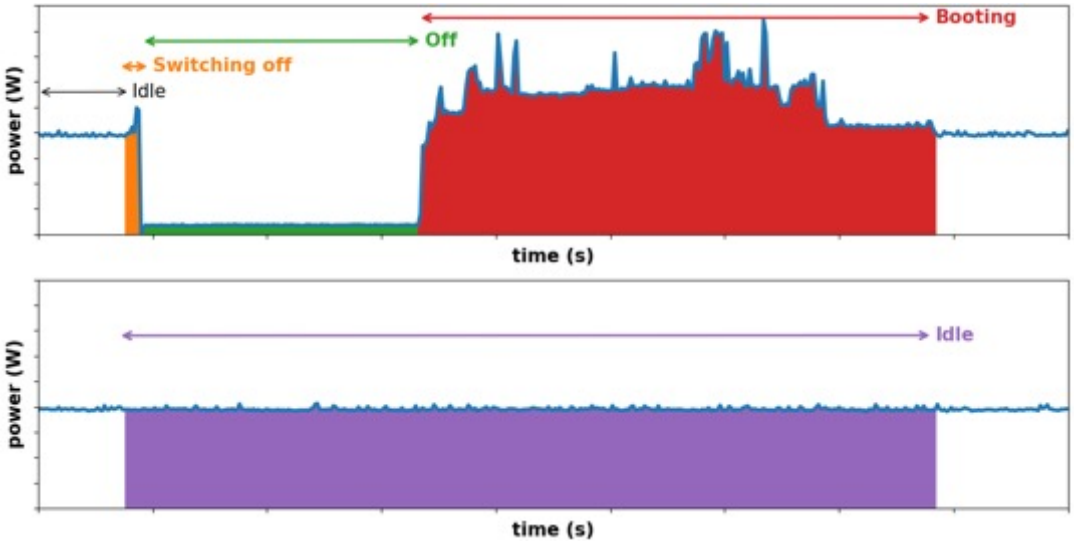
Trafic agrégé sur l'ensemble des ports réseaux du point d'échange Internet d'Amsterdam (AMS-IX) le 8 février 2022.

[Source : https://www.ams-ix.net/ams]

Stratégies d'optimisations énergétiques



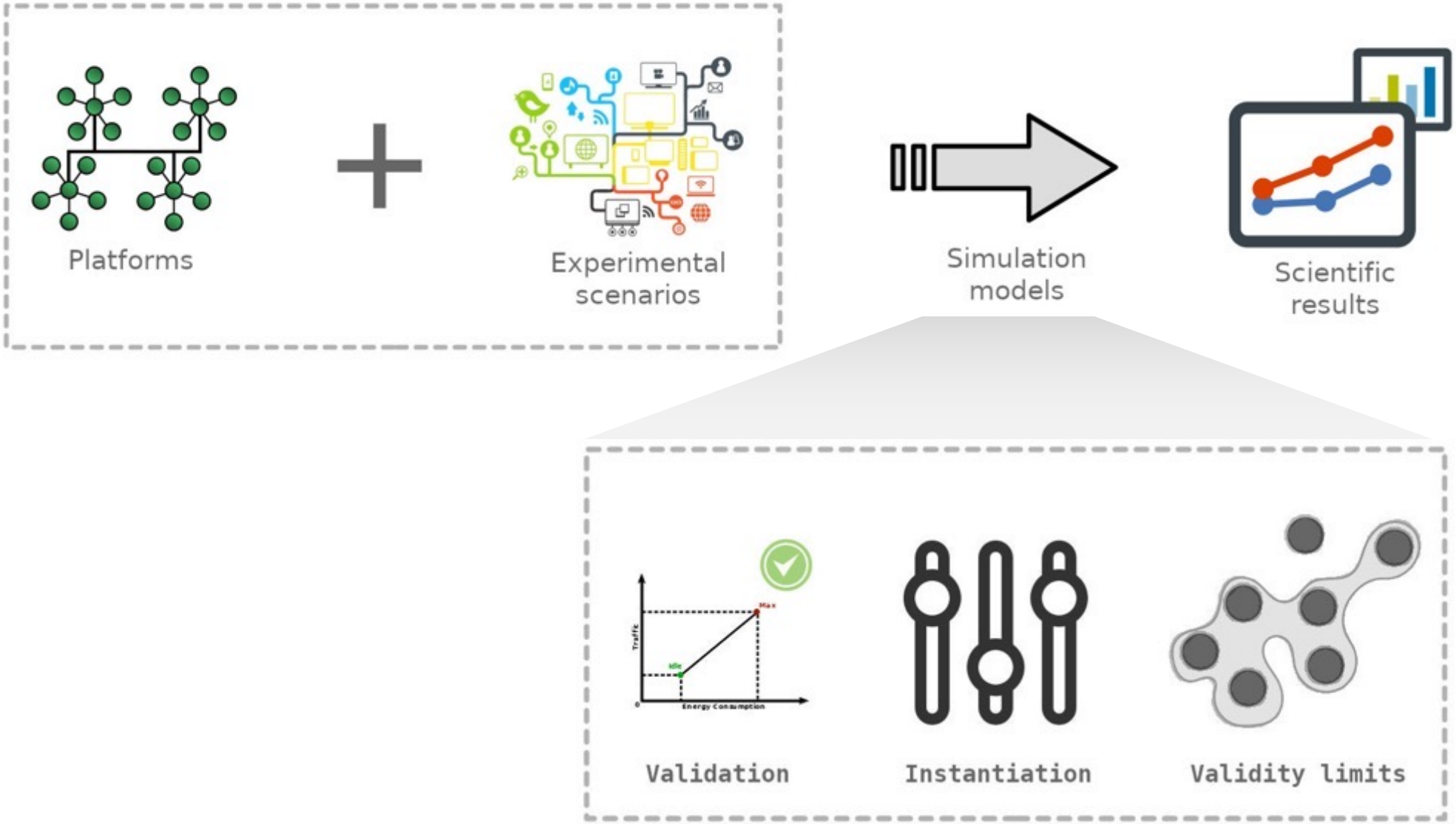
[Source : Benjamin Camus]



[CCPE 2018]

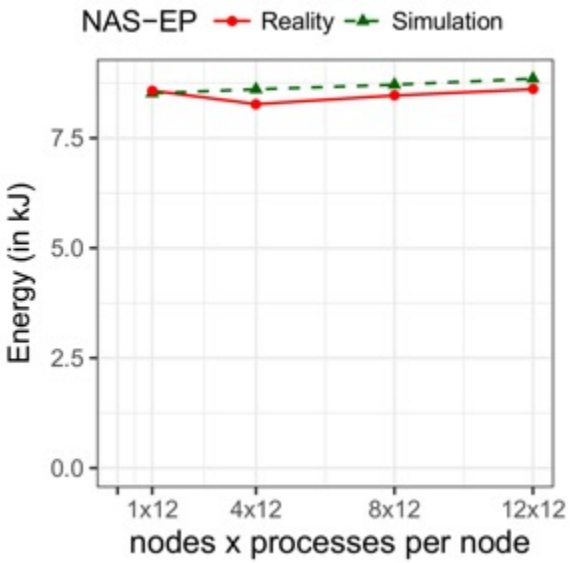
[Cluster 2018]

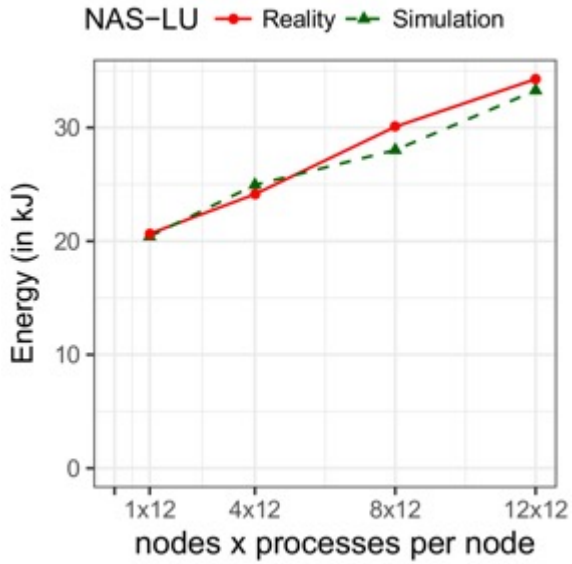
Simuler des systèmes distribués

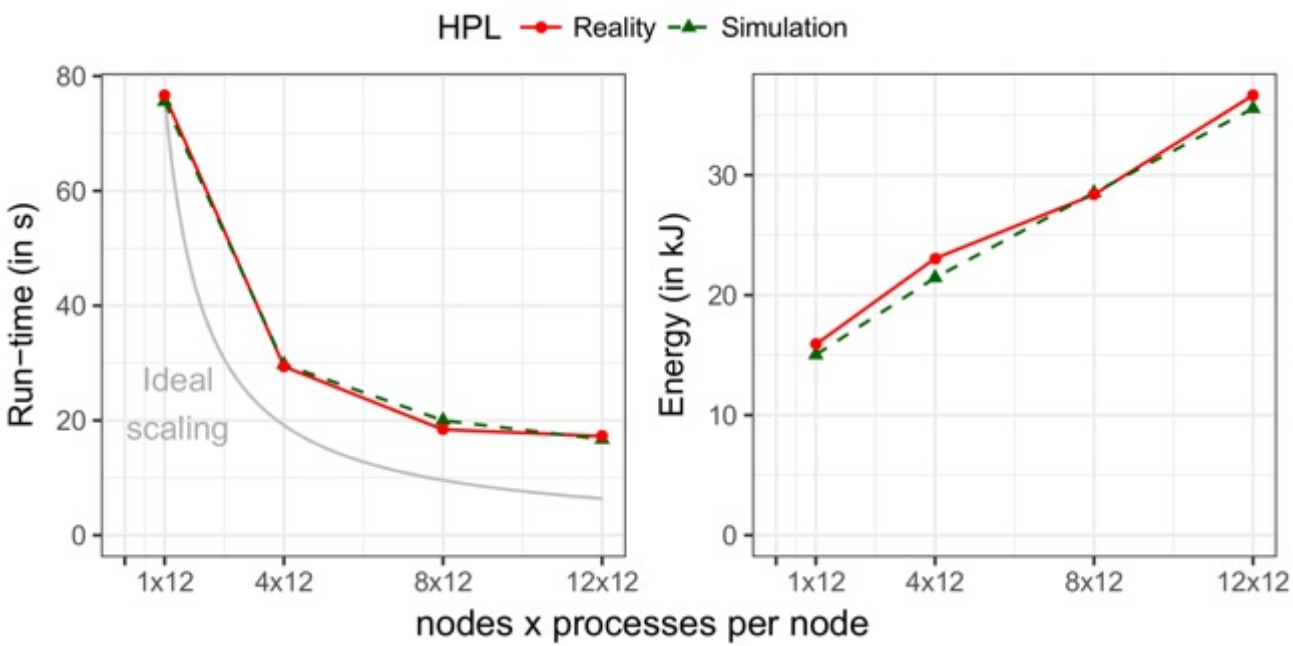


[Source : Loic Guegan]

