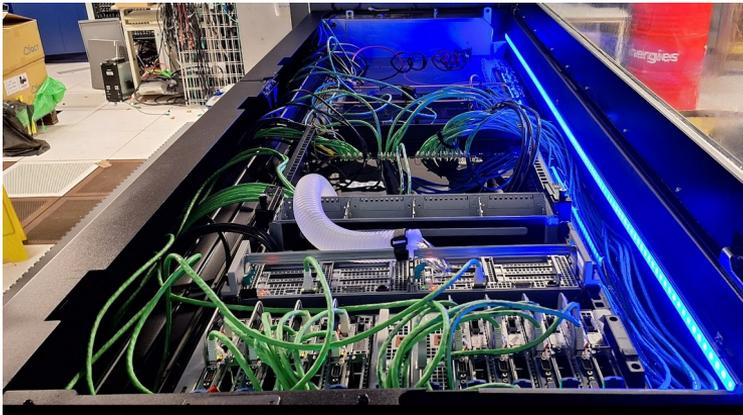


Les stratégies de consommation : consommer moins ou consommer mieux ?



Emmanuel Quémener



Des nouveaux usages... ... aux nouveaux contextes !

- L'IT a une part croissante (de consommation) :
 - Parce qu'on s'en sert de plus en plus (pas une discipline épargnée)
 - Parce que les outils sont plus gourmands (*Machine Learning* en tête)
- Mais un contexte de rentrée 2022 unique :
 - Émergence (tardive ?) d'une « conscience écologique »
 - Augmentation drastique du coût de l'énergie (électricité en tête x10)
 - Risques de coupure électrique l'hiver 2022-2023
- Questionnement tous azimuts de direction & laboratoires
- Comment aborder le problème ?

Ce que ce ne sera pas...

Pas de « reprise » de Labos 1.5 !

Estimation de l'empreinte carbone d'une heure de calcul
sur un cœur CPU ou sur un GPU

Methodologie

Par Labos 1point5

[Janvier 2024]

Dans cette version de la méthodologie, il a été décidé d'exclure la fin de vie, compte tenu de son très faible impact relatif compte tenu des circonstances de traitement des équipements électriques et électroniques (DEEE) professionnels en France. La récupération de la chaleur fatale est exclue car hors périmètre.

Sommaire

Introduction	3
Estimation de l'empreinte carbone d'un service de calcul	4
Estimation des unités fonctionnelles	11
Incertitudes	13
Annexes	15
Références	21

Justement, on va s'intéresser à ces éléments !

Ce que ce sera : approche locale sur une « meute » de machines...

- Si cela ne fonctionne pas à petite échelle...
 - Aucune chance que cela fonctionne à une plus grande !

Cloud@CBP : État des ressources

Recherche : utilisation d'adresse IP 140.177.20.226
Vous souhaitez savoir avec le navigateur Mozilla avec Cloud@CBP

Le 2023-10-01, Heures Locales 18:03 | 131 machines "chargées" à 54.27 et utilisées par 72 utilisateurs

A ce instant, CPU : 209 unités avec 2028 cœurs dans 54 machines différents. GPU : 158 unités dans 12 machines différents.

Liens rapides : Configuration, Demande d'accès ou d'assistance

Sélection d'une machine

- Machine générique
- Machine multi-cœurs (n=2)
- Machine à processeur ARM (n=2000)
- Machine avec gros GPU Au Gamer
- Machine avec GPU(FPGA)

Submit Query | Reset

Liste des machines avec caractéristiques techniques

Hostname	IDCPU	Architecture	OS	RAM	Cores	GPU	GPU Mem	GPU Type	GPU Vendor	GPU Model	GPU Mem	GPU Vendor	GPU Model	GPU Mem	GPU Vendor	GPU Model
apollon001	1.13	3	12	960	1000	2400										
apollon002	1.19	2	12	180	1000	1800										
apollon003	1.32	3	12	180	1200	1817										
apollon004	1.25	3	12	1200	1200	1310										
apollon005	1.19	3	12	180	1200	1817										

Cluster@CBP : État des ressources

Recherche : utilisation d'adresse IP 140.177.20.226
Vous souhaitez savoir avec le navigateur Mozilla avec Cloud@CBP

Le 2023-10-01, Heures Locales 18:03 | 156 machines "chargées" à 24.10 et utilisées par 1 utilisateur

A ce instant, CPU : 312 unités dans 248 cœurs dans 14 machines différents. GPU : 5 unités dans 3 machines différents.

