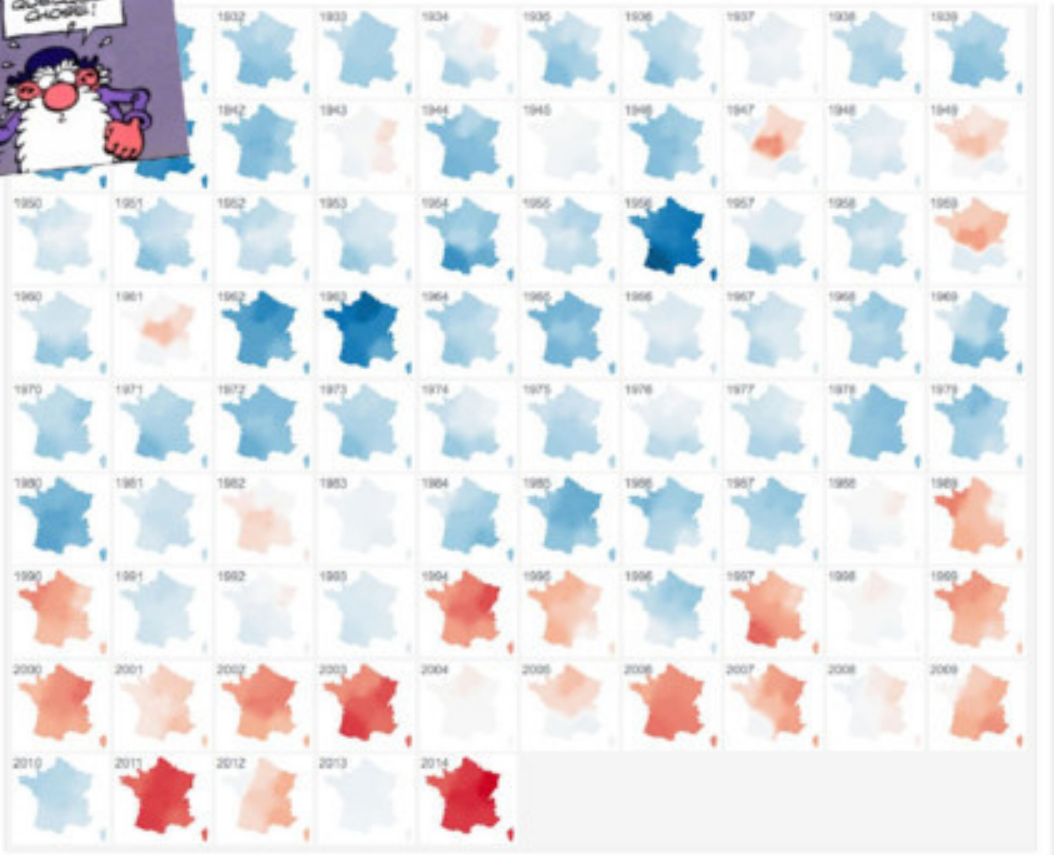
Quelques recherches sur la consommation et l'efficacité énergétique des équipements informatique

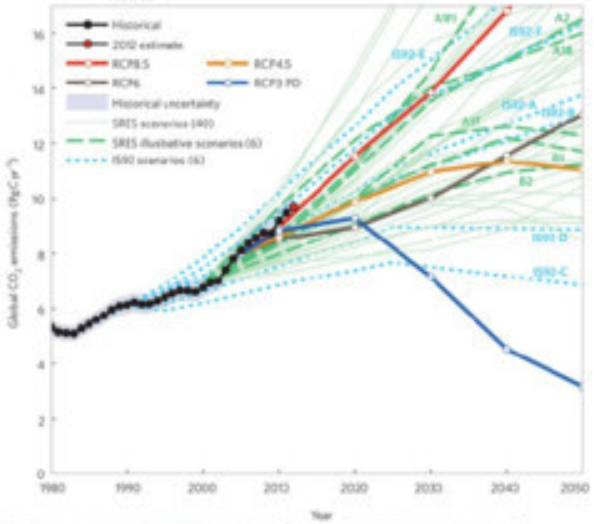
Laurent Lefèvre

10 ans d'Ecoinfo, Paris, 23 Avril 2015

INRIA AVALON / LIP
Ecole Normale
Supérieure de Lyon

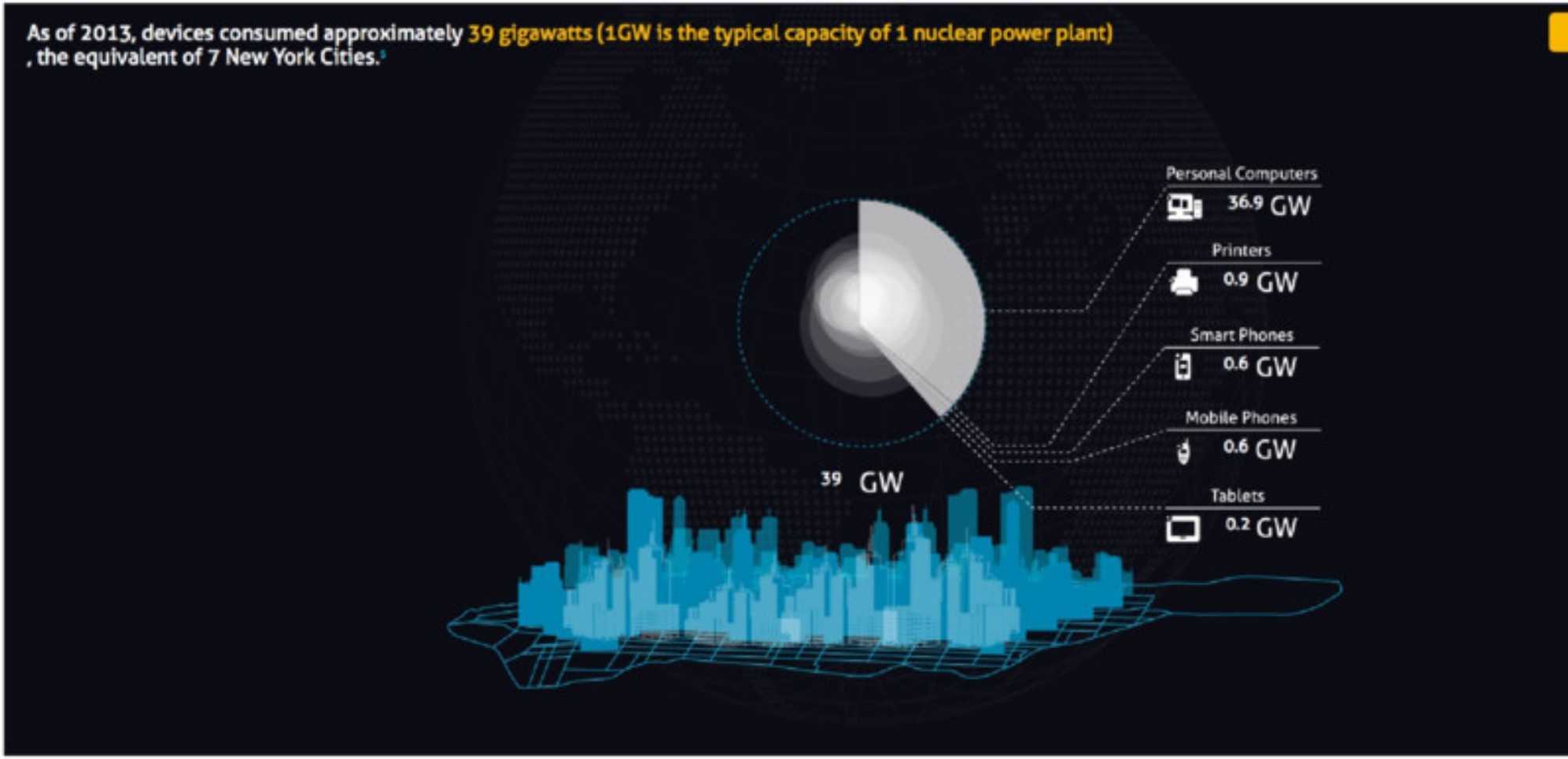
Ça chauffe !





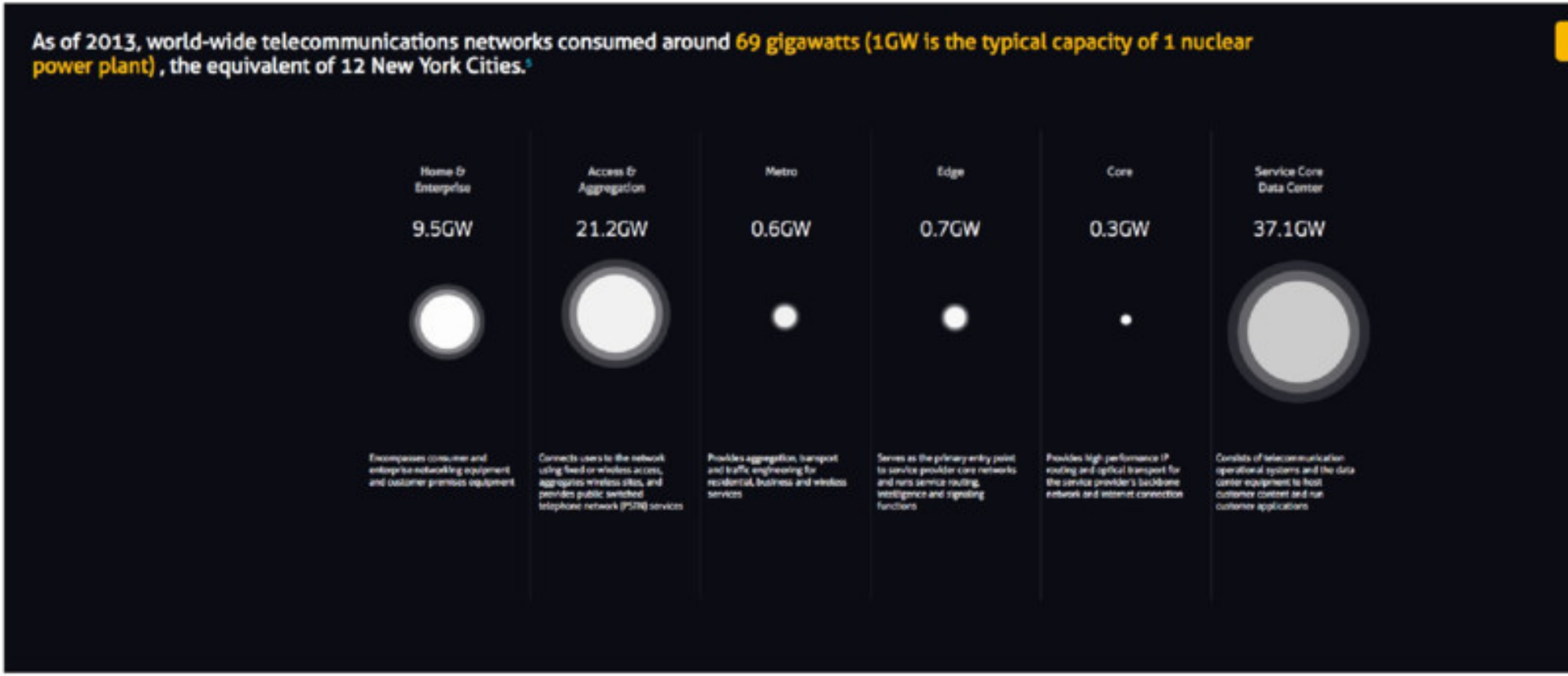
Quel rapport avec moi ? je ne suis qu'un chercheur en informatique…