Simuler des serveurs de calcul



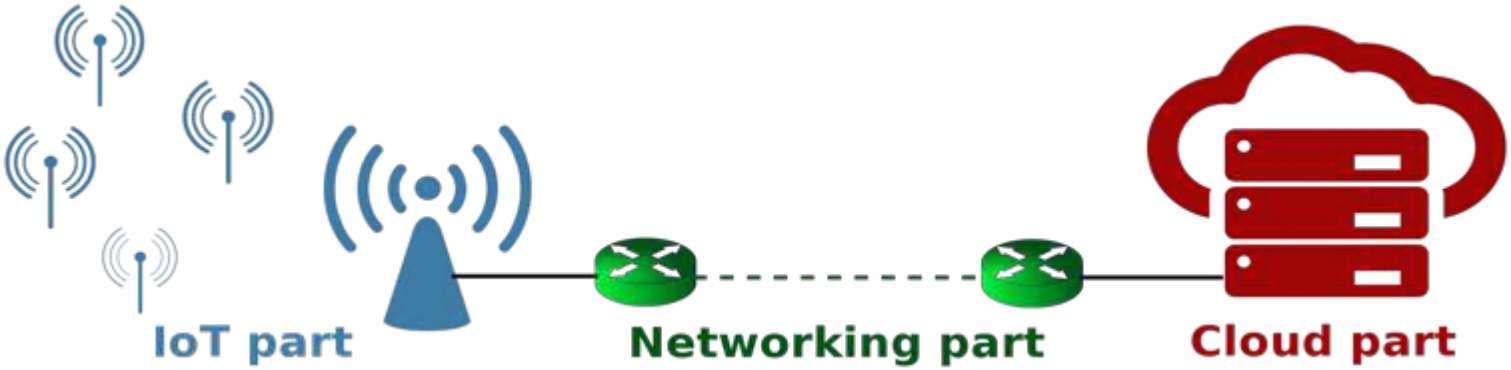




[Cluster 2017]

Résultats reproductibles : https://gitlab.inria.fr/fheinric/paper-simgrid-energy

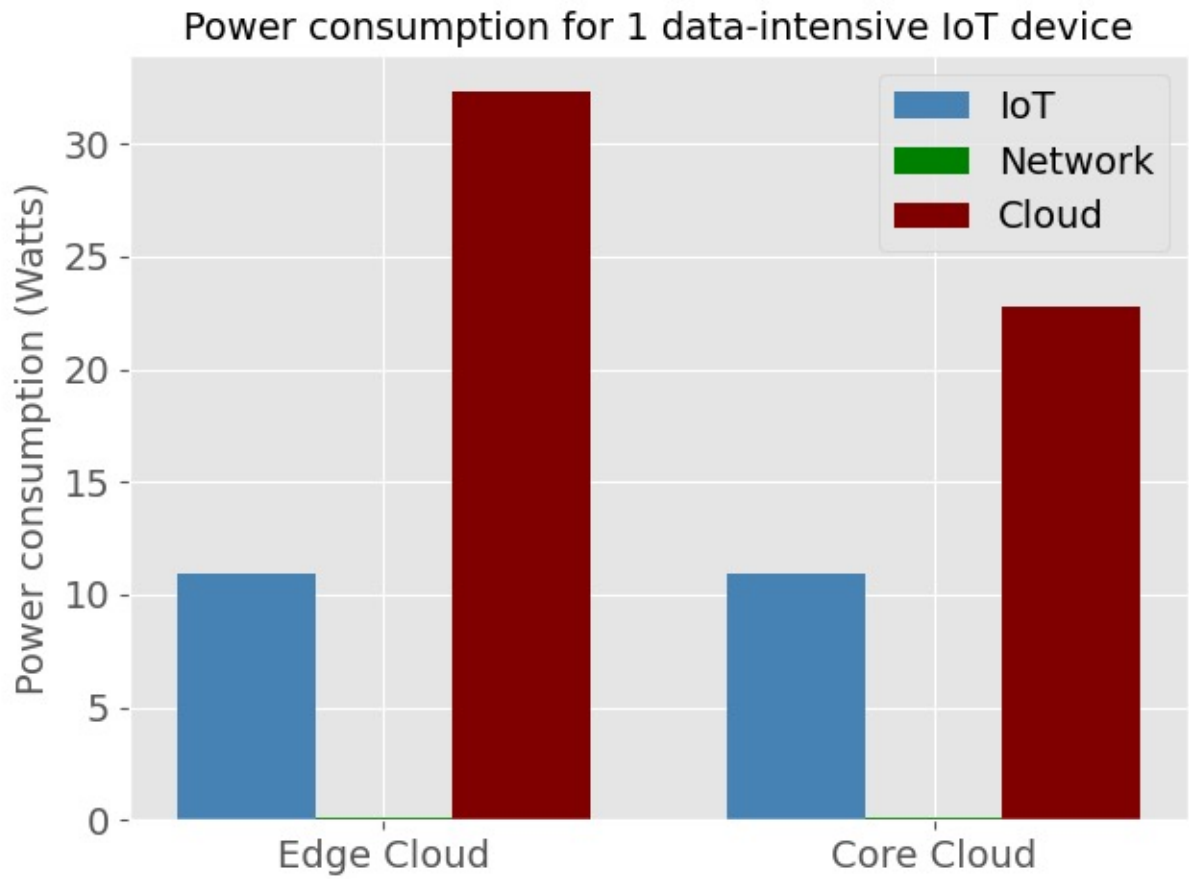
Evaluer la consummation de l'Internet des objets



THE STATE UNIVERSITY OF NEW JERSEY RUTGERS

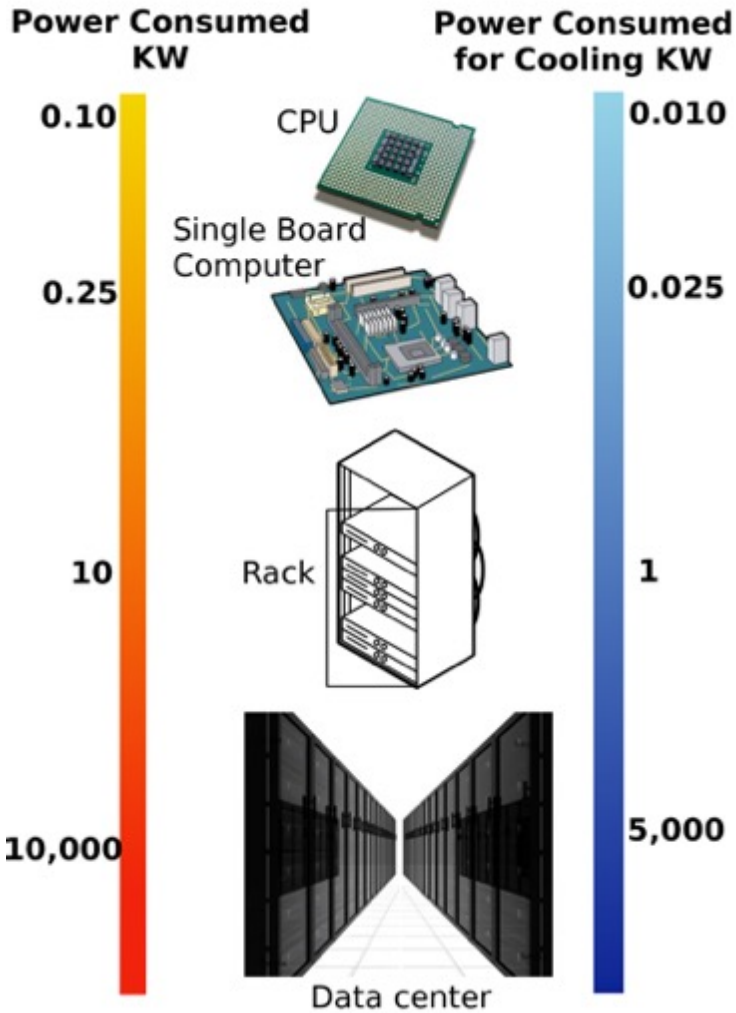
[FGCS 2018]

Compromis entre:
- Performance
- Précision de détection
- Consomation énergétique



Ça dépend.

Des économies possibles à tous niveaux



Processeurs basse-consommation
Réduction de l'énergie par opération

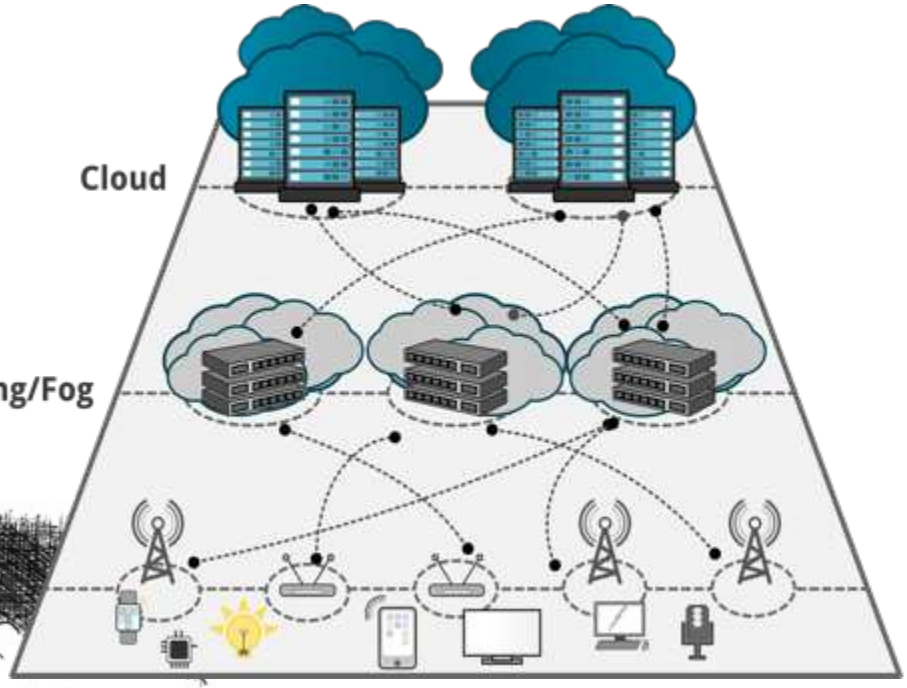
Architectures multi-cœurs
Architectures dédiées (FPGA, GPU)

Placement des données
Utilisation du parallélisme
Optimisation des communications

Gestion des points chauds
Réduction des pics de charge
Adaptation dynamique

Mais est-ce que réduire la facture électrique des centres de calculs permet réellement de réduire la consommation du numérique ?