Liens rapides : Configuration, Demande d'accès ou d'assistance

Liste des machines avec caractéristiques techniques

Hostname	IDCPU	Architecture	OS	RAM	Cores	GPU	GPU Mem	GPU Type	GPU Vendor	GPU Model	GPU Mem	GPU Vendor	GPU Model	GPU Mem	GPU Vendor	GPU Model
apollon001	1.13	3	12	960	1000	2400										
apollon002	1.19	2	12	180	1000	1800										
apollon003	1.32	3	12	180	1200	1817										
apollon004	1.25	3	12	1200	1200	1310										
apollon005	1.19	3	12	180	1200	1817										

Servers@CBP : État des ressources

Recherche : utilisation d'adresse IP 140.177.20.226
Vous souhaitez savoir avec le navigateur Mozilla avec Cloud@CBP

Le 2023-10-01, Heures Locales 18:04 | 26 machines "chargées" à 22.88 et utilisées par 6 utilisateurs

A ce instant, CPU : 54 unités dans 432 cœurs dans 17 machines différents. GPU : 26 unités dans 5 machines différents. Stockage : 160 unités dans 10 ports et 60 disques 2T5.

Liens rapides : Configuration, Demande d'accès ou d'assistance

Liste des machines avec caractéristiques techniques

Hostname	IDCPU	Architecture	OS	RAM	Cores	GPU	GPU Mem	GPU Type	GPU Vendor	GPU Model	GPU Mem	GPU Vendor	GPU Model	GPU Mem	GPU Vendor	GPU Model
apollon001	1.13	3	12	960	1000	2400										
apollon002	1.19	2	12	180	1000	1800										
apollon003	1.32	3	12	180	1200	1817										
apollon004	1.25	3	12	1200	1200	1310										
apollon005	1.19	3	12	180	1200	1817										

Les 3 infrastructures Cloud@CBP, Cluster@CBP et Servers@CBP

- Plus de 300 machines (plus d'une centaine éteintes démarrables à discrétion)
- Un taux de disponibilité supérieur à 99.8 % (0.1 % coupure électrique annuelle)
- Tout ça avec 90 % d'équipements hors garantie et 70 % d'occasion...

Inéquation impossible ?

Moins consommer & mieux servir...

- Nécessité de placer des « nombres » sur des « faits »
- Indicateurs de « consommation » :
 - Croissance (« infinie » dans un environnement « fini » :-/ du physicien)
 - **ADP** : potentiel d'épuisement des ressources abiotiques :
 - Abiotic Depletion Potential (unité kgSbeq)
 - **PRG** : Potentiel de Réchauffement Global ou « empreinte carbone »
 - GWP : Global Warming Potential (unité kgCO₂eq)
 - **PE** : Consommation de ressources énergétiques
 - Primary Energy (unité MJ)

Quelle empreinte carbone ?

Application d'une vieille approche...

- Analyse « comptable », « **3 coûts** » pour tout « service »
 - **Coût d'entrée** : appropriation, développement, intégration, ...
 - **Coût d'exploitation** : MCO, évolutions réglementaires, sanitaires, ...
 - **Coût de sortie** : remplacement, abandon, délégation, ...
- Pour du matériel (et son usage) : même combat !
 - **Avant** : sa fabrication (et son transport)
 - **Pendant** : son exploitation (et sa maintenance)
 - **Après** : son recyclage (et son transport, stockage)
- **Pendant** : consommation électrique (et chaleur fatale)...

Quelle empreinte carbone ?

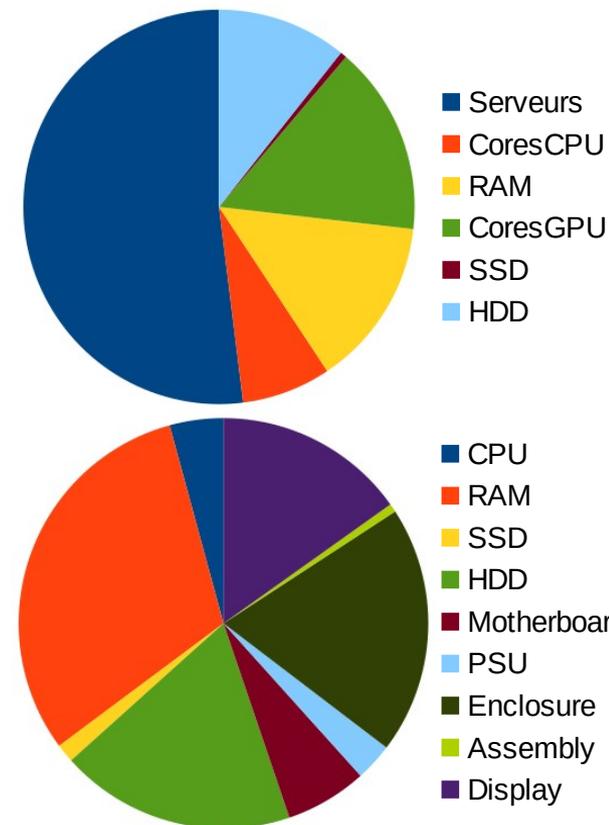
Interrogations sur la « littérature »