Informatique en phase d'usage



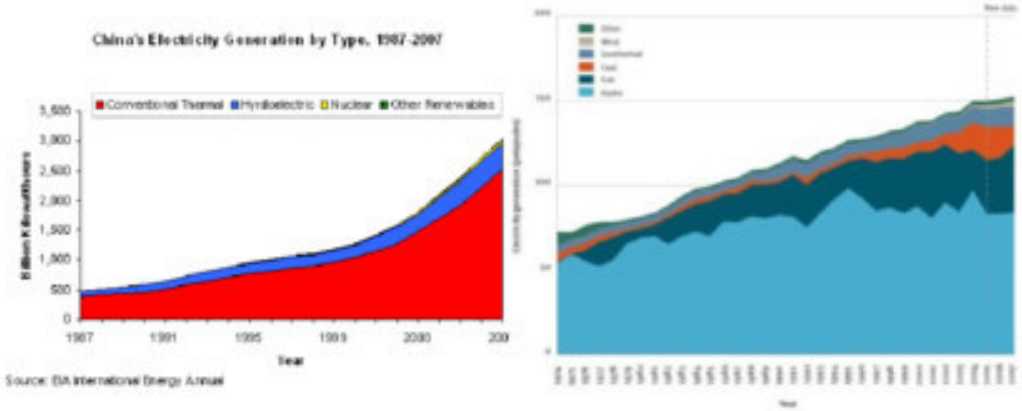
http://gwatt.net

Réseaux et Dcs



Recherches sur l'efficacité énergétique en phase d'usage

- Réduire la consommation électrique des équipements
- Emballement récent propulsé par la vague « Green IT »
- On se focalise sur des watts !
- Des systèmes embarqués à récupération d'énergie aux gros datacentres (20 → 100 MW)
- Approches multi-dimensionnelles et intégrées : matériel et logicielles

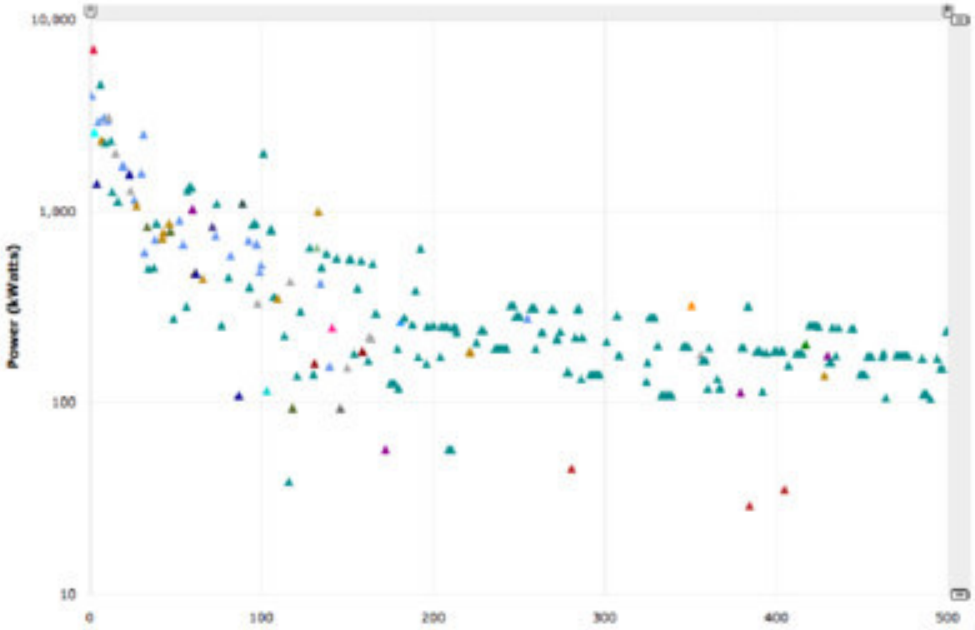


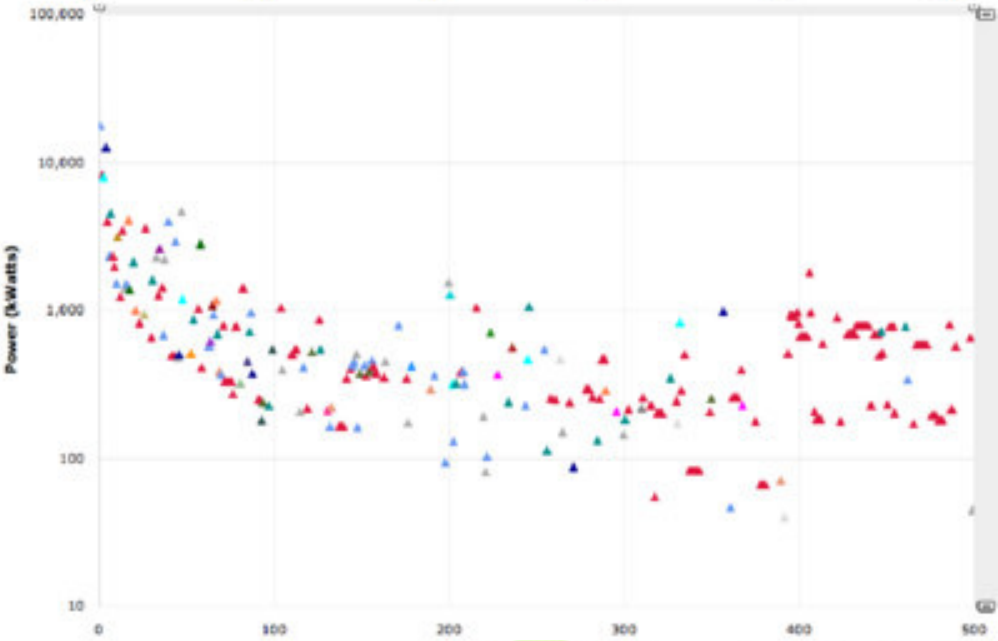
Energie : facteur limitant pour le déploiement de système à grande échelle (Datacentres, Cloud, réseaux) ?

Sans doute !

Mais la consommation et le nombre de datacentres continuent d'augmenter

- Top500 : Nov 2010 : 127 MW – Nov 2014 : 215 MW (pas tous référencés)
- Green500 : 590 MW (Nov. 14 – référence complète)



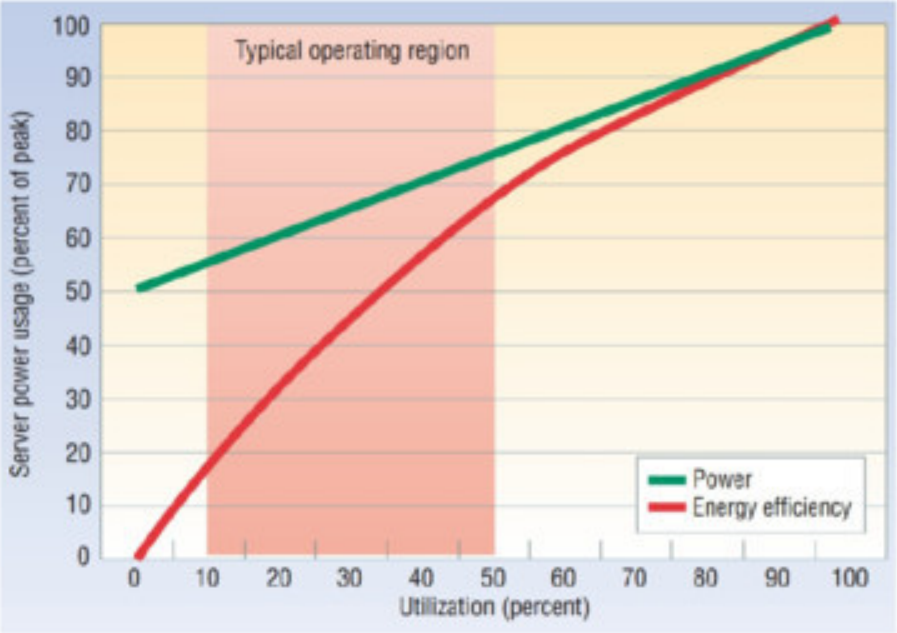


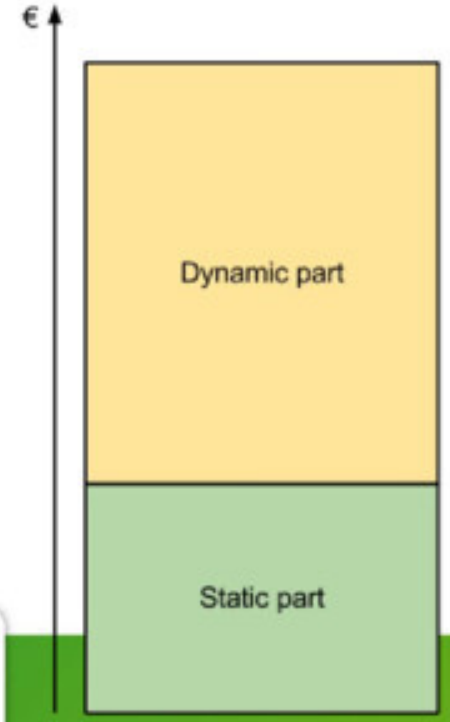
Proportionalité Energétique

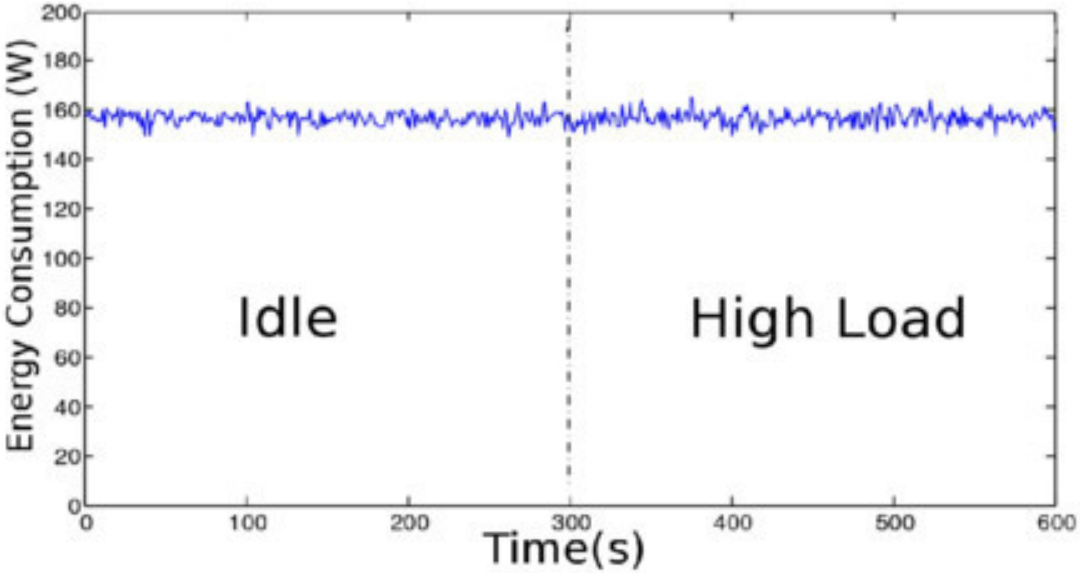
Luiz André Barroso and Urs Hölzle, « The case for Energy-Proportional Computing », IEEE Computer, 2007