Le monde des STIC côté utilisateur

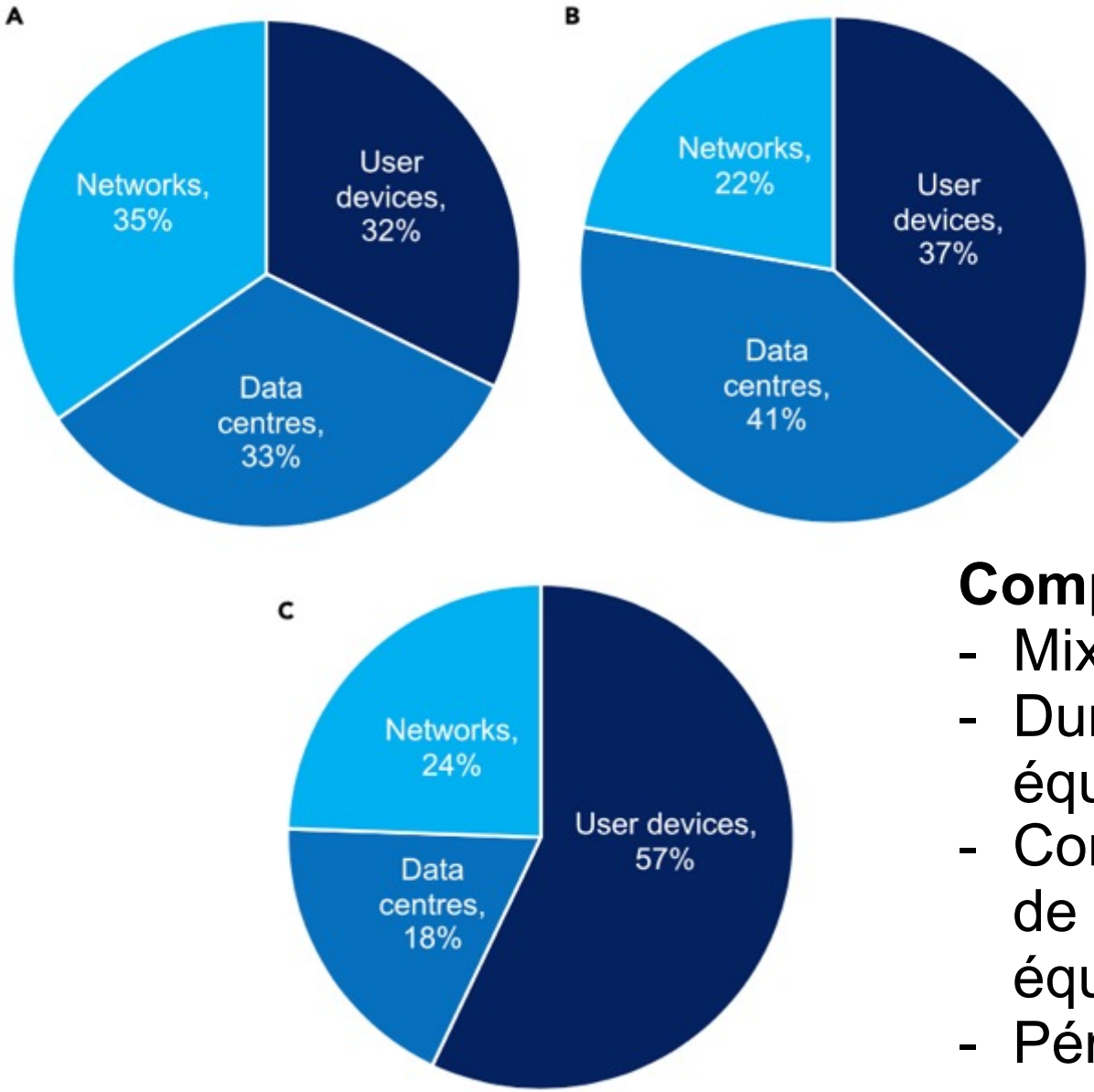


[Source : Loic Guegan]

[Crédits : Eric Drezet]

Environ 4 à 8% des émissions de gaz à effet de serre mondialement.

Répartition de l'empreinte carbone

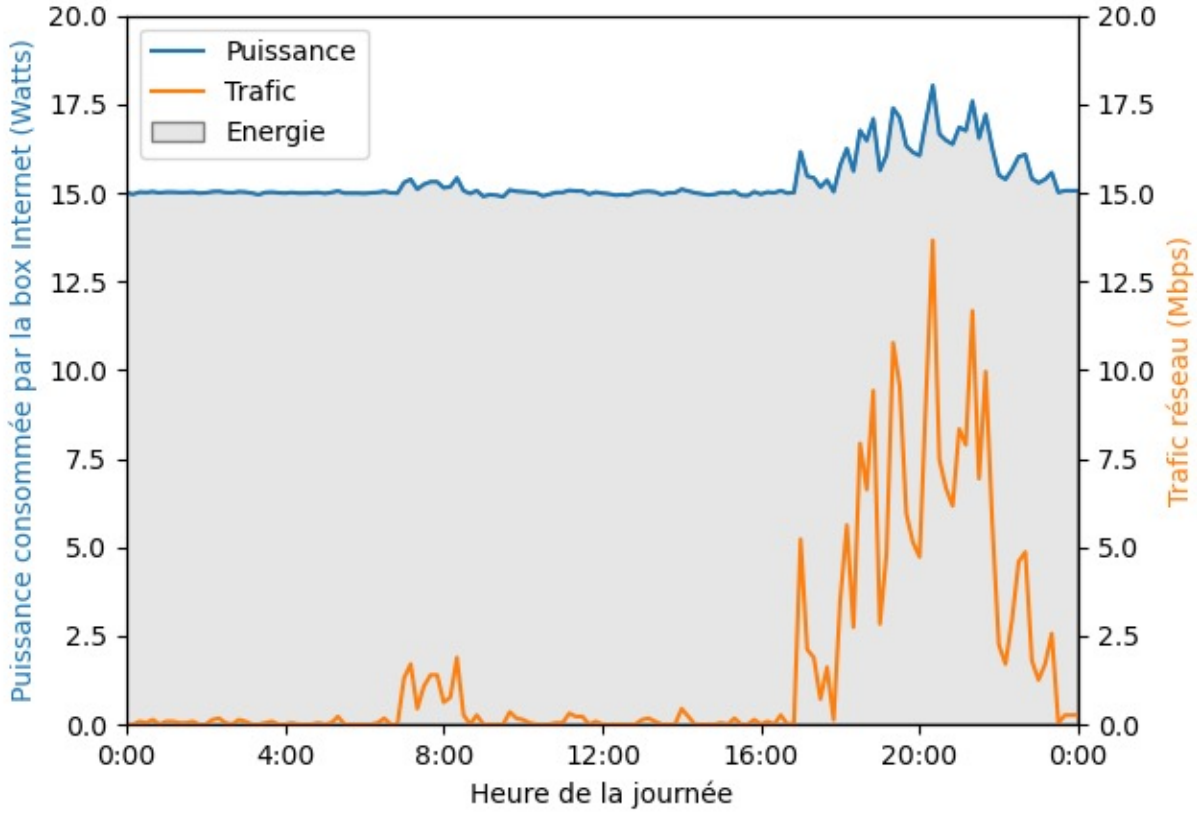


[Source : The real climate and transformative impact of ICT: A critique of estimates, trends and regulations, C. Freitag, M. Berners-Lee, K. Widdicks, B. Knowles, G. Blair, A. Friday, Patterns, Sept. 2021.]

Complexité :

- Mix électrique par pays
- Durée de vie des équipements
- Complexité des processus de fabrication des équipements
- Périmètre de l'informatique
- …

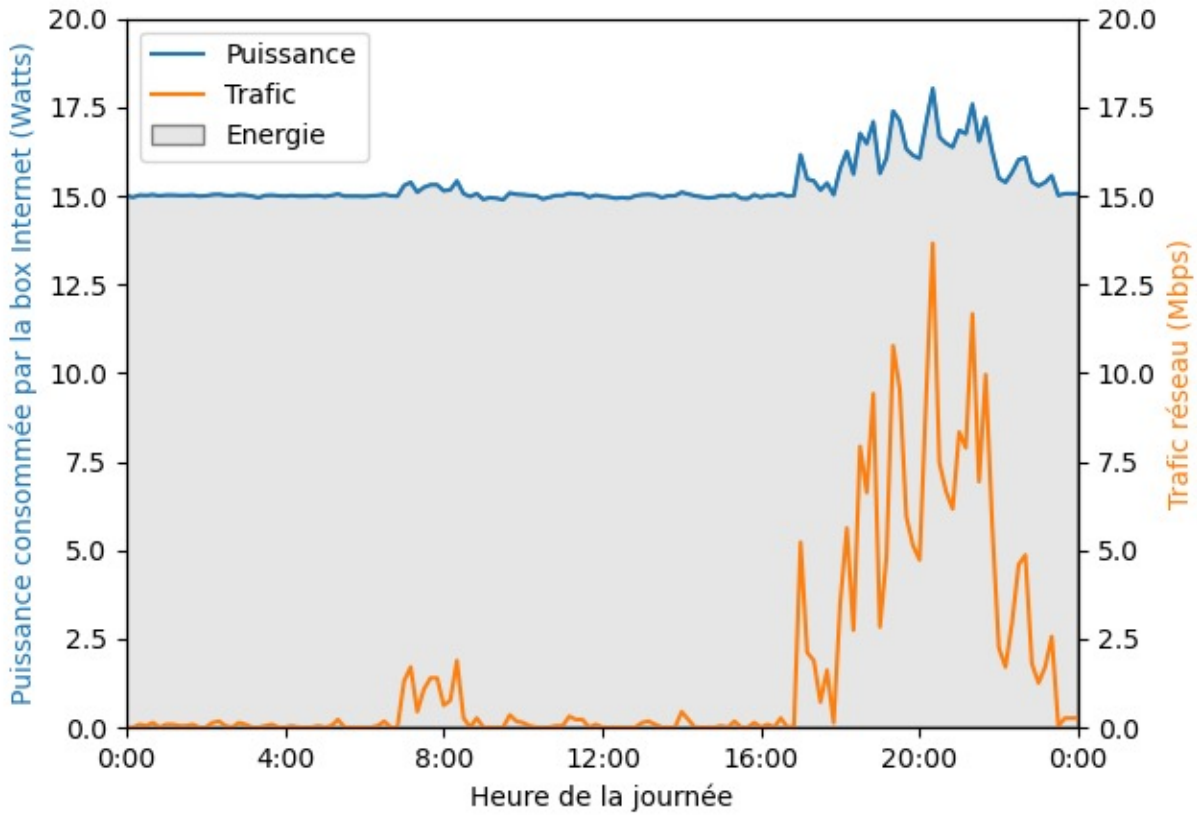
Efficacité énergétique



Efficacité énergétique en Joules/bit ?