- Petite expérience : *établir un devis sur Matinfo5*
 - Empreintes carbone identiques quel que soit : le modèle CPU, la RAM
- Littérature plus « pertinente, cohérente et consistante » :
 - <https://boavizta.org/blog/empreinte-de-la-fabrication-d-un-serveur>
 - **Approche 2 : « facteurs d'émission arbitraires par composant »**
 - $\text{servergwp}(\text{kgCO}_2\text{eq}) = 900(\text{kgCO}_2\text{eq}) + \text{cpuunits}(\text{unit}) \times 100(\text{kgCO}_2\text{eq}/\text{unit}) + \text{ramsize}(\text{GB}) \times 150/128(\text{kgCO}_2\text{eq}/\text{GB}) + \text{ssdunits}(\text{unit}) \times 100(\text{kgCO}_2\text{eq}) + \text{hddunits}(\text{unit}) \times 50(\text{kgCO}_2\text{eq}) + \text{gpuunits}(\text{unit}) \times 150(\text{kgCO}_2\text{eq}/\text{unit})$
 - **Approche 3 : « vers une formule de calcul d'impact multicritère » basé sur les semiconducteurs**
 - $\text{server} = \text{cpu} + \text{ram} + \text{ssd} + \text{hdd} + \text{motherboard} + \text{psu} + \text{enclosure} + \text{assembly}$
 - $\text{cpu} = \text{cpuunits} \times ((\text{cpucoreunits} \times \text{cpudiesize} + 0,491) \times \text{cpu_die} + \text{cpu_base})$
 - $\text{ram} = \text{ramunits} \times ((\text{ramsize} / \text{ramdensity}) \times \text{ram_die} + \text{ram_base})$
 - $\text{ssd} = \text{ssdunits} \times ((\text{ssdsize} / \text{ssddensity}) \times \text{ssd_die} + \text{ssd_base})$
 - $\text{hdd} = \text{hddunits} \times \text{hdd_unit}$
 - $\text{psu} = \text{psuunits} \times \text{psuunitweight} \times \text{psu_weight}$
 - $\text{enclosure} = \text{rack ou enclosure} = \text{blade} \times \text{bladeenclosure}/16$

Analyse : CBP comme « pollueur »

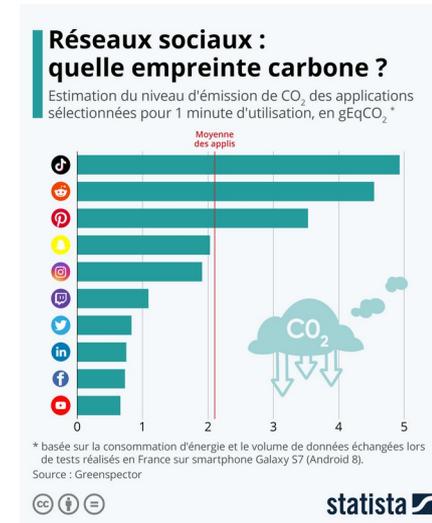
Fabrication & Exploitation

- Fabrication : 322 machines
 - 10 % sous garantie, 1/3 achetées neuves
 - 6572 coeursCPU, 9245 coeursGPU,
 - 63 TiB RAM, 4 PB, 1200 HDD, 53 SSD
 - 557 ou 277 tonnes CO₂ à la fabrication
- Exploitation : 322 avec 150W H24 7/7
 - 21 tonnes CO₂ par an soit 2 françaises « moyennes »
 - Mais 400 utilisatrices différentes chaque mois !
- « Amortissement » carbone : jamais l'égalité...
 - Ratio : 1 pour 26 à 1 pour 12...



Pour fixer les idées sur l'empreinte carbone individuelle...

- Un smartphone :
 - 80 kgCO₂ à la fabrication, 800 gCO₂ à l'utilisation par an
 - « Amortissement » carbone sur 100 ans...
 - Durée de vie de 3 ans...
- Sauf qu'un smartphone n'est qu'un « terminal » !
- Usage des réseaux sociaux : forte disparité...
 - En moyenne 2 gCO₂/minute
 - Pour un jeune utilisant TikTok 4h/jour : 450kgCO₂/an
 - Équivalent carbone de 7 machines du CBP/H24 7/7

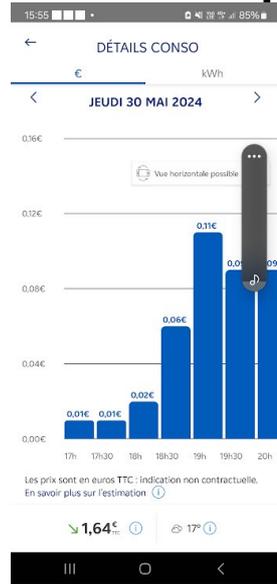


Actions à envisager

- D'abord réduire les coûts d'entrée (ou de sortie)
 - En exploitant le maximum les machines déjà fabriquées
- Puis réduire les coûts d'exploitation :
 - En exploitant la chaleur fatale directement
 - En améliorant l'évacuation la chaleur fatale
- Mais, avant tout « plan de bataille », « *intelligence* » :
 - Connaître sa consommation à petite échelle (machine ou composant)
 - Connaître sa consommation à grande échelle (Data Center)

Appréhension : récupérer les infos... A son échelle, c'est « simple »...