Serveurs : Consommation *idle* (a vide / statique) importante – faiblement proportionnelle

Réseau : consommation constante insensible à l'usage





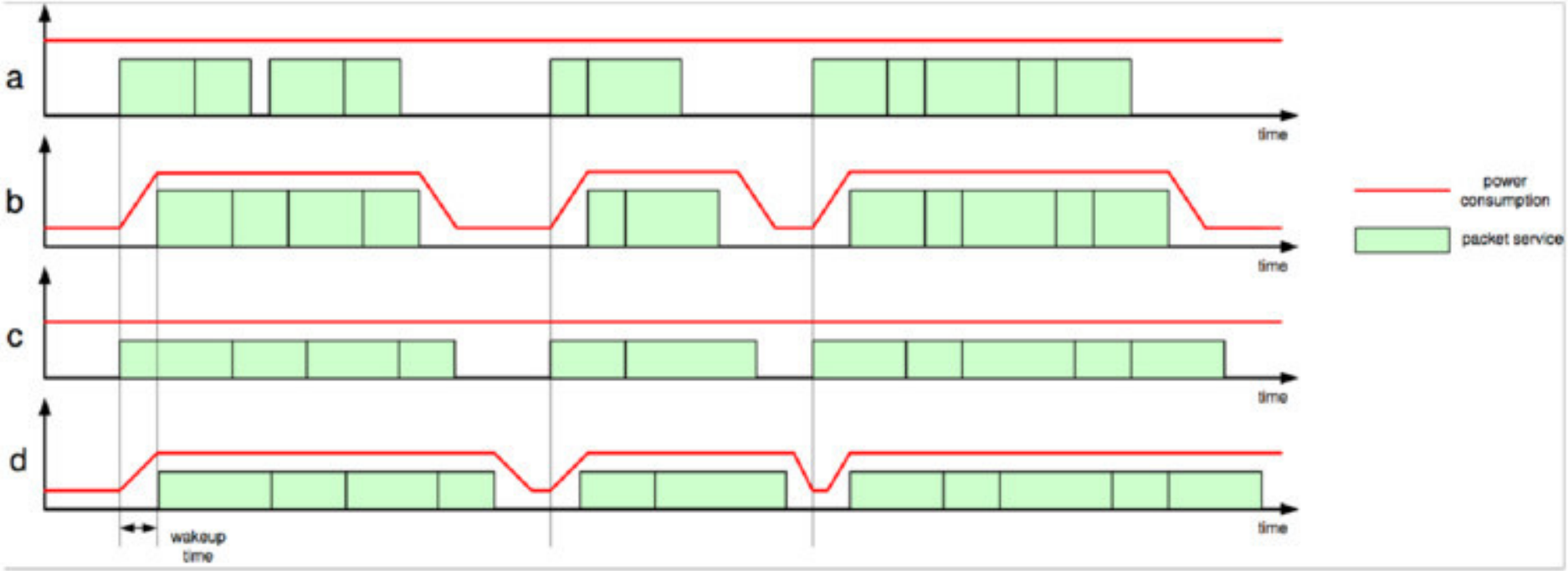


Les leviers verts !!!

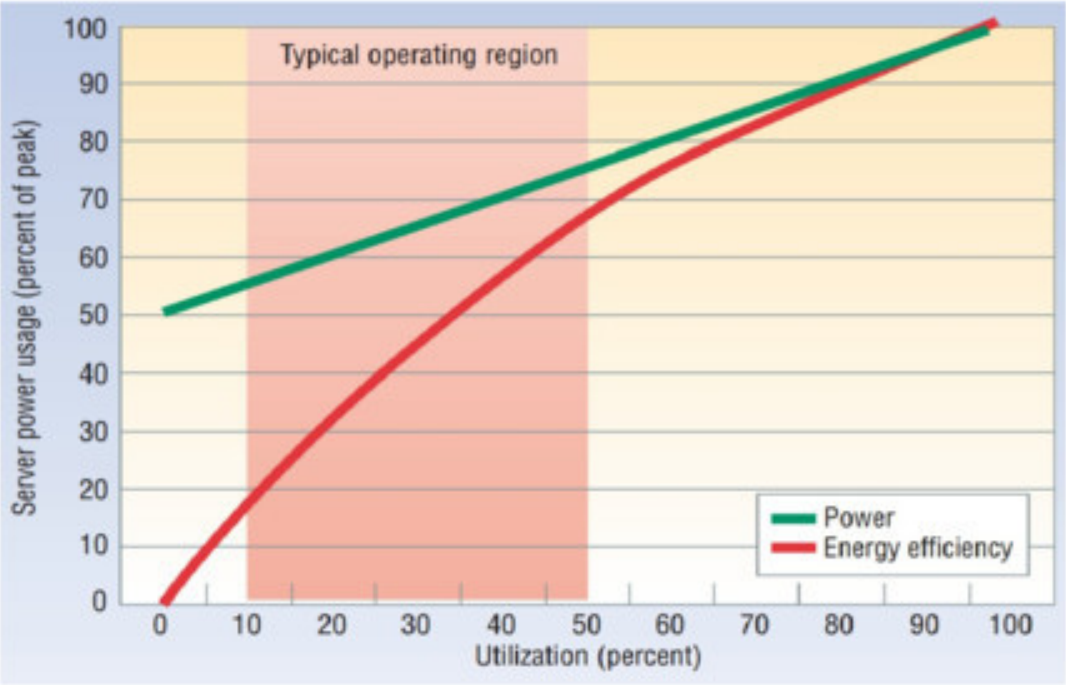
- **Extinction** (*shutdown*) : réduire le nombre de ressources alimentées inutilement
- **Dimensionner** (*slowdown*) : adapter les performances des ressources aux réels besoins
- **Optimiser** : modifier les applications et les services pour les rendre plus « verts »
- **Consolider / Aggréger** : relocaliser / regrouper des services et applications sur un nombre réduit de ressources physiques

Exemple d'applications de leviers dans les réseaux de communications

[Bolla 2011] : no Green, shutdown, slowdown, shutdown+slowdown

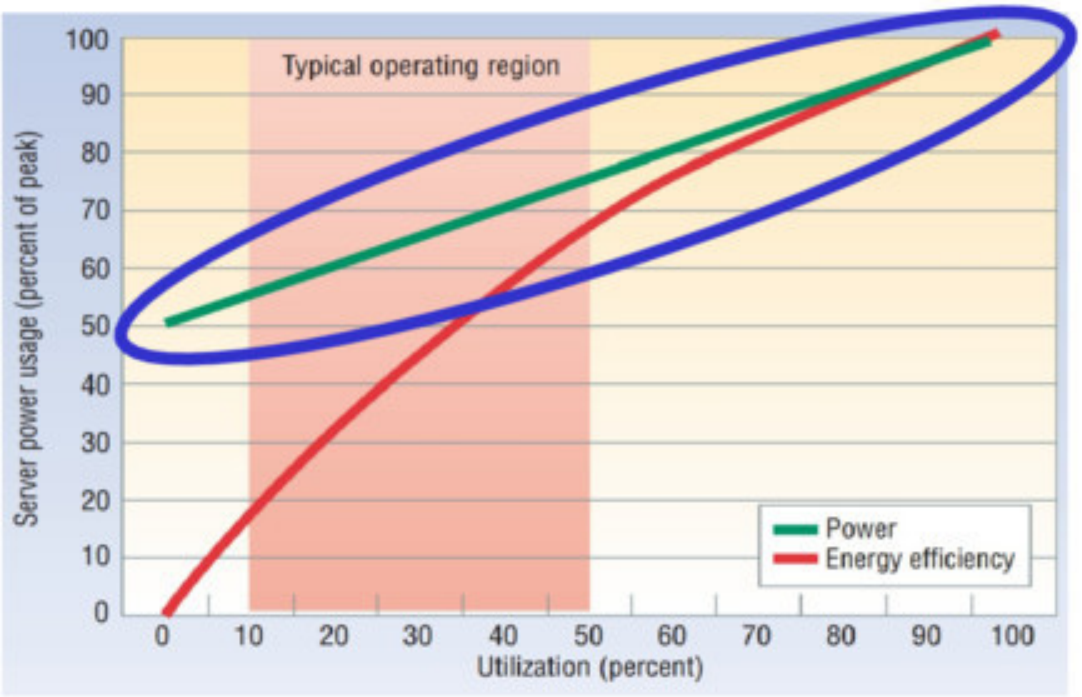


Challenges scientifiques : exposé d'aujourd'hui



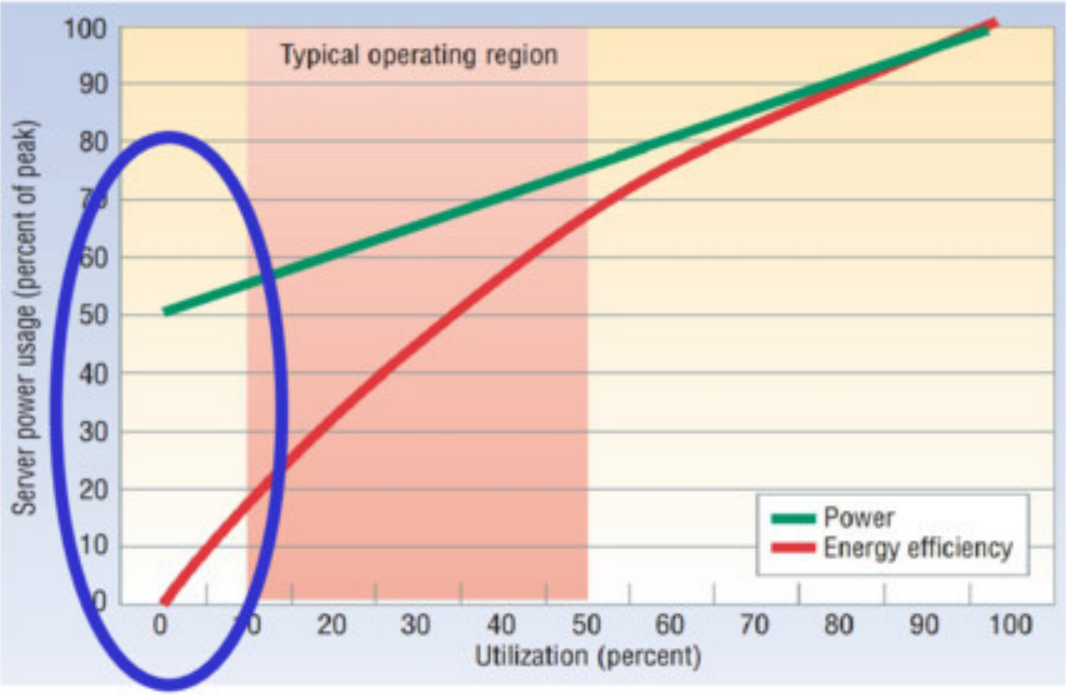
Challenges scientifiques : exposé d'aujourd'hui