Efficacité = énergie / volume de données

Efficacité énergétique



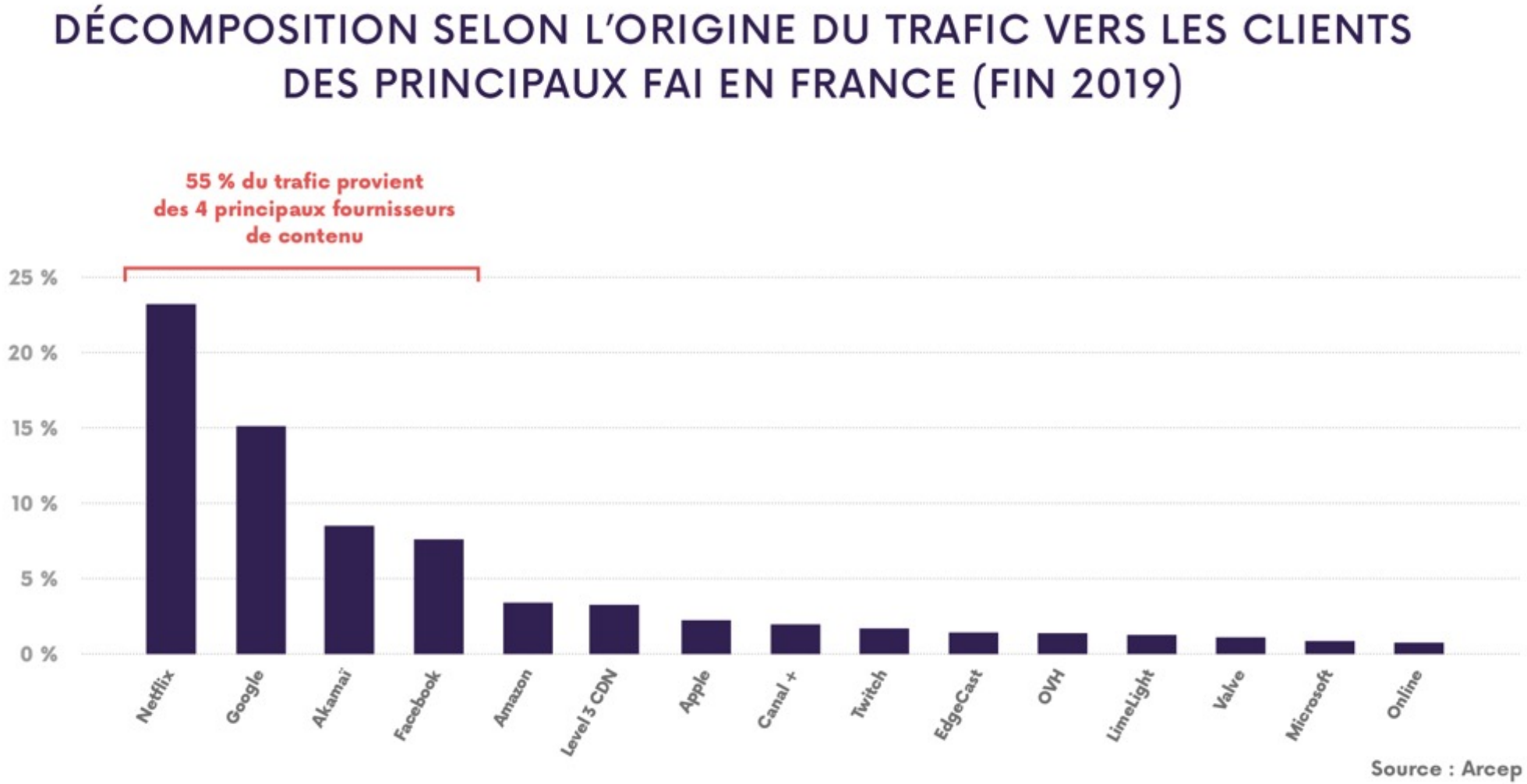
Efficacité énergétique en Joules/bit ?

Efficacité = énergie / volume de données

Métriques d'efficacité	Conditions	µJoules/ bit
Coût dynamique instantané	À 8h, 0.15 Watts 0.65 Mbps	**0.23**
Coût total instantané	À 20h, 15.15 Watts 0.65 Mbps	**23.3**
	À 8h, 16.05 Watts 4.75 Mbps	**3.38**
Coût journalier total	En moyenne, 15.35 Watts 1.55 Mbps	**9.9**
Coût minimal total	Au max, 25 Watts 60 Mbps	**0.42**

2 ordres de grandeur de différence

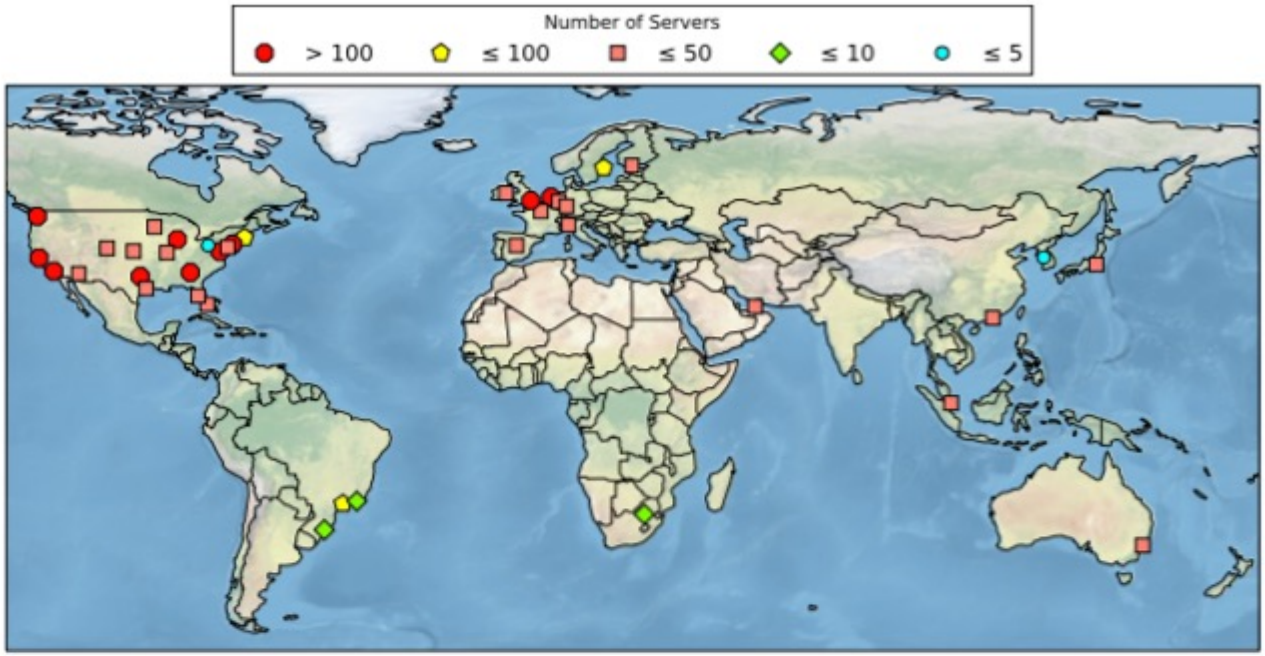
Trafic Internet en France



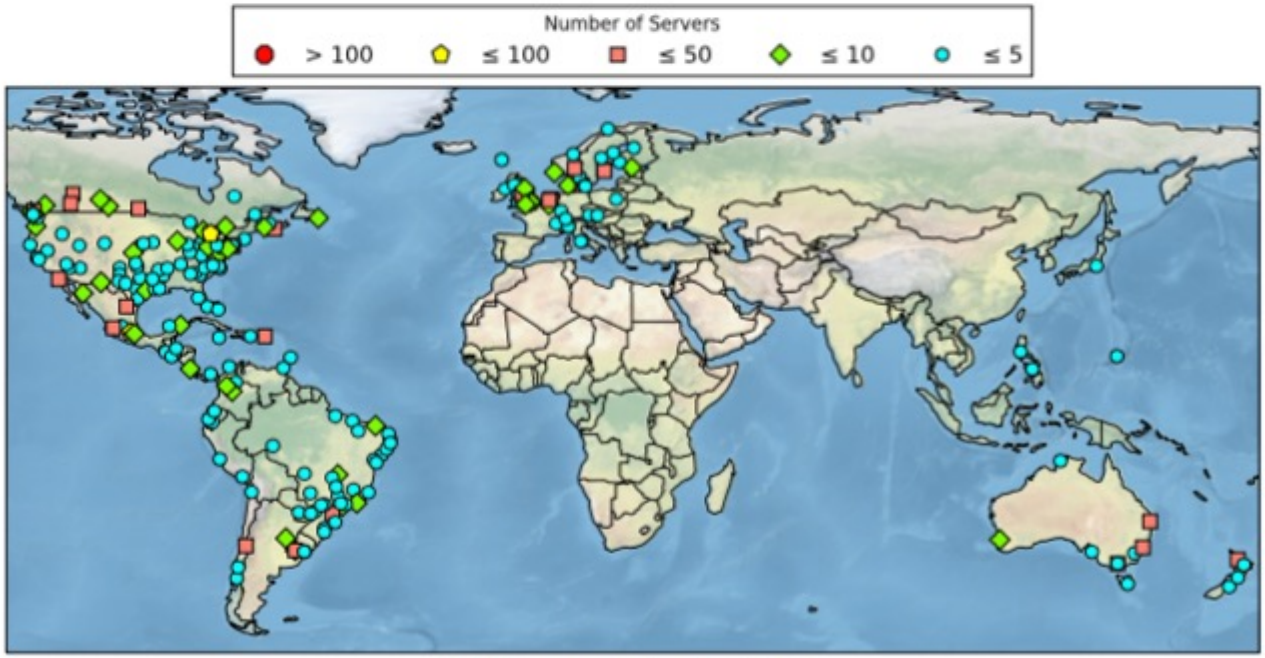
[Source : *Trafic Internet en France selon l'Arcep en 2019.*]

Ressources de Netflix

[Source : Open Connect Everywhere: A Glimpse at the Internet Ecosystem through the Lens of the Netflix CDN, T. Boettger, F. Cuadrado, G. Tyson, I. Castro, S. Uhlig, ACM SIGCOMM Computer Communication Review, 2018.]

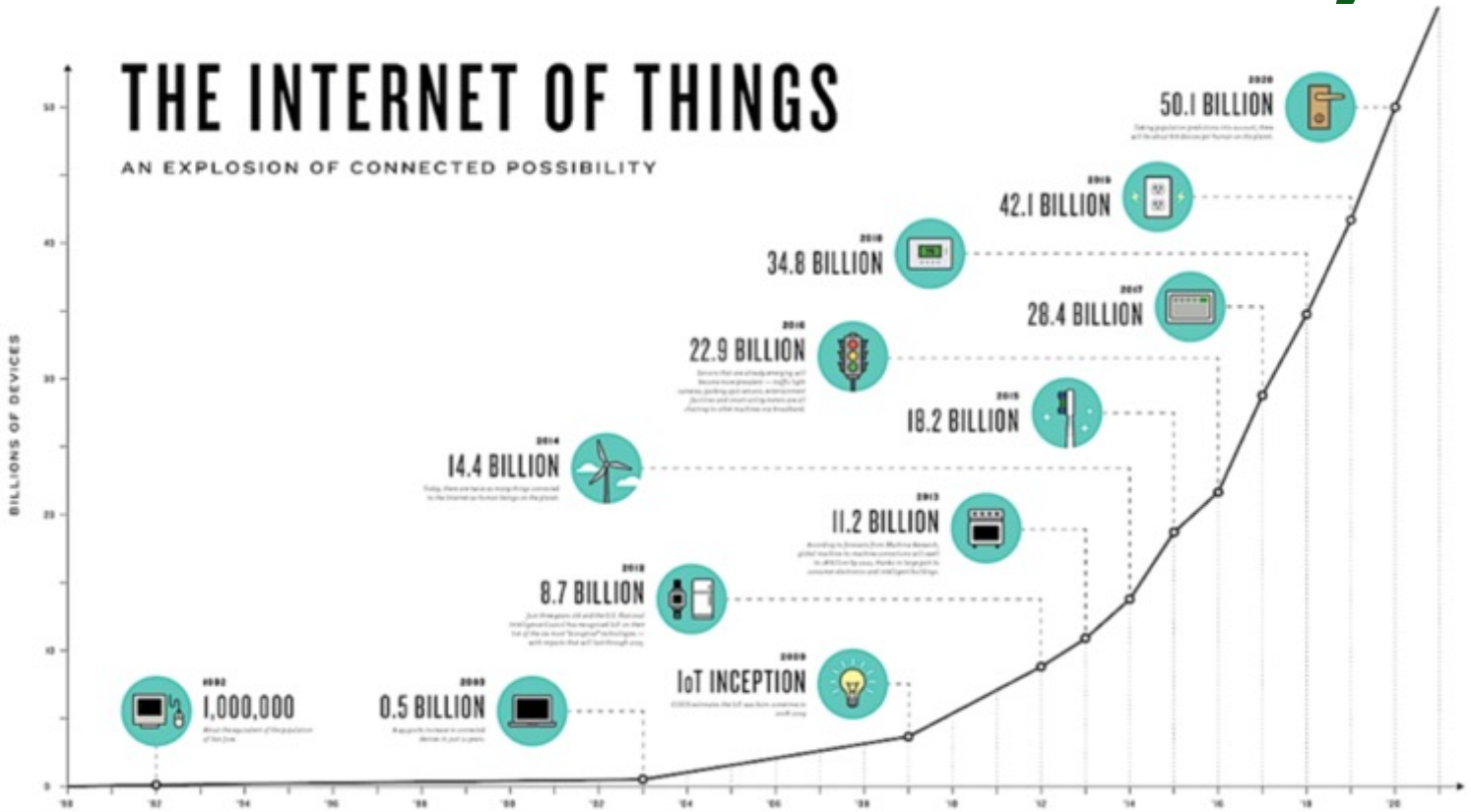


(a) CDN servers operated by Netflix at IXPs.