- Chez soi, le compteur Linky et son application...



```
numa@pound: ~  
File Edit View Search Terminal Help  
Device: /org/freedesktop/UPower/devices/DisplayDevice  
power supply: yes  
updated: Tue 04 Jun 2024 12:45:37 PM CEST (1 seconds ago)  
has history: no  
has statistics: no  
battery  
present: yes  
state: discharging  
warning-level: none  
energy: 24.7152 Wh  
energy-full: 61.3662 Wh  
energy-rate: 7.923 W  
charge-cycles: N/A  
time to empty: 3.1 hours  
percentage: 40%  
icon-name: 'battery-good-symbolic'  
  
Daemon:  
daemon-version: 0.99.20  
on-battery: yes  
lid-is-closed: no  
lid-is-present: yes  
critical-action: PowerOff  
numa@pound: ~$
```

- Sur son laptop, la commande « upower -d » en décharge...

Appréhension : récupérer les infos...

En fait, pas si simple !

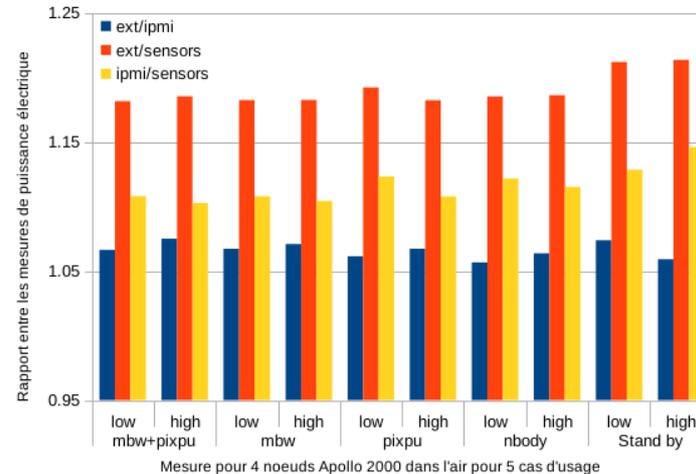
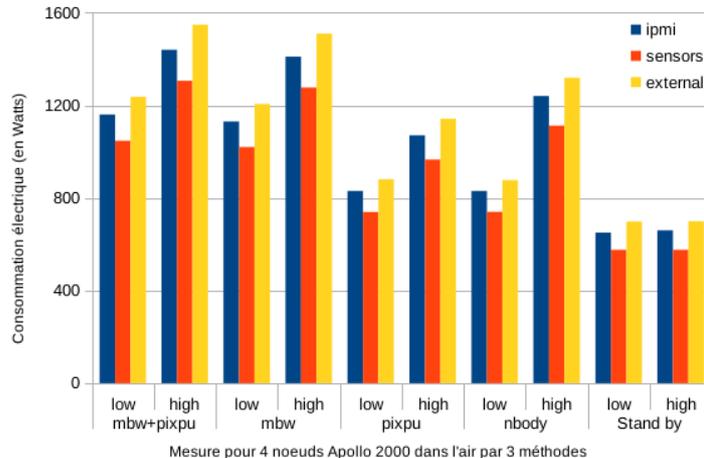
- Récupérer « localement » : échelle du composant
 - Via le système d'exploitation, directement : sensor
 - Via l'IPMI et l'OS : « ipmitool » ou mieux « ipmi-sensors »
 - Via un wattmètre (et une webcam)
 - Via une pince ampèremétrique
 - Via la supervision du PDU (à son échelle...)
- Récupérer « globalement » : Data Center
 - Un site web authentifié (en Java)
 - « page Web » = « framebuffer »

SOUS SOL	JOUR EN COU	CUMUL
GENERAL TGBT	5450 kWh	20130286 kW-hr
GENERAL ECLAIRAGE	0 kWh	1694 kW-hr
GENERAL CVC	27 kWh	118004 kW-hr
GENERAL CVC TT	1057 kWh	3990965 kW-hr
GENERAL TGHQ NDC	2899 kWh	9464892 kW-hr
GENERAL TGHQ CORPORA	116 kWh	826766 kW-hr
GENERAL TGHQ STOCKAGE	399 kWh	2046484 kW-hr

Mesures « locales »

Quel comportement à la charge ?