- Mesurer



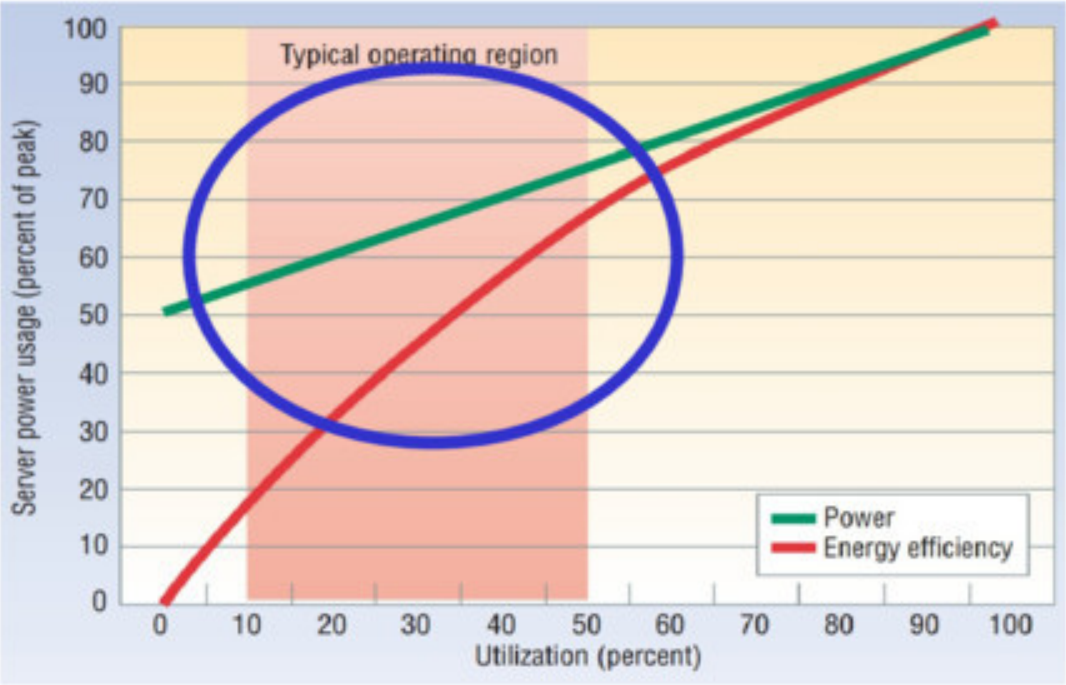
Challenges scientifiques : exposé d'aujourd'hui

- Mesurer
- ON/OFF



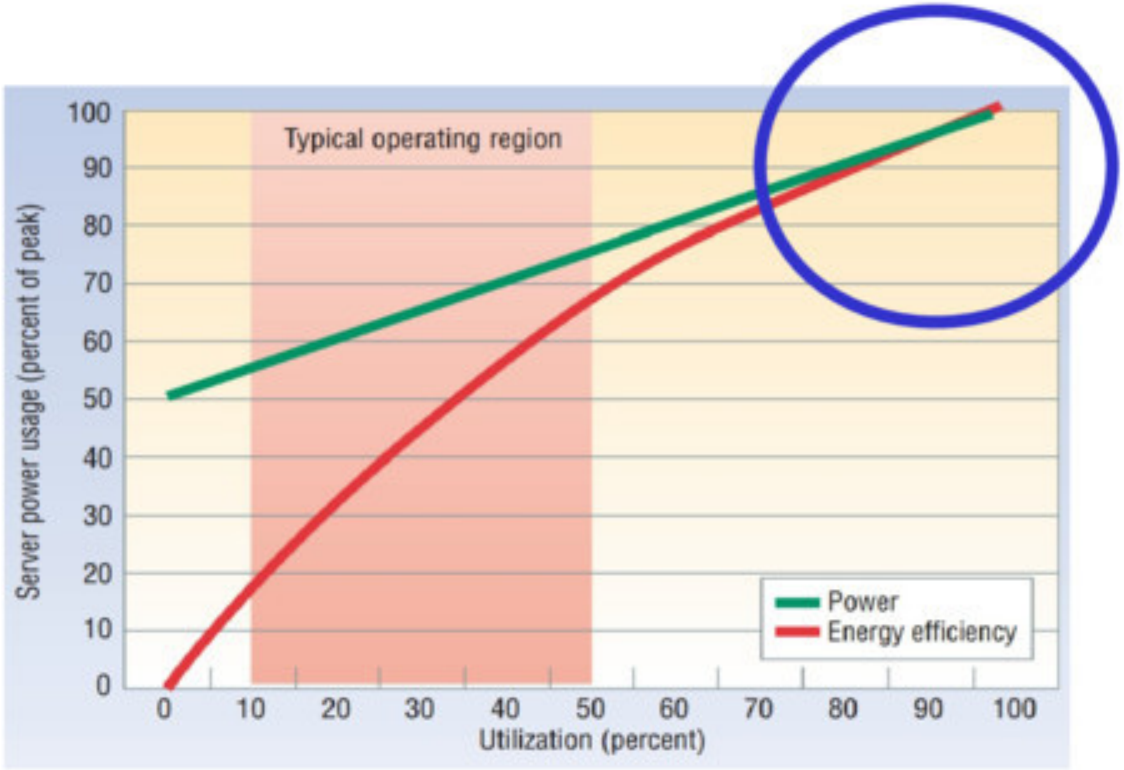
Challenges scientifiques : exposé d'aujourd'hui

- Mesurer
- ON/OFF
- Dimensionner



Challenges scientifiques : exposé d'aujourd'hui

- Mesurer
- ON/OFF
- Dimensionner
- Consolider



Une multitude d'approches !

Level	Technology	Existing work
Node	Sleep state	[Chase et al. 2001] [Gunaratne et al. 2005] [Leangsuksun et al 2006]
	DVFS	[Snowdon et al. 2005] [Kappiah et al. 2005] [Wang et al. 2010] [Lim et al. 2006] [Chen et al. 2005] [Yan et al. 2005]
	Software improvements	[EcoInfo 2011] [LessWatts 2010] [Steele 1998]
	Hardware capabilities	[Carrera et al. 2003]
Infrastructure	Green sources	[Greenpeace 2011] [Figuerola et al. 2009]
	Thermal management	[Patel et al. 2002] [Sharma et al. 2005] [Merkel and Bellosa 2006]
	Workload consolidation	[Chase and Doyle 2001] [Doyle et al. 2003] [Urgaonkar et al. 2008] [Verma et al. 2008] [Kusic et al. 2008] [Jung et al. 2009] [Srikantaiah et al. 2008] [Freeh et al. 2005] [de Langen and Juurlink 2006] [Lefèvre and Orgerie 2009]
	Task scheduling	[Jejurikar and Gupta 2006] [Mishra et al. 2003] [Yang et al. 2009] [Zhuo and Chakrabarti 2008] [Yao et al. 1995] [He and Ji Kim 2004] [Fan et al. 2001] [Merke [Chen et al. 2008 20
Virtualized environments	Virtual machines	[Nathuji and Sch et al. 2007] [Tal [Torres et al. 200 Gardner 2005] [He
	VM migration	[Clark et al. 2005] 2006] [Voorslu
	Cloud level	[Nurmi et al. 20 2008] [Barham et et al. 2002] [Liu et al. 2008] [Jung and Orge

Un écosystème de recherches..

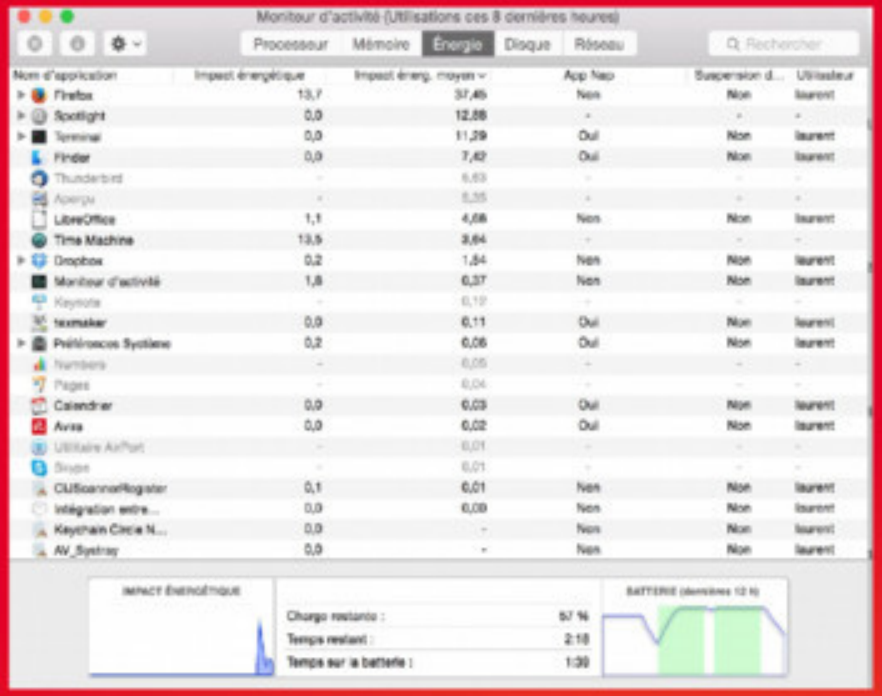
mais beaucoup restent à un niveau théorique et modèle et ne sortent pas des laboratoires