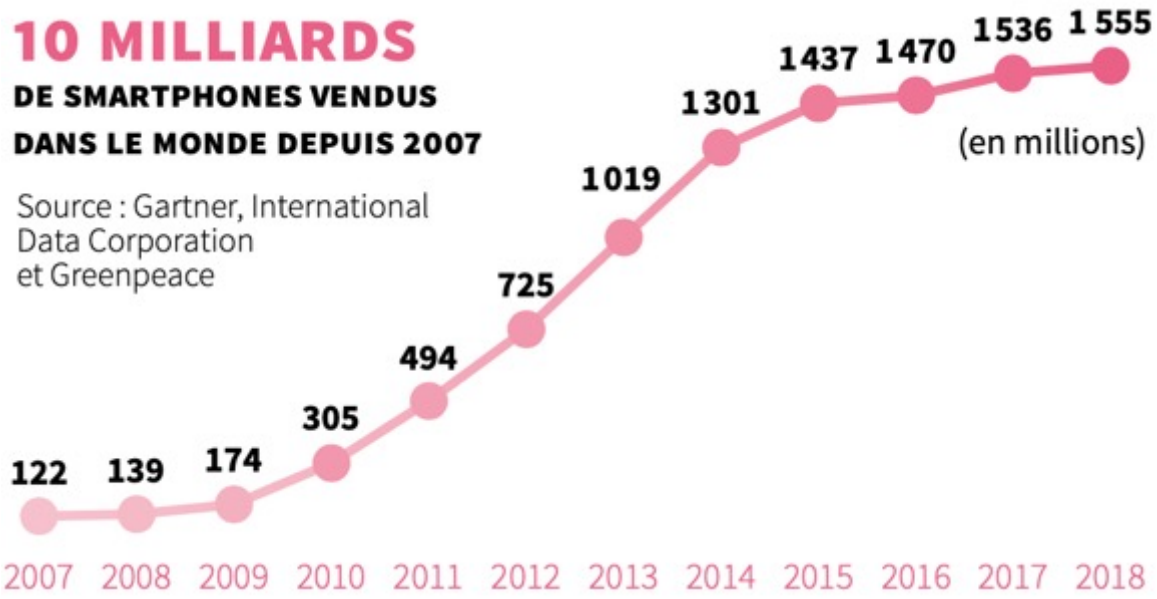
(b) CDN servers deployed within ISPs.

L'Internet des nombreux objets



[Source : https://www.mesh-net.co.uk/what-is-the-internet-of-things-iot/]

Des objets de plus en plus indispensables

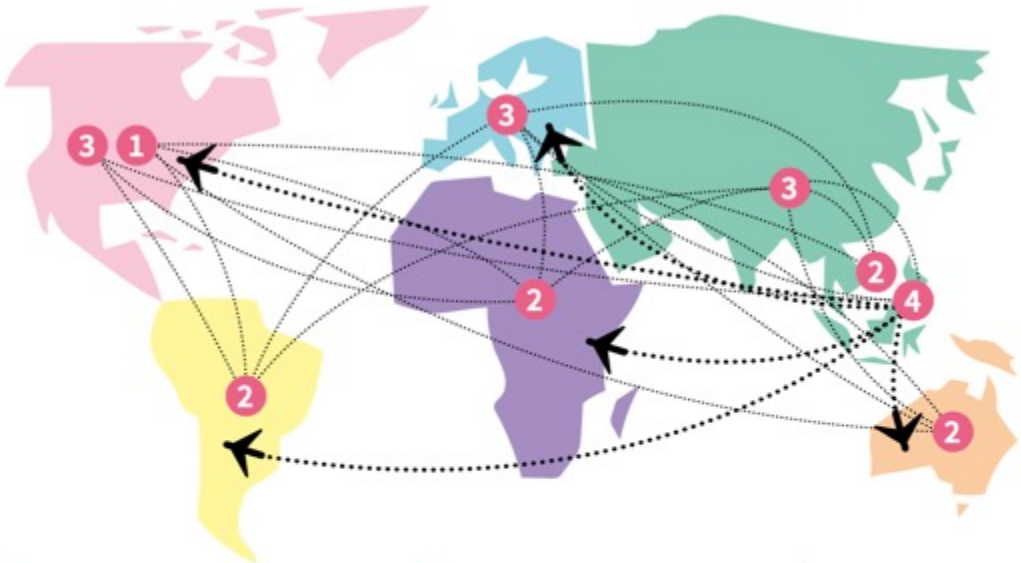


88 % DES FRANÇAIS CHANGENT LEUR TÉLÉPHONE PORTABLE ALORS QU'IL FONCTIONNE ENCORE...

[Les impacts du smartphone, ADEME, 2019]

Une fabrication complexe

QUATRE TOURS DU MONDE POUR FABRIQUER UN SMARTPHONE



1. **Conception** le plus souvent aux États-Unis

2. **Extraction et transformation des matières premières** en Asie du Sud-Est, en Australie, en Afrique centrale et en Amérique du Sud

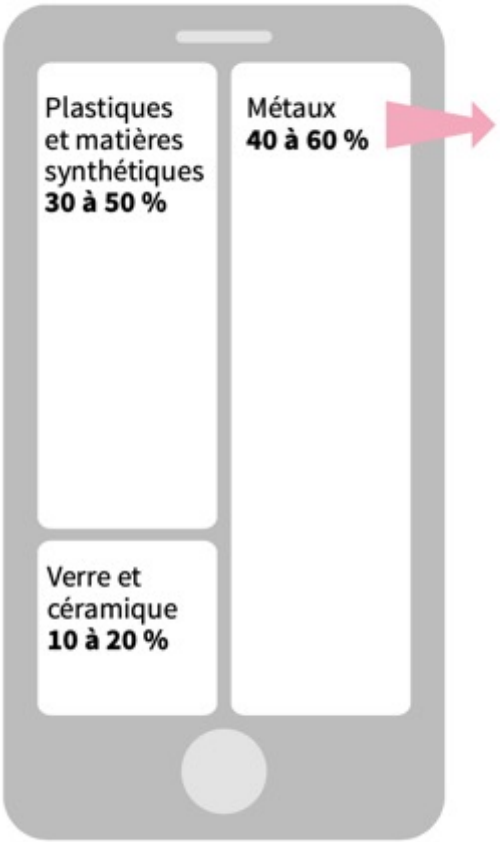
3. **Fabrication des principaux composants** en Asie, aux États-Unis et en Europe

4. **Assemblage** en Asie du Sud-Est

Distribution vers le reste du monde, souvent en avion.

70 MATÉRIAUX POUR FABRIQUER UN SMARTPHONE

RÉPARTITION DU POIDS DES MATÉRIAUX DANS LA COMPOSITION D'UN SMARTPHONE



PROPORTION DES MÉTAUX

80 à 85 % de métaux ferreux et non ferreux : cuivre, aluminium, zinc, étain, chrome, nickel…

0,5 % de métaux précieux : or, argent, platine, palladium…

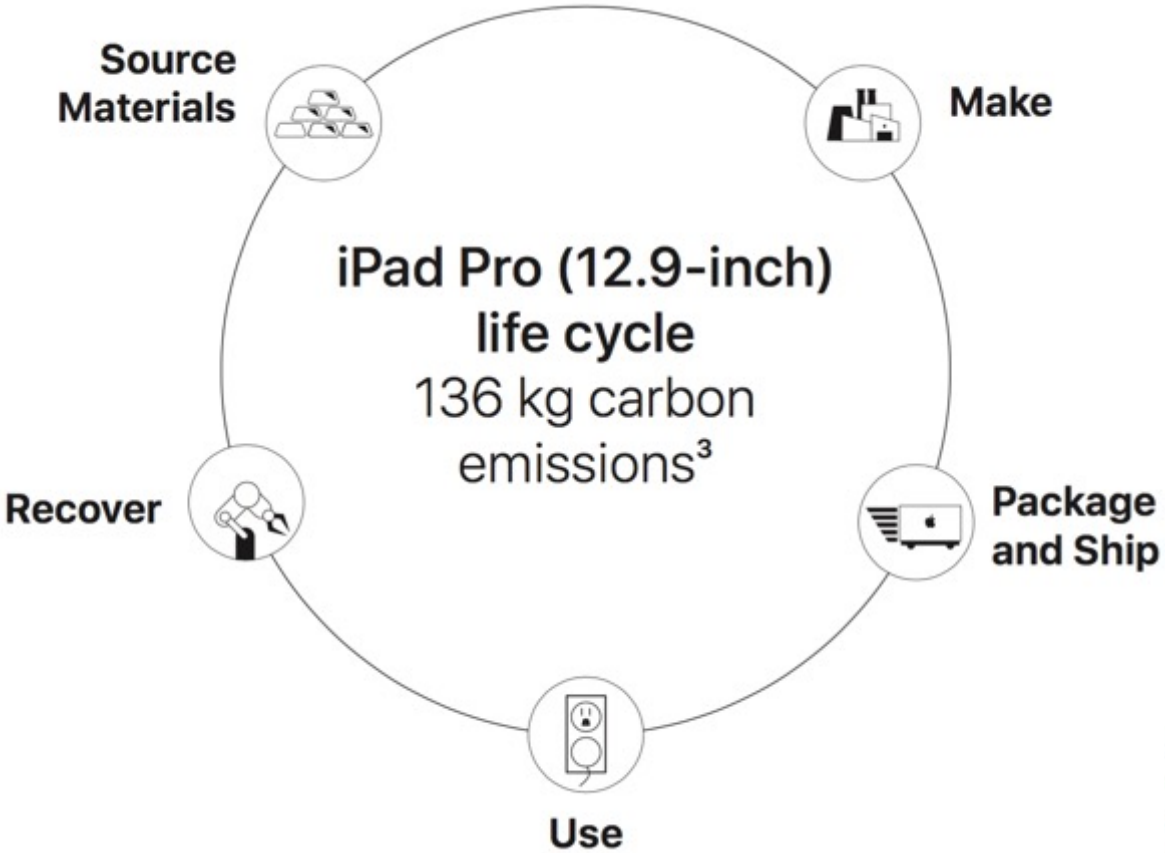
0,1 % de terres rares et métaux spéciaux : europium, yttrium, terbium, gallium, tungstène, indium, tantale…