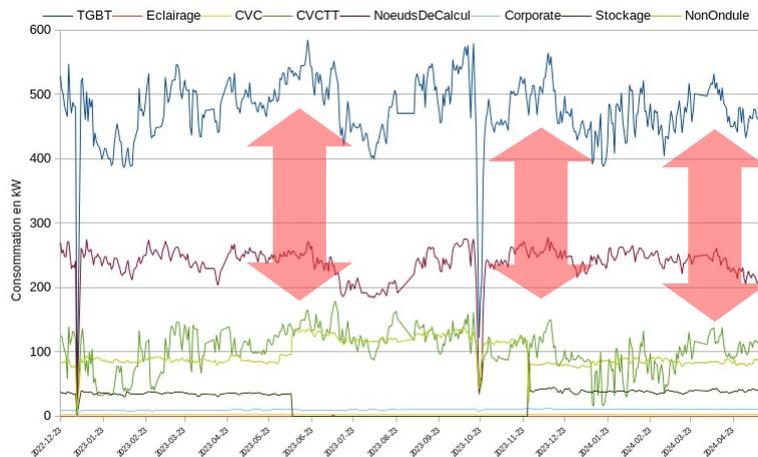
Expérimentation : basse/haute fréquences, 5 cas d'usage



- Des mesures à appréhender avec précautions !
- Solution : MesureIPMI (compensée au besoin...)
- Mais le troupe Wattmètre/Webcam/Led reste une solution systématique...

Analyse du DataCenter la statistique (et son pentacle)

Période du 23/12/2022 au 13/05/2024 : plus de 500 jours

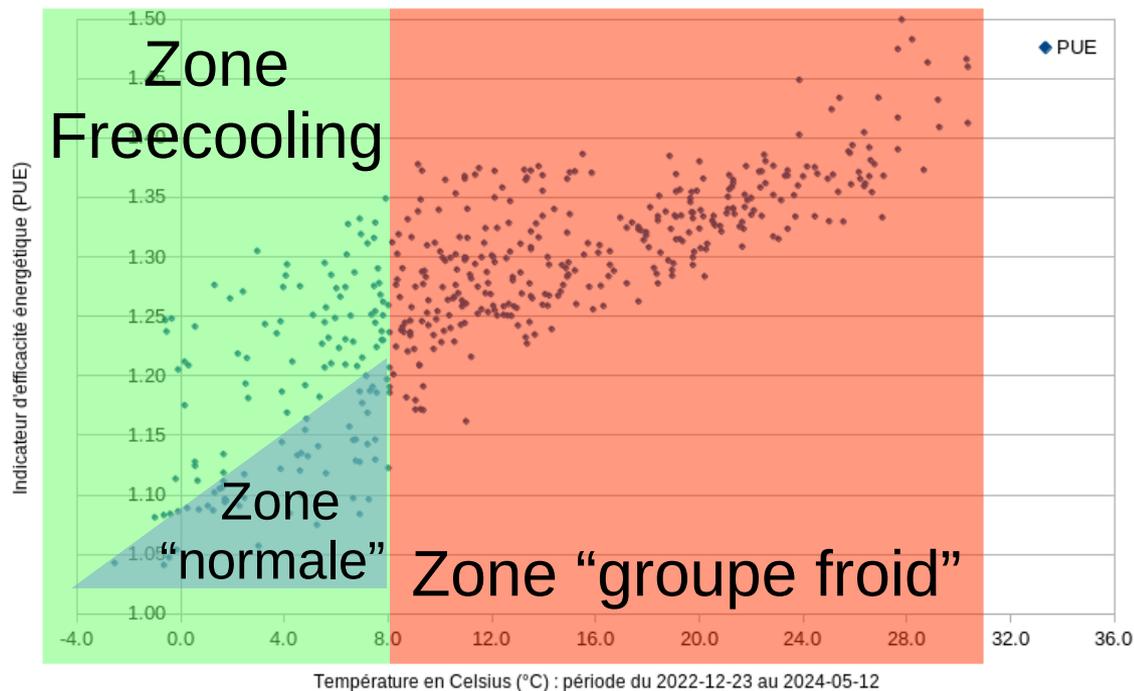


Légende sommaire...

- **TGBT** : arrivée électrique
- **CVCTT** : climatisation
- **NoeudsDeCalcul** : éponyme...
- **Corporate** : machines DSI
- **Stockage** : machines labos
- **Non ondulé** : tout sauf

- Moyenne : 478 kW, Médiane : 476 kW, Max-Min ~200 kW
- Une forte variabilité, mais quelle est sa « nature » ?

Mais qu'en est-il du PUE sur cette période de 500 jours ?