Exemples : calculs et réseaux

Approach	Technology	Existing work
Hardware	Improving and re-engineering	[Gupta and Singh 2007b] [Ananthanarayanan and Katz 2008] [Roberts 2009]
Shutdown	On/off and sleeping	[Gupta and Singh 2003] [Nedevschi et al. 2008] [Gupta et al. 2004] [Bolla et al. 2011] [Hu et al. 2011] [Chiaraviglio et al. 2008] [Yamanaka et al. 2010] [Soteriou and Peh 2003] [Christensen et al. 2010] [Gupta and Singh 2007a]
	Proxying	[Irish et al. 1998] [Christensen et al. 2004] [Jimeno et al. 2008] [Sabhanatarajan et al. 2008] [Agarwal et al. 2009] [Agarwal et al. 2010] [Gunaratne et al. 2005]
Slowdown	Rate adaptation	[Zhang et al. 2008] [Gunaratne et al. 2006] [Nedevschi et al. 2008] [Alonso et al. 2004]
Coordination	Network protocols	[Irish et al. 1998] [Wang and Singh 2004] [Cianfrani et al. 2010] [Zhang et al. 2010] [Van Heddeghem et al. 2010] [Chabarek et al. 2008] [Gelenbe and Silvestri 2009] [Blackburn and Christensen 2009]
	Clean-slate	[Baldi and Ofek 2009] [Hayenga et al. 2009] [Tucker 2010] [Baliga et al. 2009]
	Frameworks	[Chiaraviglio et al. 2009] [Shang et al. 2006] [Chiaraviglio et al. 2009] [Chiaraviglio and Matta 2010] [Steinder et al. 2008] [Restrepo et al. 2009] [Wang et al. 2008] [Tzanakaki et al. 2011] [Orgerie and Lefèvre 2011]

« On the Road to Energy-Efficient Computing and Network: Where Exactly are We? » - Anne-Cécile Orgerie, Marcos Dias de Assuncao, Laurent Lefevre ACM Compuyting Surveys - December 2014

EcoInfo


Moniteur d'activité (Utilisations ces 8 dernières heures)
Processeur
Mémoire
Énergie
Disque
Réseau
Rechercher
Nom d'application
Impact énergétique
Impact énerg. moyen
App Nap
Suspension d...
Utilisateur
Firefox 13,7 37,45 Non Non laurent
Spotlight 0,0 12,88 - - -
Terminal 0,0 11,29 Oui Non laurent
Finder 0,0 7,42 Oui Non laurent
Thunderbird - 6,63 - - -
Aperçu - 5,35 - - -
LibreOffice 1,1 4,08 Non Non laurent
Time Machine 13,5 3,64 - - -
Dropbox 0,2 1,54 Non Non laurent
Moniteur d'activité 1,8 0,37 Non Non laurent
Keynote - 0,19 - - -
texmaker 0,0 0,11 Oui Non laurent
Préférences Système 0,2 0,06 Oui Non laurent
Numbers - 0,05 - - -
Pages - 0,04 - - -
Calendrier 0,0 0,03 Oui Non laurent
Avira 0,0 0,02 Oui Non laurent
Utilitaire AirPort - 0,01 - - -
Skype - 0,01 - - -
CUSlicannerRegister 0,1 0,01 Non Non laurent
Intégration entre... 0,0 0,00 Non Non laurent
Keychain Circle N... 0,0 - Non Non laurent
AV_Systray 0,0 - Non Non laurent
IMPACT ÉNERGÉTIQUE
Charge restante : 57 %
Temps restant : 2:18
Temps sur la batterie : 1:30
BATTERIE (dernières 12 h)

Mesurer !!!

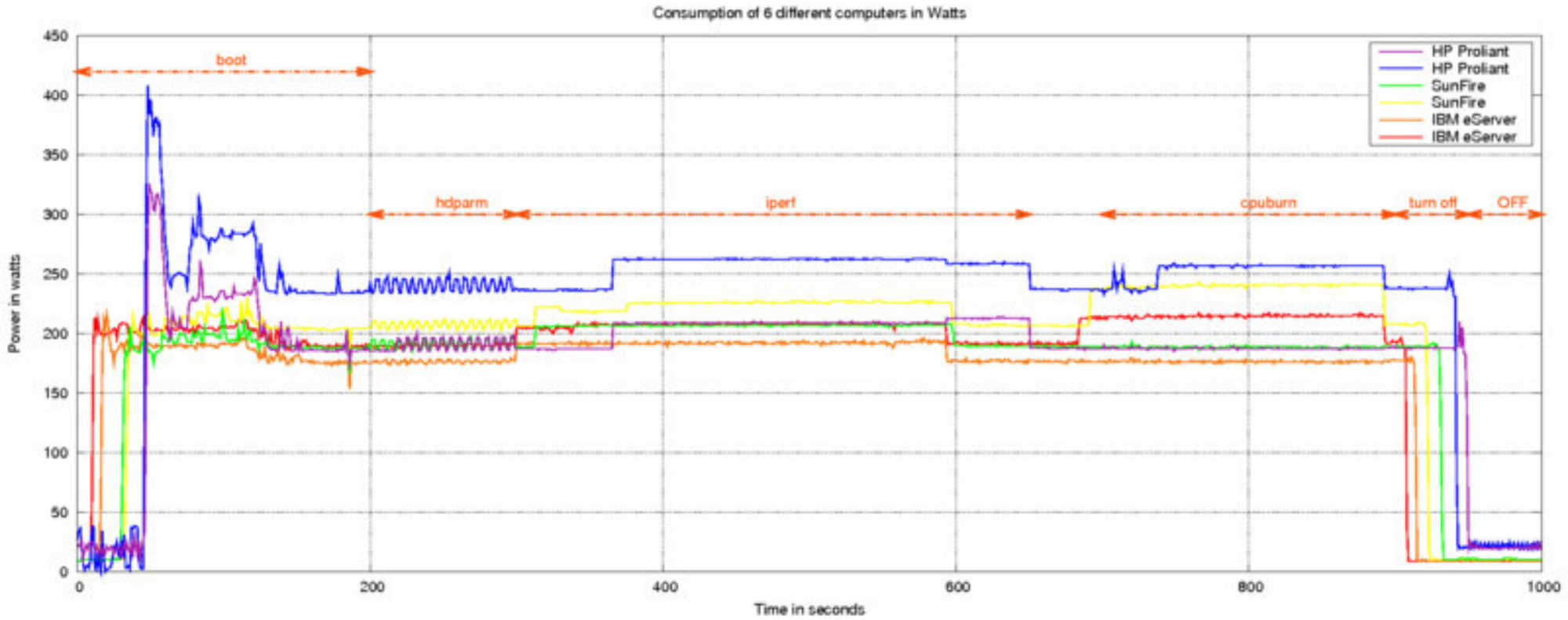
Inria
INVENTEURS DU MONDE NUMÉRIQUE

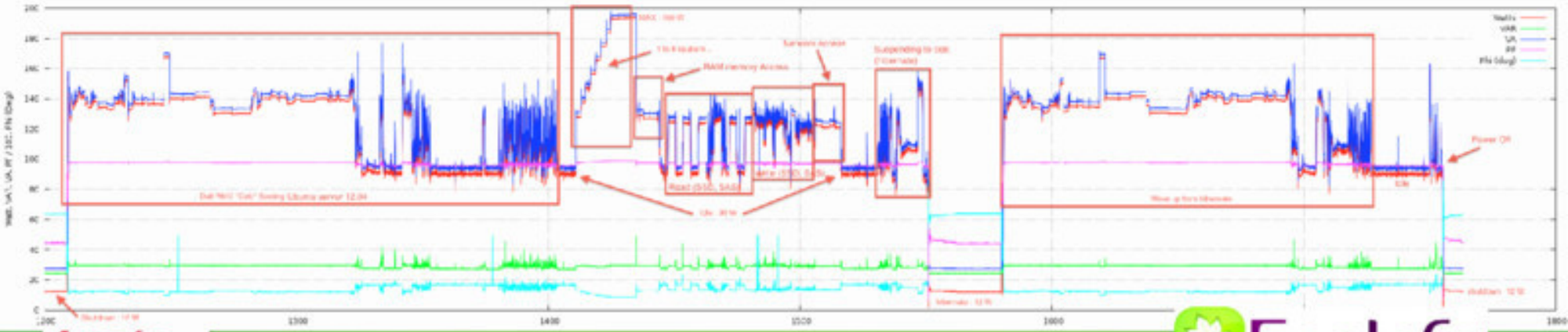
1er challenge : mesurer l'énergie consommée en phase d'usage

- Besoin d'équipements spécifiques : capteurs, wattmètres…
- Capteurs logiciels

Ex : Mesure d'un serveur avec un PDU Raritan

Mesurer finement !