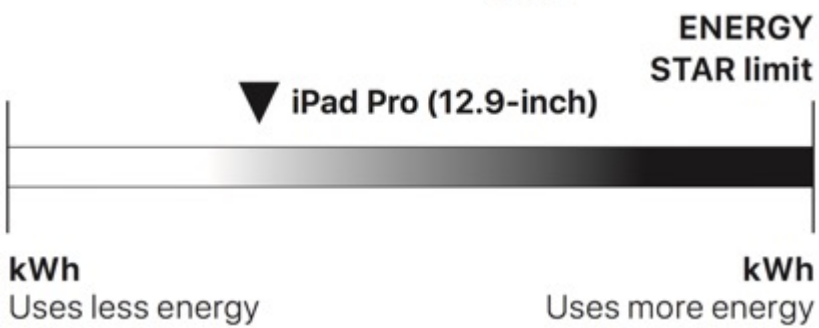
15 à 20 % d'autres substances : magnésium, carbone, cobalt, lithium…

Source : Oeko-Institut, EcoInfo et Sénat

[Les impacts du smartphone, ADEME, 2019]

La phase usage : pointe de l'iceberg





iPad Pro (12.9-inch) life cycle carbon emissions

83%	Production
11%	Transport
6%	Use
<1%	End-of-life processing

Pour 3 ans d'utilisation

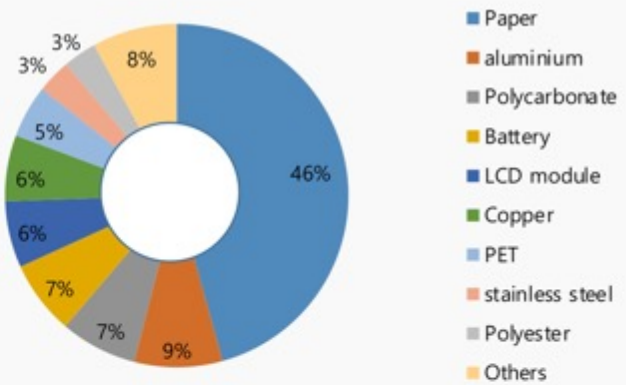
[Source : Product environmental report, Apple, 2018.]

Beaucoup d'autres impacts

Product Features

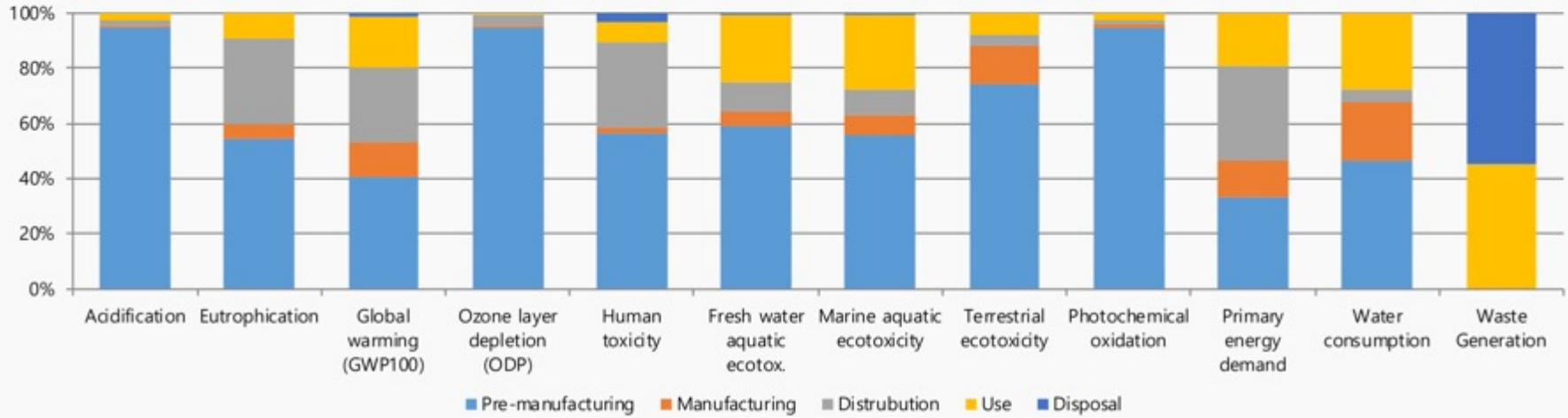
Model name	SM-N950U (Galaxy Note8)
Processor	Qualcomm 2.35GHz, 1.9GHz Octa-Core 64bit
Dimension	162.5 x 74.8 x 8.6 mm
Display	6.3" 2960 x 1440, 16M In-Cell Touch LCD
Battery	Li-Ion 3300 mAh
Camera	12 MP / 5MP
Wt.(g)	186.34g

Material Use



Characterized Environment Impact

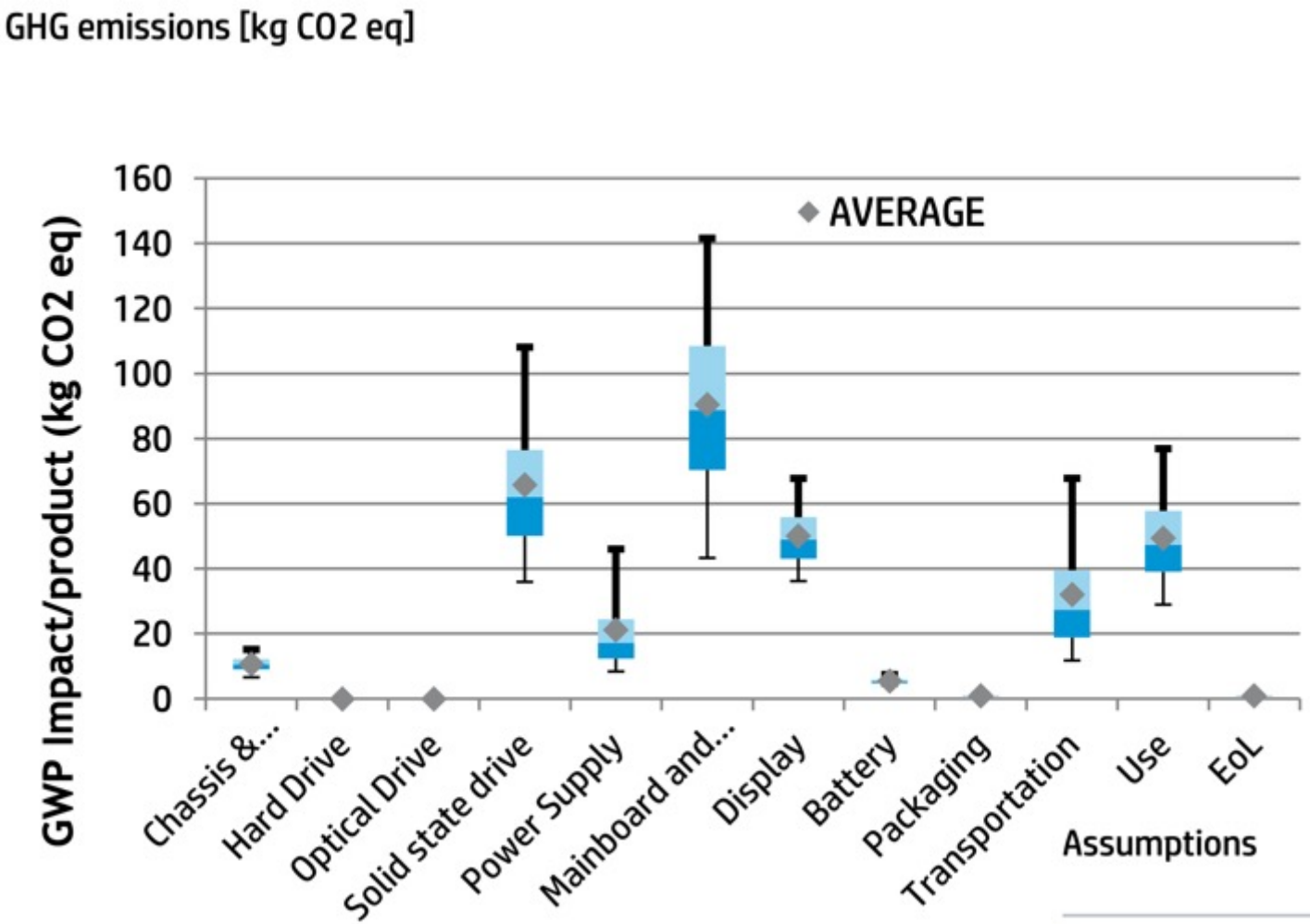
[Source : Life Cycle Assessment for Mobile Products, Samsung, 2018.]



Standard	ISO 14040:2006 and 14044:2006
Database	Ecoinvent 2.2
Method for impact assessment	Life cycle impact assessment classification and characterization factors according to CML 2001 as provided in the SimaPro 7.1.5 LCA tool
LCA software	SimaPro 7.1.5

Pre-manufacturing	Parts and materials constituting the products and its transportation (from supplier to Samsung factory)
Manufacturing	Product assembly by Samsung Electronics (Data collection period : 3 months ahead of assessment)
Distribution	From China or Vietnam to United States
Usage	2 years use
Disposal	Waste treatment of parts and material

Des marges d'incertitudes

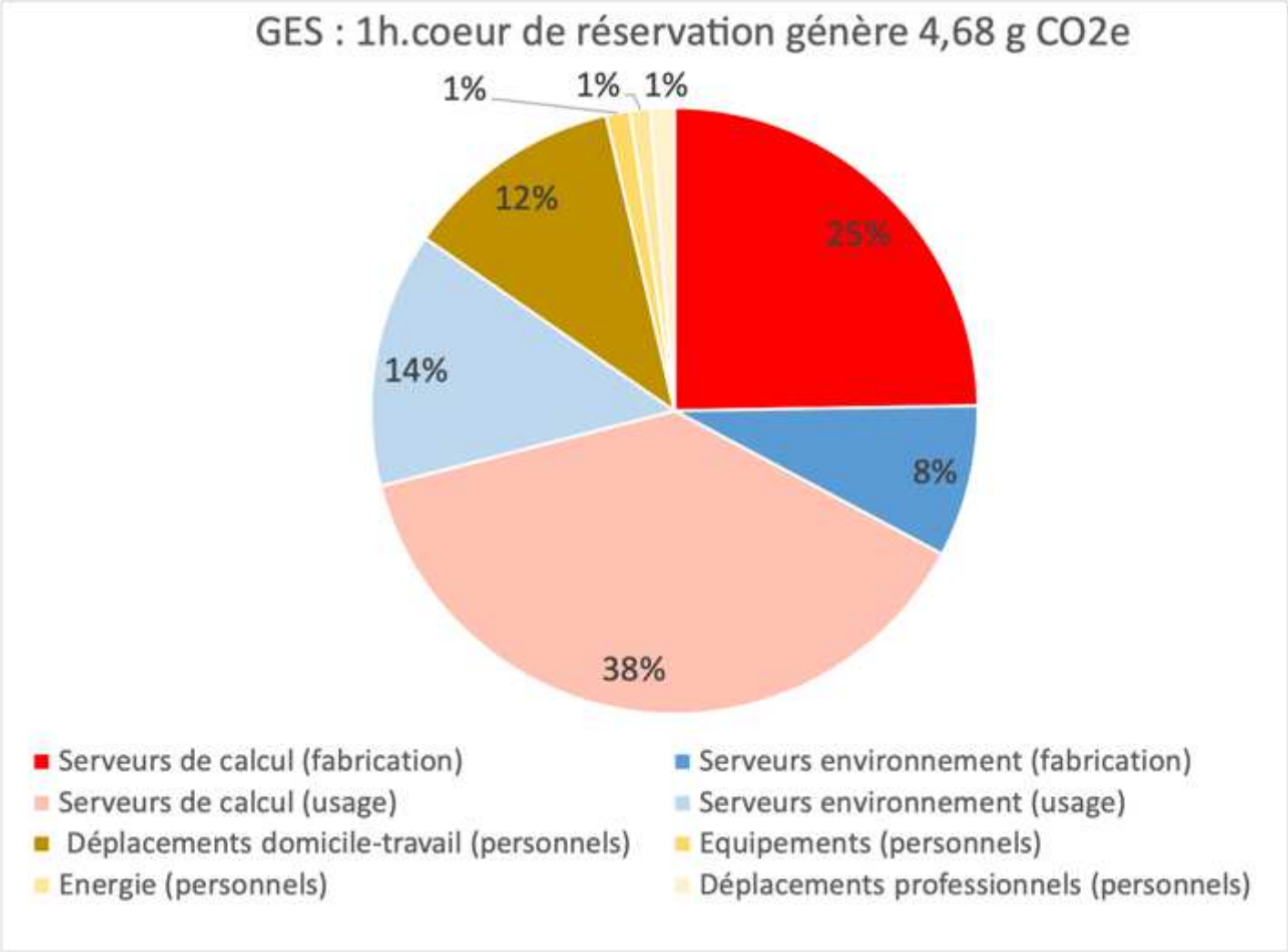


[Source: Product carbon footprint HP mt44, HP, 2018]