DataCenter « état de l'art » (PUE 1.29) mais **perfectible !**

Analyse du DataCenter

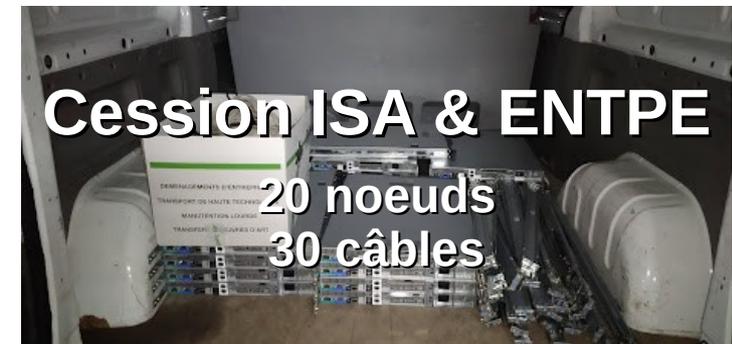
Mais que faire en « exploitation » ?

- Des pistes d'améliorations :
 - Améliorer le PUE, mais hors champ « utilisateurs »
 - Limiter la consommation des serveurs consommateurs (à GPU)
 - Limiter la consommation du centre de calcul (600 machines)
 - Dépayser des machines dans les bureaux
 - Éteindre les machines anciennes
- Avec la « nouvelle donne » :
 - Explosion du prix le jour : entre 6h et 22h, x8 du prix
 - Risques de coupure électrique...

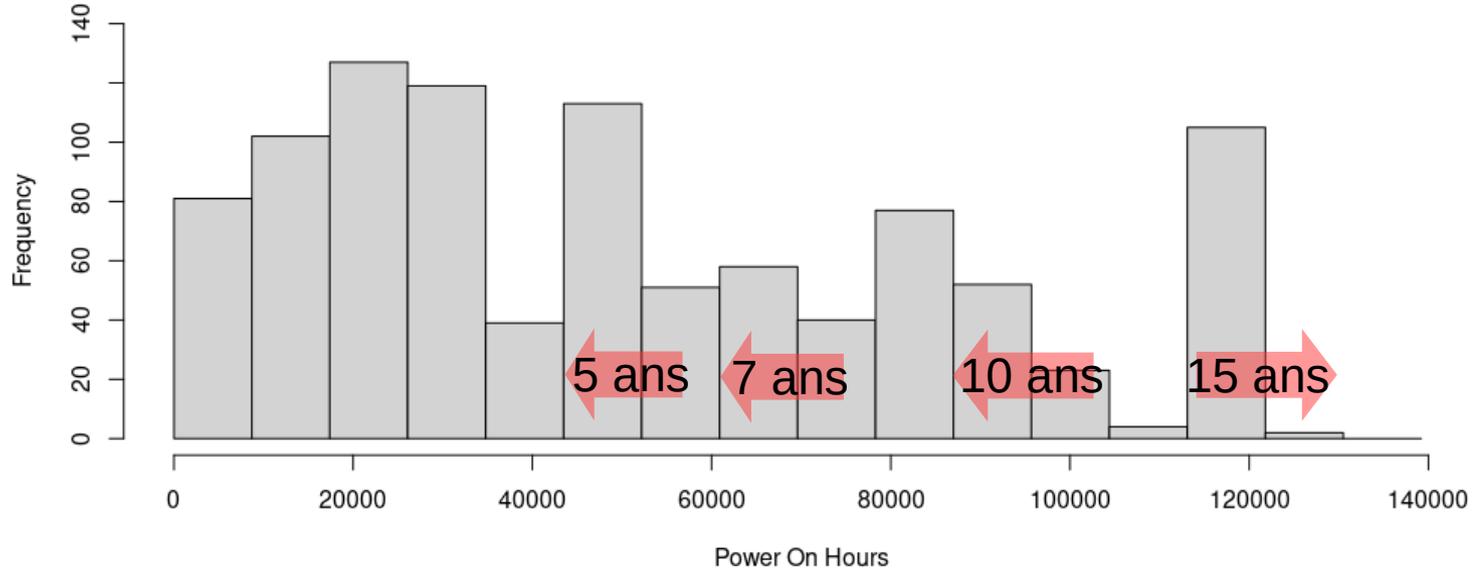
Action #1 « Avant & Après » : privilégier les « cycles courts »

« Les déchets des uns sont les ressources des autres. »

- Constats (implacables) :
 - Pas de fabrication « locale »
 - Ressources inexploitées à proximité
- Quelles actions « Avant » :
 - récupération, requalification,
 - démontage, détournement,
 - achat d'occasion chez broker
- Quelles actions « Après » :
 - Cession des machines inexploitées



Et les disques durs (50kgCO₂/HD) Réaffectation plutôt que recyclage...