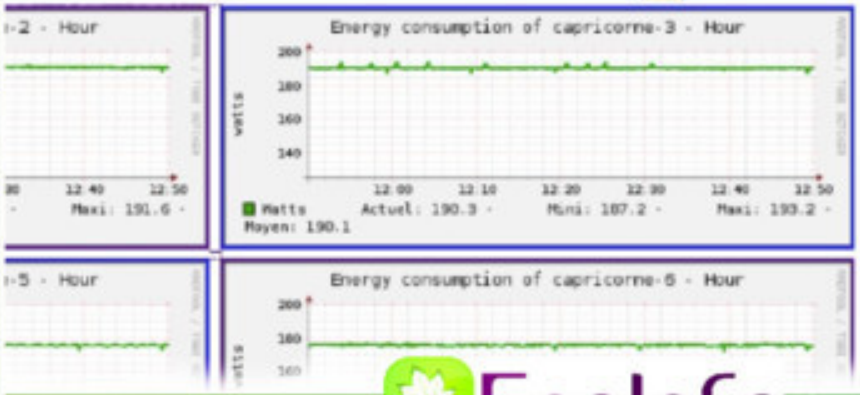




inria

EcoInfo

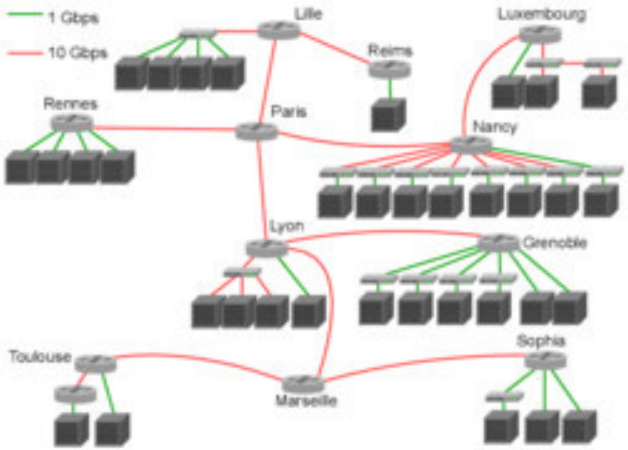
Donner des informations utiles

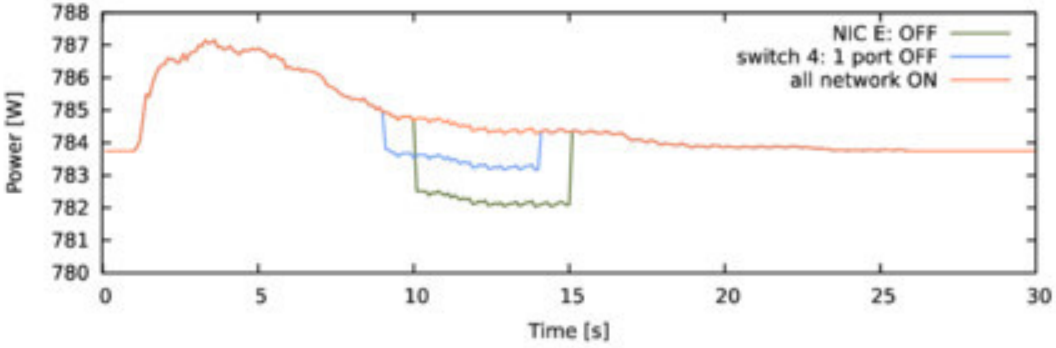


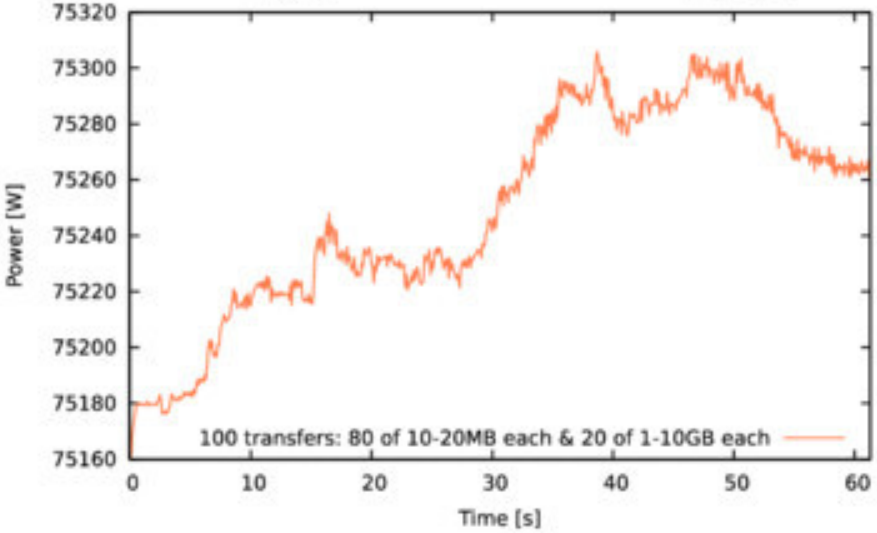
EcoInfo

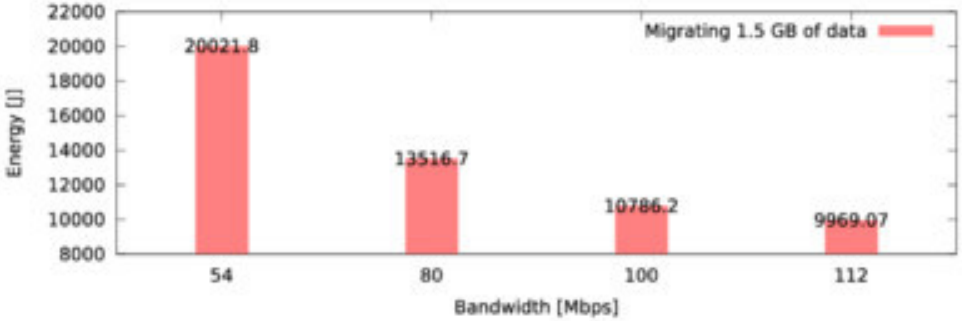
Mesurer la consommation d'énergie dans les réseaux ?

- Besoin de mesures de bout en bout
- Evaluer l'impact des équipements, protocoles et services
- Difficultés: besoin d'outils de simulation : ex : Ecofen
- Evaluer les leviers verts dans les réseaux





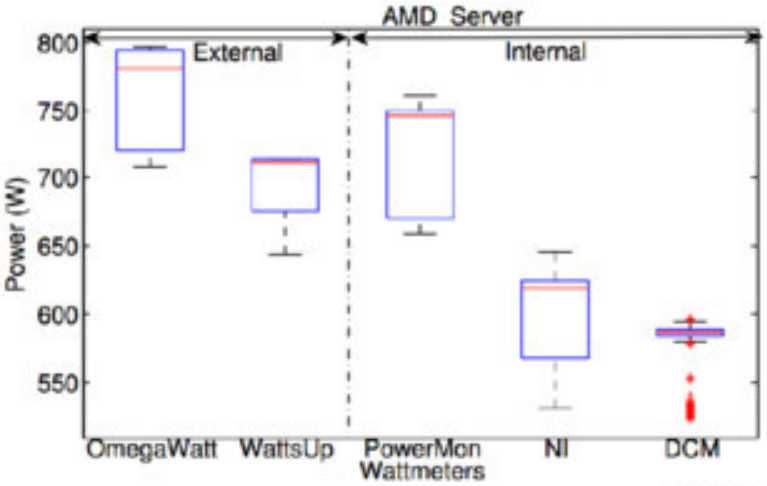


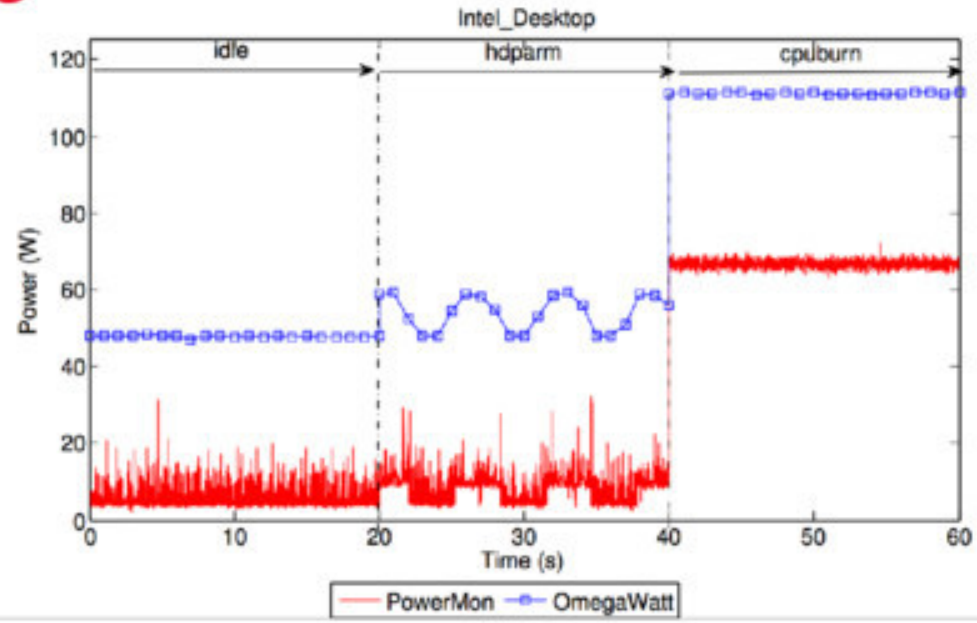


[CloudNet14] : Cornea, Orgerie, Lefevre

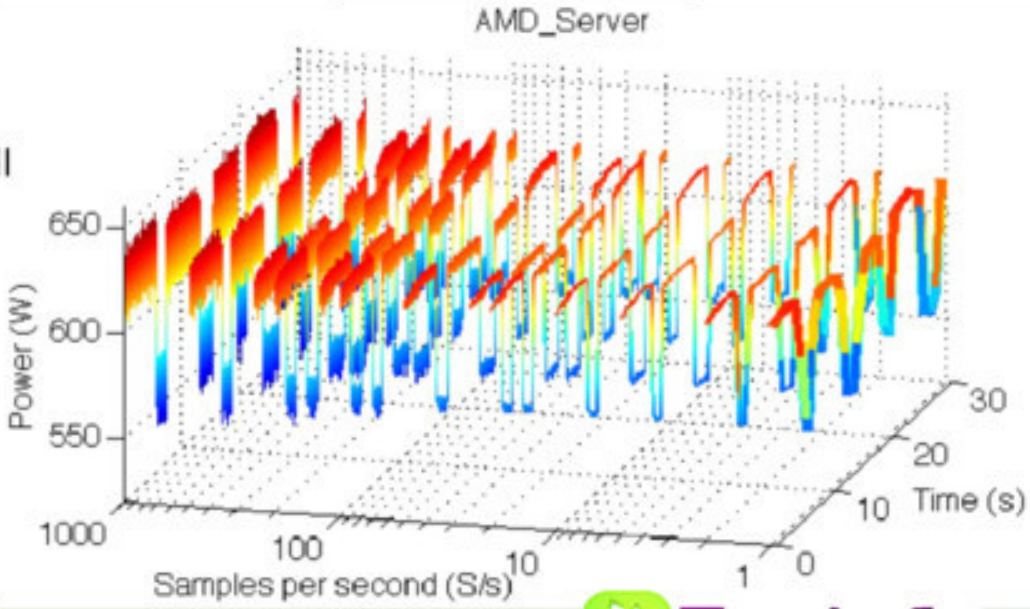
Problème ouvert : analyser et faire confiance aux mesures des wattmètres

- Fréquence / précision





- CPUBurn / NI



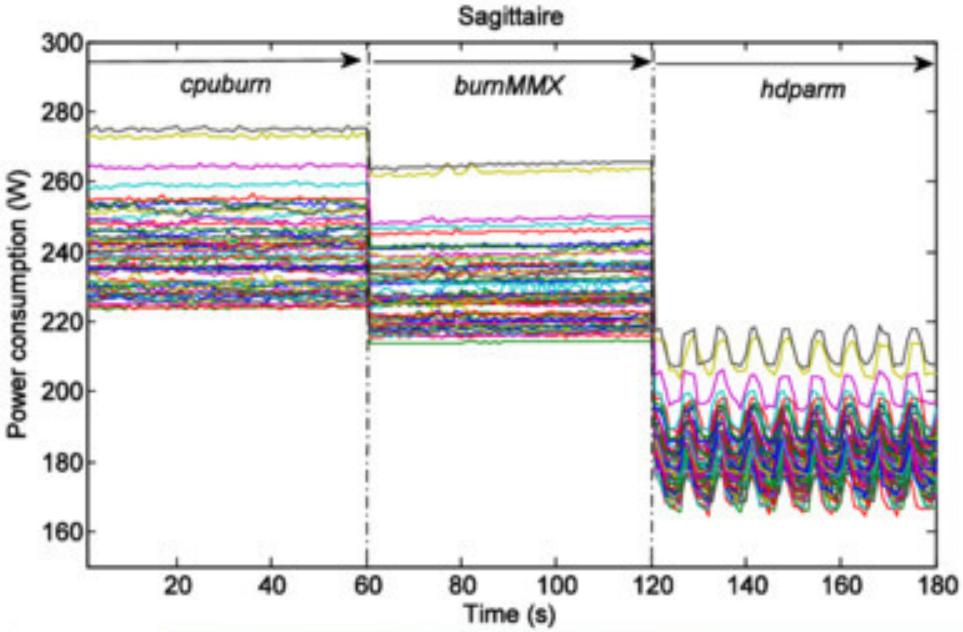
M. Diouri, M. Dolz, O. Glück, L. Lefevre, P. Alonso, S. Catalan, R. Mayo, E. Quintan-Orti. **"Solving some Mysteries in Power Monitoring of Servers: Take Care of your Wattmeters! "**, *EE-LSDS 2013 : Energy Efficiency in Large Scale Distributed Systems conference* , Vienna, Austria, April 22-24, 2013