Assumptions

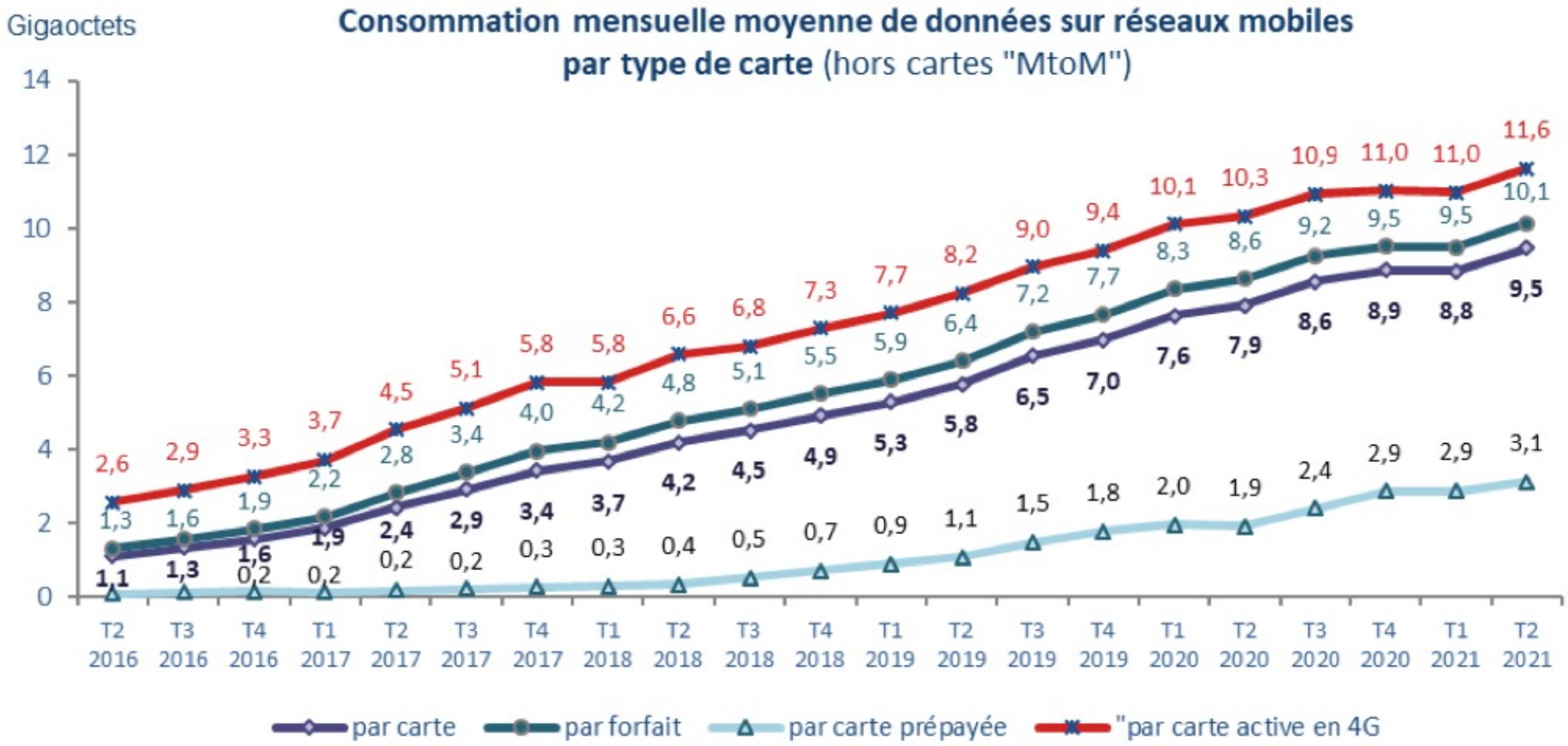
Lifetime of product	4 years
Use location	Worldwide
Use energy demand (Yearly TEC)	23.4 kWh
Product weight	1.8 kg
Screen size	14 inches
Final manufacturing location	China

Des périmètres variables



[Source : “Estimation de l'empreinte carbone d'une heure.coeur de calcul”, F. Berthoud, B. Bzeznik, N. Gibelin, M. Laurens, C. Bonamy, M. Morel, X. Schwindenhammer, rapport, 2020]

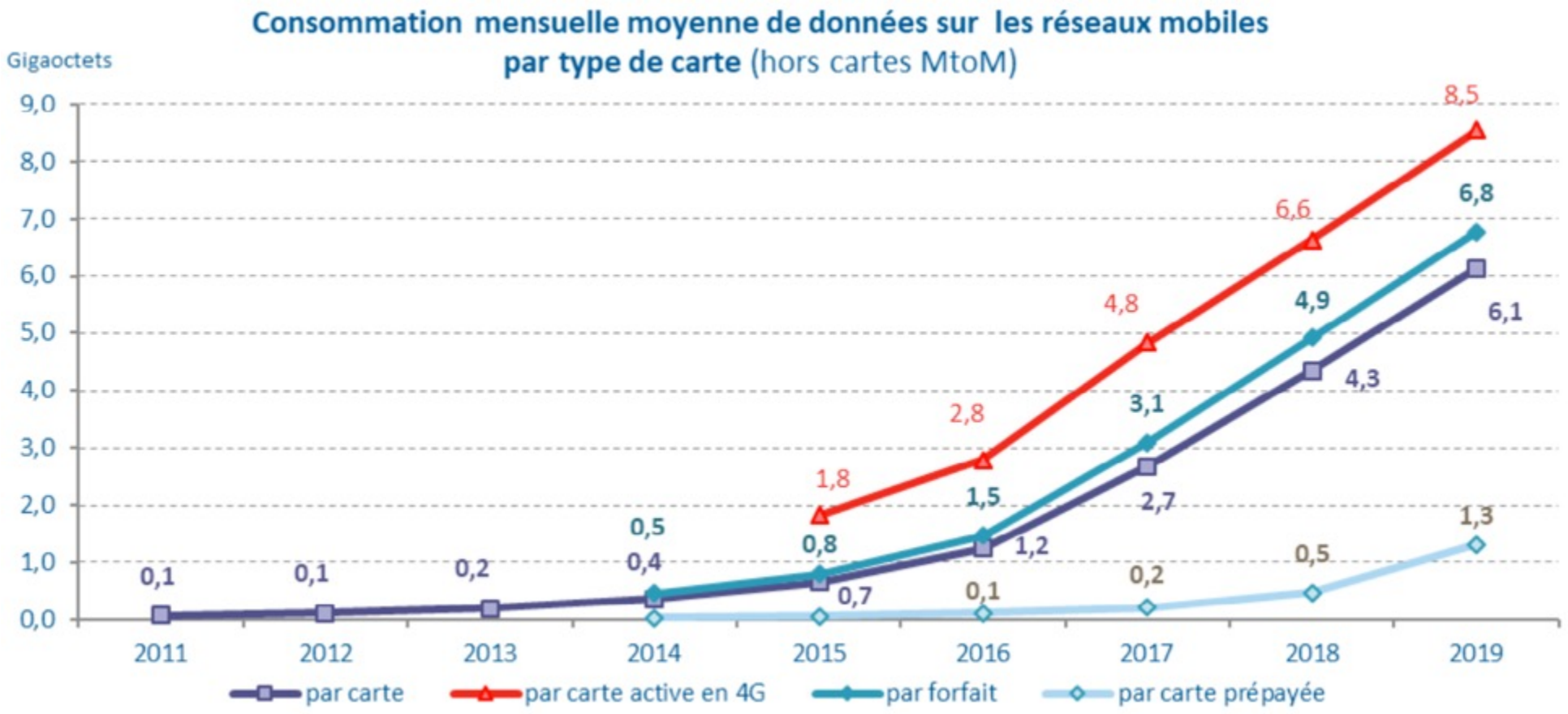
Toujours plus de trafic réseau



[Source : Marché des communications électroniques en France - Les chiffres au 2eme trimestre 2021, ARCEP]

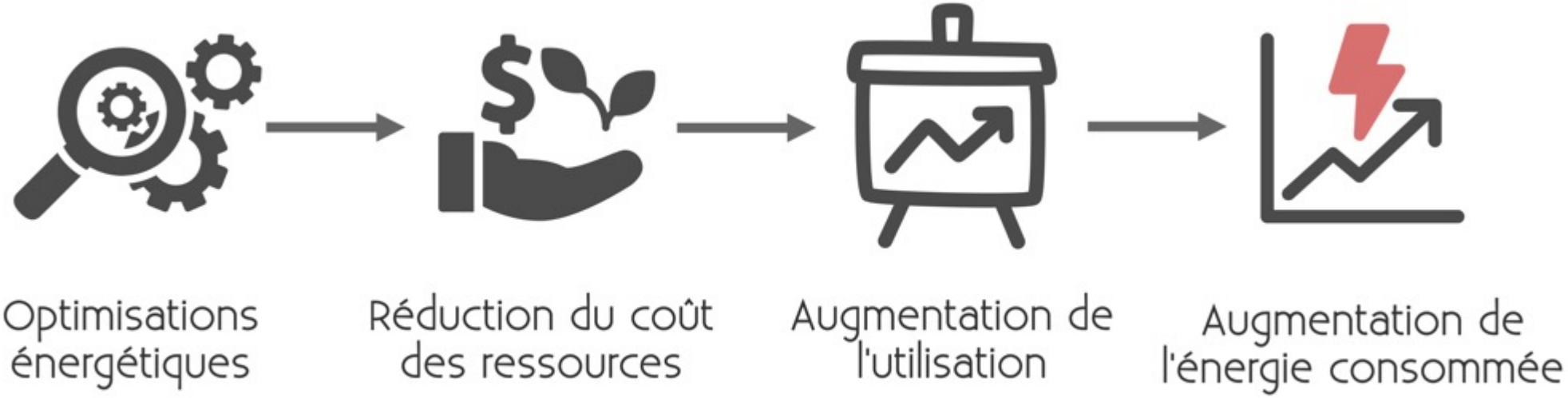
Plus de bande passante, plus de données, *quels impacts sur les infrastructures ?*