- Des disques plus « résistants » que la garantie,
 - Donc la garantie n'est pas une DLC (Date Limite de Consommation)
 - Une garantie n'est pas une garantie de fonctionnement (loin de là !)

Action #1 : les « cycles courts » généralisables ?

Oui, une question de volonté mais il faut :

- Disposer de systèmes « résilients » (solutions propriétaires exclues)
- Avoir de quoi tester rapidement les arrivées de matériel
- Stocker le matériel de secours dans des endroits « sûrs »
- Connaître les usages pour adapter le matériel
- Privilégier les systèmes d'exploitation libres

Action #2 : cycles arrêt/redémarrage sur 4 salles & 64 machines...

- Objectif : éteindre les machines des salles « hors cours »
- Opérations : allumage à 7h30, extinction à 19h30
- Craintes : vieillissement prématuré des HDD
 - Critères : cycle marche/arrêt, variation température, température max
- Adaptations :
 - Dans le BIOS, activation WoL & désactivation autres modes...
- Retour sur presque 2 années : indisponibilité rare...
 - Vieillissement du matériel à évaluer

Actions #3 : relocaliser le matériel déployer des « AnchiAles »

Truisme : chauffer les bureaux en récupérant la « chaleur fatale »

- Pas nouveau (ancien chauffage de véhicules « thermiques »)
- Contexte favorable : interdiction des chauffages d'appoint...
- Passage de 6 machines à 20 machines entre fin 2022 et mi 2024



Action #3 : déploiement AnchiAles

Quel bilan ? Généralisation ?

- Plus de demandes que d'offres en 2023 :
 - Plus de demandes que d'offres possibles !
- La chaleur « ventilée » offre un confort meilleur
 - Une machine « à vide » (de 100W) offre un confort de radiateur de 1kW
- La généralisation exige une « modification » d'approche
 - Des équipements « plus » génériques (salle machine, bureau, ...)
 - Une gestion de la transhumence entre les saisons
 - Un placement estival dans des zones « hors d'eau »
 - Des bacs d'immersion « individuels » à diffuser dans les locaux...

Action #4 : immerger les machines

Limiter la PUE à ~ 1 voire < 1

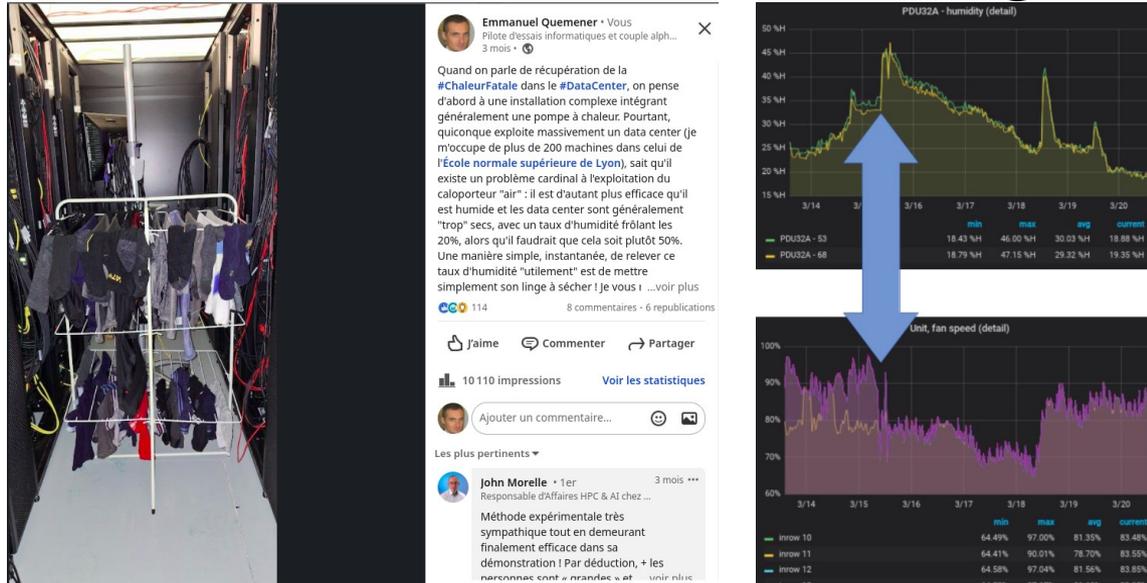


Action #4 : avantage de l'immersion

Des avantages aux inconvénients...

- Constats d'exploitation après 2 ans :
 - Réduction de consommation « directe » de 10 % à 20 %
 - Température de l'huile jusqu'à 55°C sans pertes de performance
 - Bonne tenue des équipements (un taux de panne comparable)
- Avantages :
 - Refroidissement de tous les composants sans distinction de nature (barrettes de RAM)
 - Adaptation possible de l'existant ou exploitation de « matériel générique »
 - Exploitation de la chaleur fatale plus facile (eau à 50°C)
 - Exploitation d'huile « bio-sourcée » (huile de friture désoxygénée)
- Inconvénients :
 - En plus des « risques » électriques, des « risques » chimiques
 - Frilosité des intégrateurs depuis quelques années & « marchands du temple » se goinfrant !