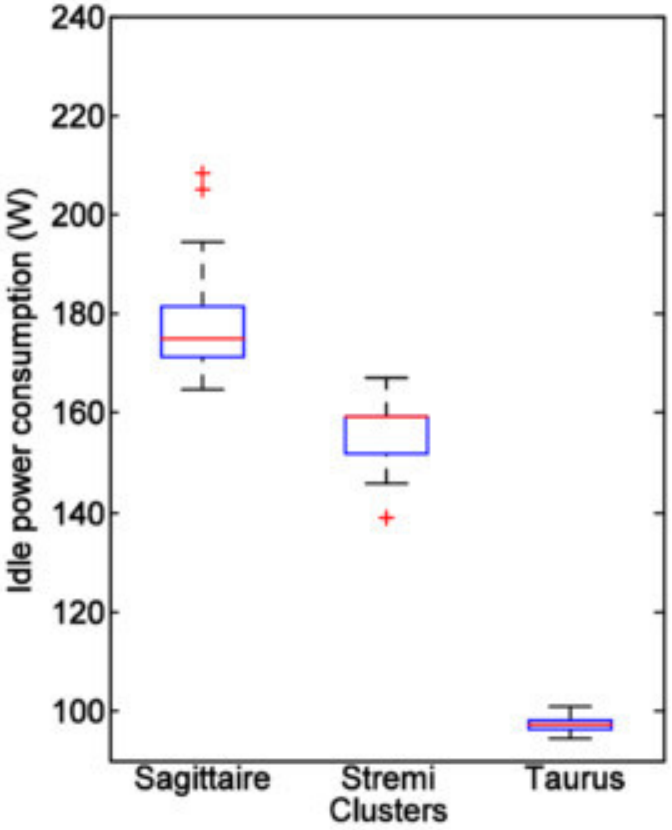
Inria

EcoInfo

Challenge : l'homogénétité en consommation n'existe pas !

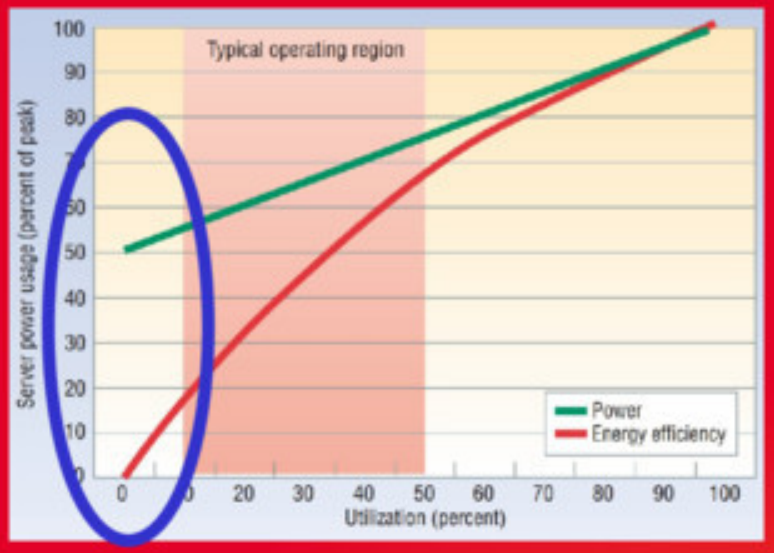
- Dependance à la technologie
- Même performances (flops, i/o par sec..) mais efficacités énergétique différentes (flops par watt)
- Processeur : principal responsable
- Besoin de tout mesurer !





Mohammed el Mehdi Diouri, Olivier Glück, Laurent Lefèvre and Jean-Christophe Mignot. **"Your Cluster is not Power Homogeneous: Take Care when Designing Green Schedulers!"**, *IGCC2013 : International Green Computing Conference*, Arlington, USA, June 27-29,

EcoInfo



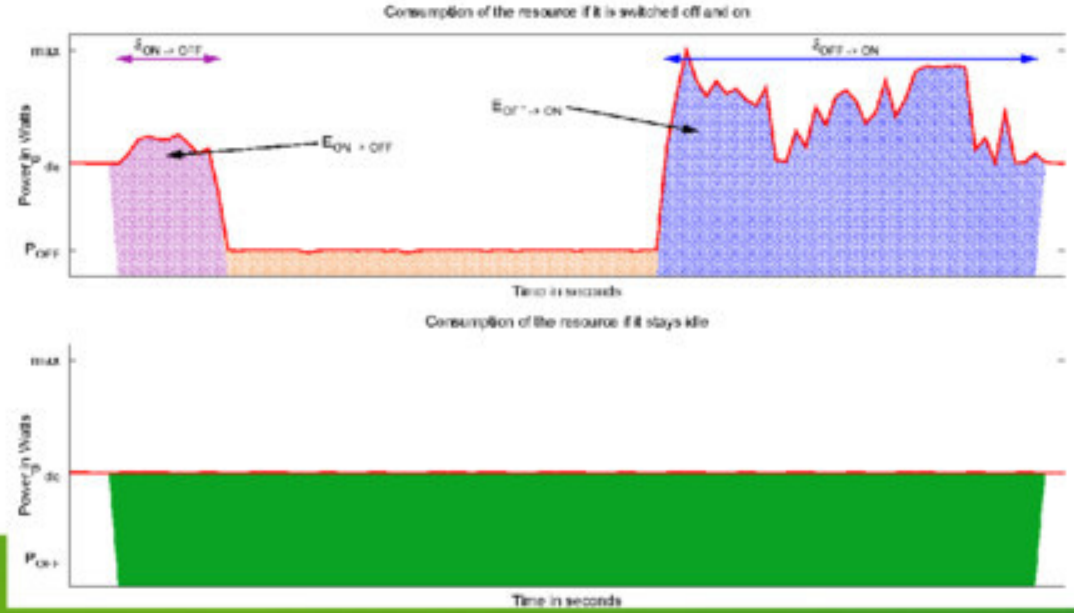
ON /OFF : Eteindre et allumer

Inria
INVENTEURS DU MONDE NUMÉRIQUE

Modèles On/Off

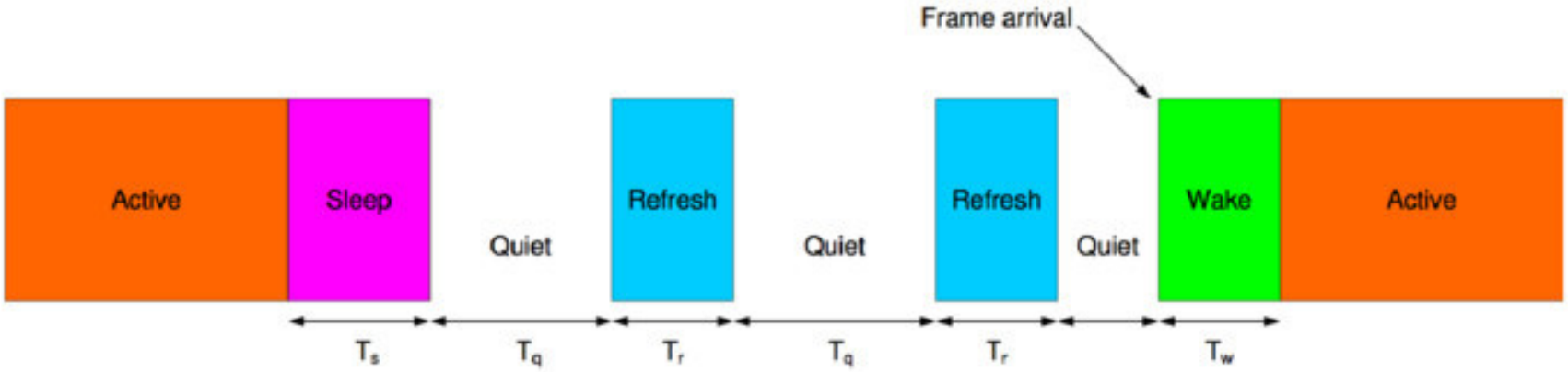
- Eteindre les ressources inutiles :
  - Gros grain : serveurs, disques, équipements réseaux
  - Grain moyen : ports réseaux, fibres, coeurs de calcul, co-processeurs
  - Grain fin : bancs mémoires, composants

- Première approche la plus explorée
- Favorise le développement de *sleep states* (veille) et de réveil rapide
- Combat les inconvénients de consommation statique



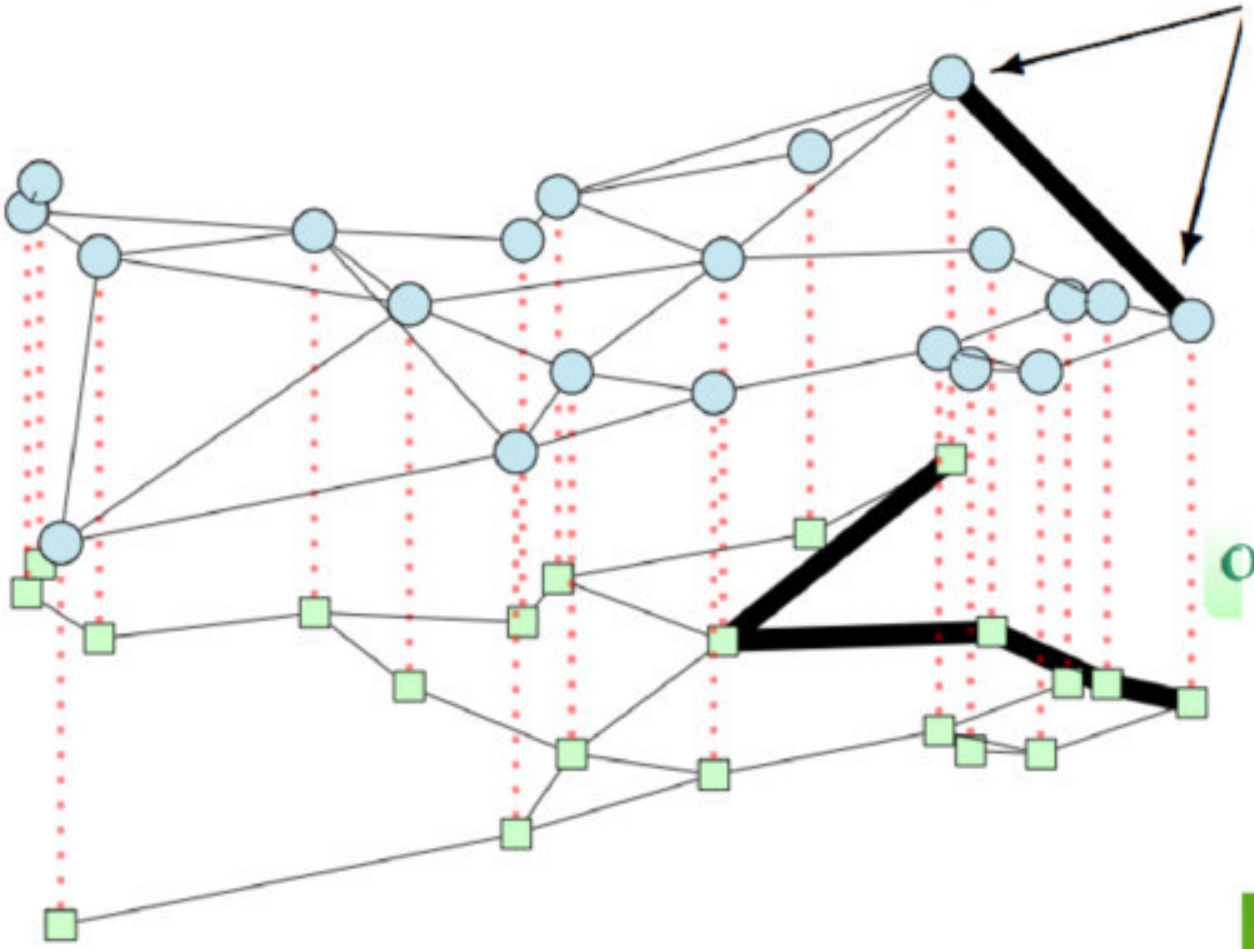
On/Off dans les réseaux : le succès du « *Energy Efficient Ethernet* »