Toujours beaucoup plus de trafic réseau



[Source : Marché des communications électroniques en France - Année 2019, ARCEP]

Effets rebond



[Source : David Guyon]

Tendances de fond :

- Accélération des taux de renouvellement des équipements
- Explosion des usages et des consommations de données
- Numérisation de tous les secteurs, sans étude préalable des impacts environnementaux

Les impacts du numérique

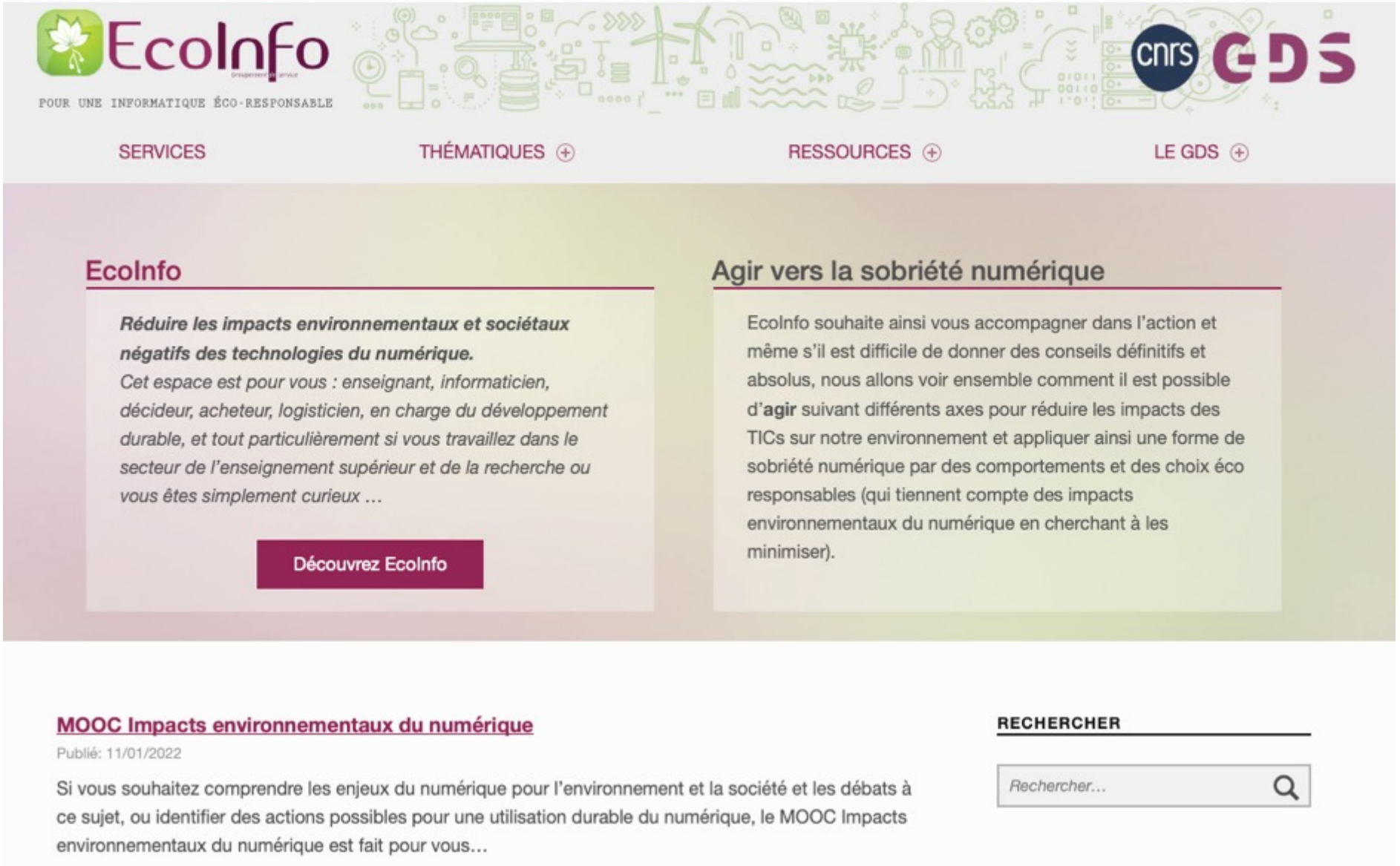
- Des effets directs à chaque étape du cycle de vie
  - Extraction : pollution, destruction d'éco-systèmes, conflits armés, épuisement des ressources
  - Transport
  - Utilisation : mix électrique
  - Déchets : collecte insuffisante, recyclage limité
- Des effets indirects plus ou moins positifs
  - Optimisation des autres secteurs
  - Obsolescence
  - Effets rebond
  - Interdépendance liée au numérique
  - Fracture numérique, santé (myopie, addictions, etc.)

Le numérique s'inquiète enfin de ses **impacts sur l'environnement** et ne se considère plus uniquement comme faisant partie de la **solution** mais comme également partie du **problème** du changement climatique.

Agir :

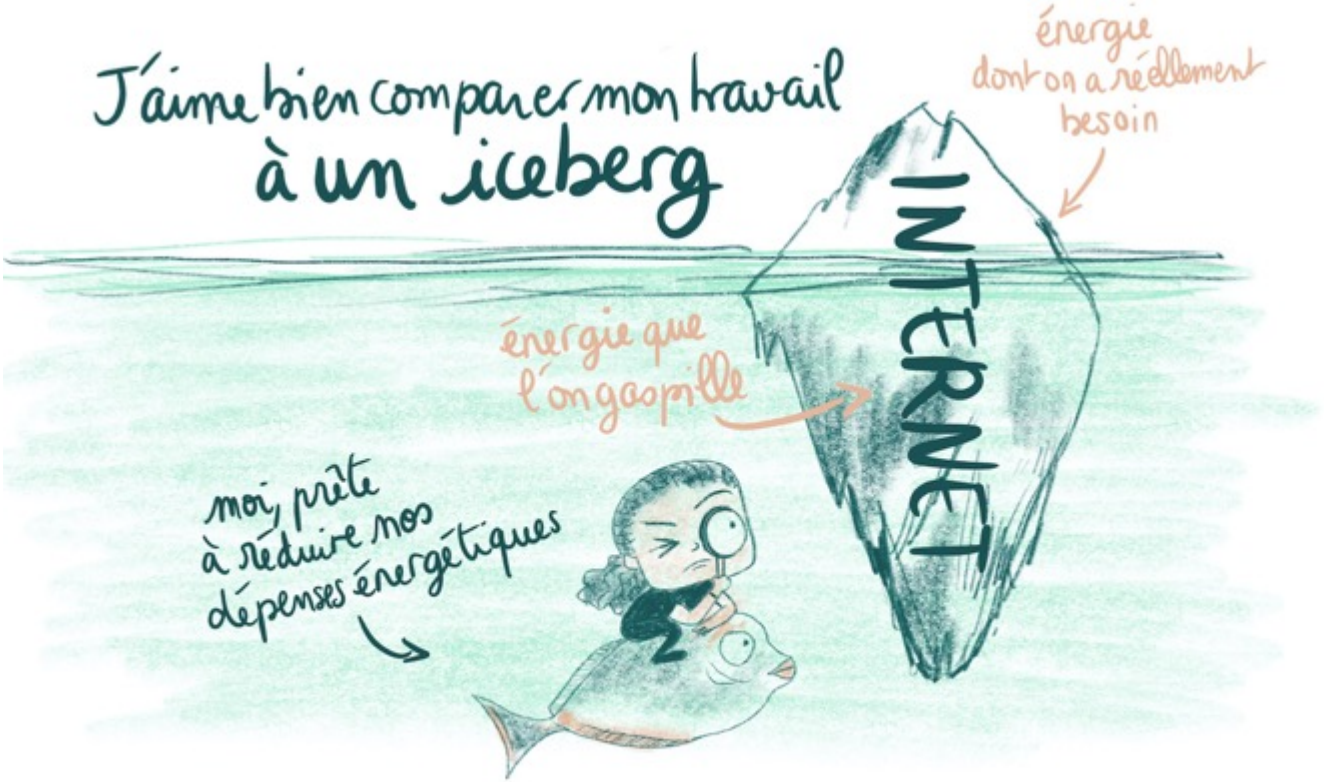
- Réfléchir
- Ralentir
- Réutiliser
- Réparer
- Recycler
- Repenser
- …

http://ecoinfo.cnrs.fr



Ce qui me plaît

- Comprendre en profondeur comment ça marche
- Échanger nos points de vue avec mes collègues
- Chercher à rendre le numérique plus vert

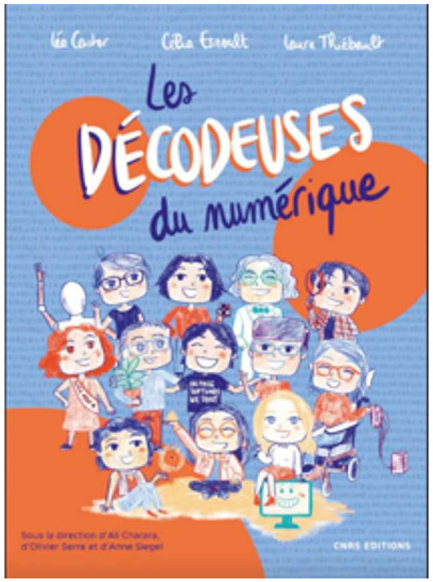


© Léa Castor / INS2I

« LES DÉCODEUSES DU NUMÉRIQUE ». incarner les sciences du numérique en BD

Décodeuses du numérique

- **Le projet**
  - **12 portraits de membres de laboratoires en numérique**
  - Illustré par Léa Castor, sur la base d'interview autour de quatre sujets : Science - Parcours - Déclic pour son métier - Sexisme ordinaire.
  - **Publiés**
    - En numérique: libre accès
    - sous la forme d'une BD : CNRS édition
  - **Envoyé** : dans tous les lycées de France et dans les collèges qui l'ont demandé



- **Les documents associés**
  - **Posters pour expositions**
  - **Fiches d'accompagnement**
  - **Fiche de lecture de l'Onisep**

https://www.ins2i.cnrs.fr/fr/les-decodeuses-du-numerique

Plus de 5 000 téléchargements

Plus de 20 000 lectures en ligne

Plus de 2 000 exemplaires papier vendus

Elsa Cazelles,
du sable pour trier le son

Magalie Ochs,
pour une IA plus humaniste

Françoise Conil,
coder pour un monde meilleur

Claire Mathieu
et l'amour des algorithmes

Emmanuelle Kristensen,
à l'écoute des neurones

Marie-Christine Rousset
au pays du web sémantique

Sarah Cohen-Boulakia,
détective publique des données biologiques

Lucile Sassatelli,
super-chercheuse en réalité virtuelle

Pauline Maurice
et les robots qui nous veulent du bien

Anne-Cécile Orgerie,
pour une informatique plus verte

Caroline Fontaine,
agent spécial de la cybersécurité

Nina Amini
et le contrôle du monde quantique

Merci de votre attention

http://people.irisa.fr/Anne-Cecile.Orgerie

[Crédits : Eric Drezet]