Action #5 : dans le Data Center Améliorer le refroidissement... En séchant son linge !

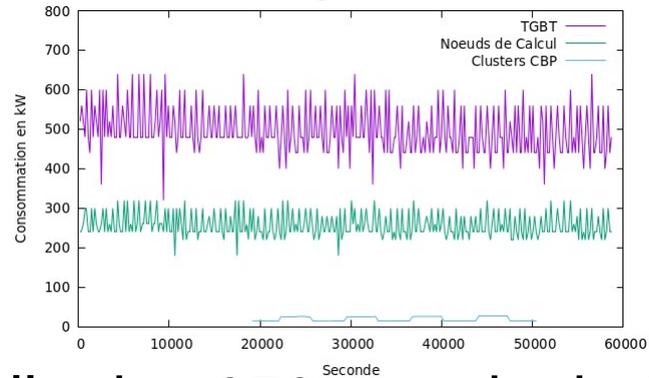
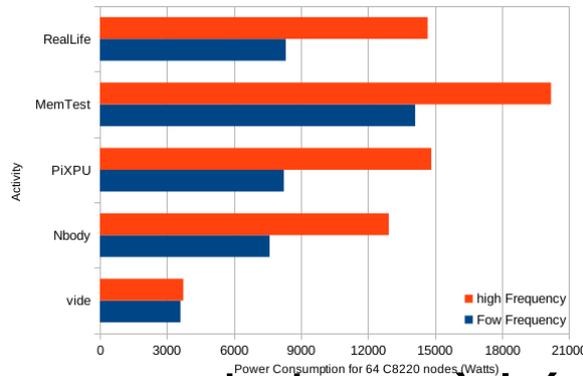


Un exemple de win-win

Action #6 : dans le DataCenter

Limiter sans gréver le HPC

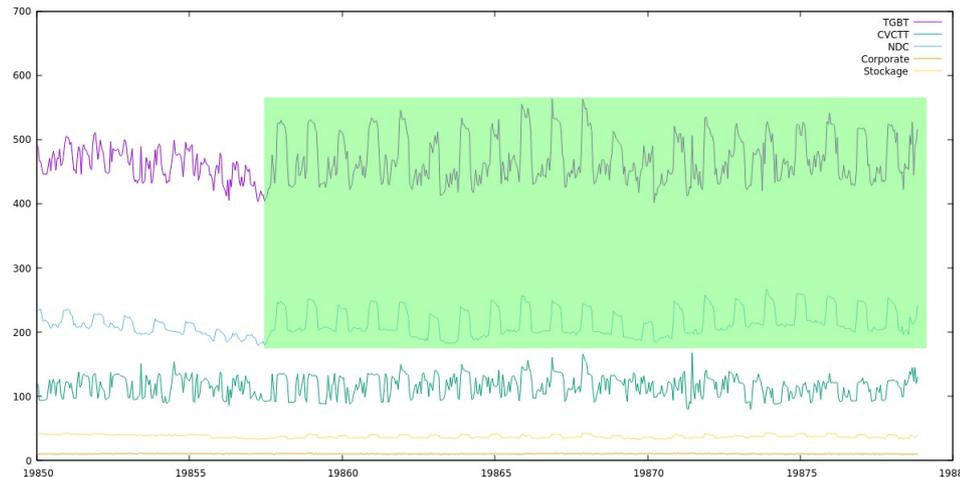
- Un prérequis : laisser à l'OS le contrôle
 - Et donc paramétrer le BIOS avec la fréquence en « OS Control »
- Expériences habituelles + cas d'usage sur 64 nœuds



- Pas grand-chose à l'échelle des 650 nœuds du DC...
 - Généralisation en cours sur le Mésocentre PSMN

Action #7 sur PSMN : approche multiple

- Limiter la fréquence le jour de 6h à 22h (divisée par 2)
- Ne pas exécuter de nouvelles tâches le jour



- De l'ordre de 20 % de consommation en moins...

En conclusion

Consommer moins ou mieux ?

- Déjà, la modulation de consommation est possible :
 - Par une extinction des installations « diurnes » (salles de cours)
 - Par une modulation de la fréquence pour les H24 7/7
- Consolider les installations ou les approches
 - Améliorer le PUE du Data Center existant : pompe à chaleur...
 - Créer un Data Center « en immersion » et exploiter la « chaleur fatale »
 - Distribuer des AnchiAles avec machines « immergées »
 - Renouveler son parc par un moins énergivore ?
 - C'est quand même externaliser son empreinte carbone à l'extérieur...