- Low Power Idle == IEEE 802.3az
- Extinction de composants électroniques (transmetteurs Ethernet) lors de période de faible trafic
- Off de 3 récepteurs sur 4
- Cycle de rafraichissement / signal d'alerte pour réveil rapide
- Gain dépend de l'usage
- Produits existants



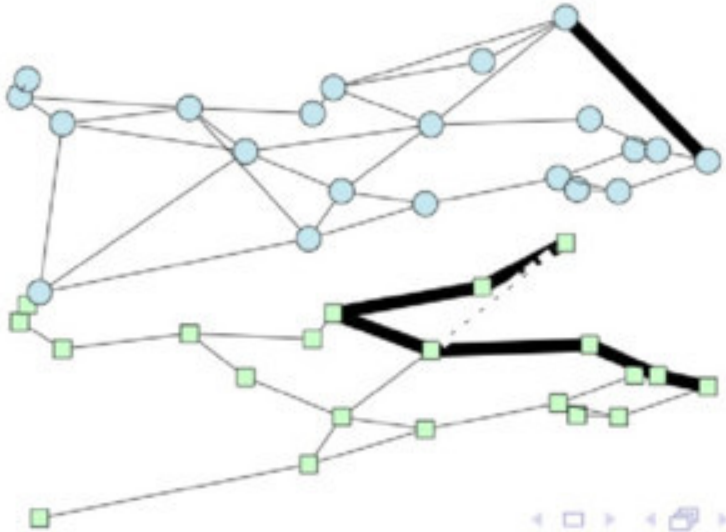
On/Off dans les réseaux de coeur

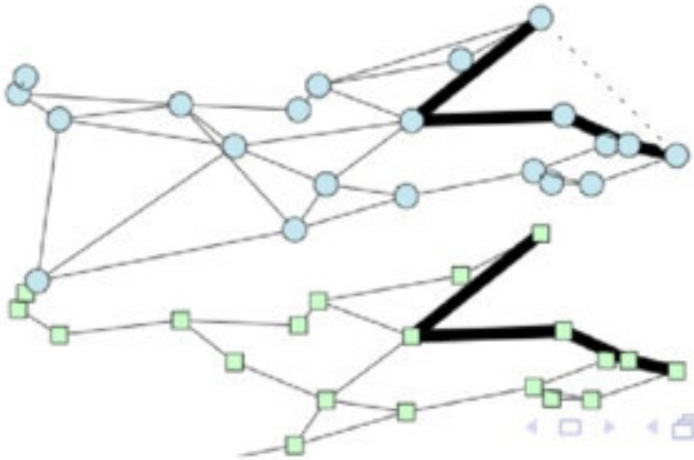
	Power consumption [Watt]
1 Gbps port	7 W
2.5 Gbps port	15 W
10 Gbps port	34 W
40 Gbps port	160 W
100 Gbps port	360 W
400 Gbps port	(1236 W)
1 Tbps port	(2794 W)



Réseaux de coeur efficaces en énergie

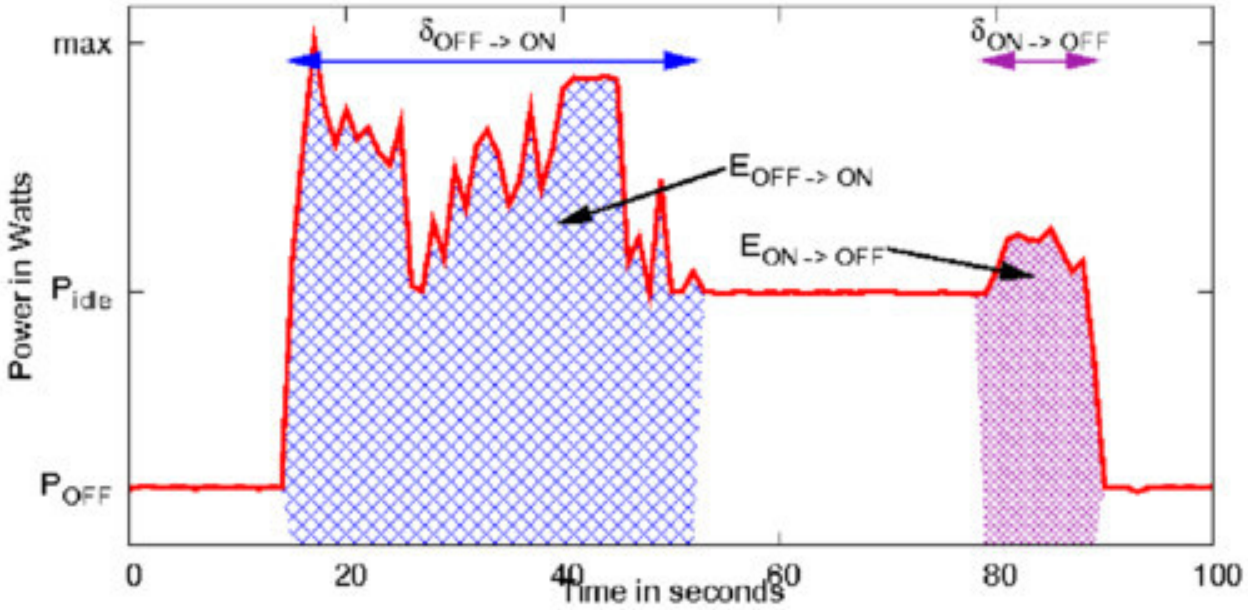
- Efficacité énergétique focalisée sur la couche IP (évite d'impacter les amplificateurs optiques qui sont très lents à rallumer)
- Extinction des ports de routeurs : Off en fonction de l'usage des liens
- Re-routage de trafic en prenant en compte les risques de congestion
- Besoin d'intelligence dans les réseaux et de bonnes connaissances sur l'usage : approche basée sur SDN (Software Defined Networks)
- Le plus difficile : qui / quand rallumer ?
- Projet CHIST-ERA STAR



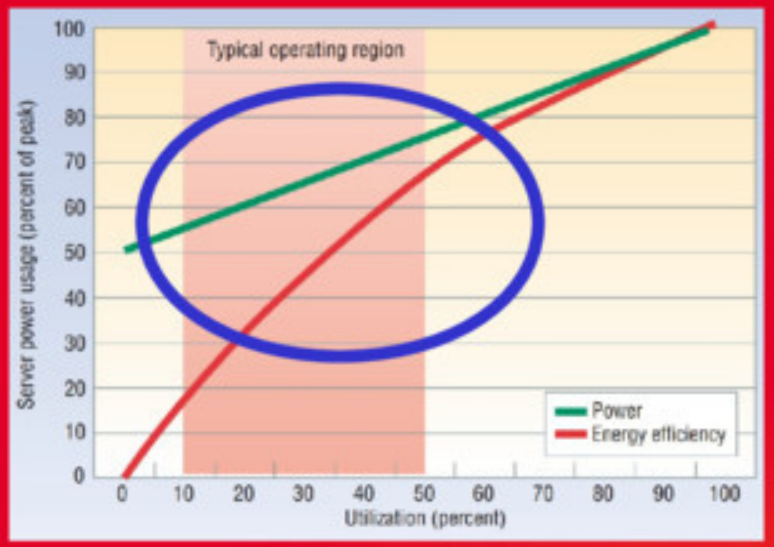


Challenge : des on/off agressifs ne sont pas toujours la meilleure solution

- Exploiter les creux entre des activités
- Seulement éteindre si les gains en énergie sont possibles
- Impact sur la durée de vie des équipements / liens avec le refroidissement
- Besoin de prédire les usages
- Systèmes à base de réservations.. même dans le « Cloud » : ex FSN XLCloud (« HPC As A Service ») : Blazar/Climate



EcoInfo

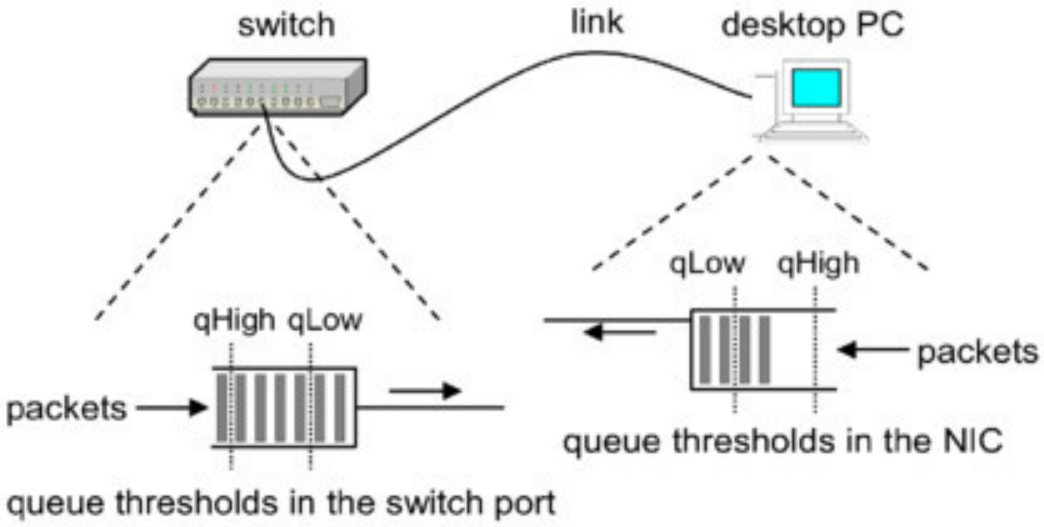


Ralentir / dimensionner

Inria
INVENTEURS DU MONDE NUMÉRIQUE

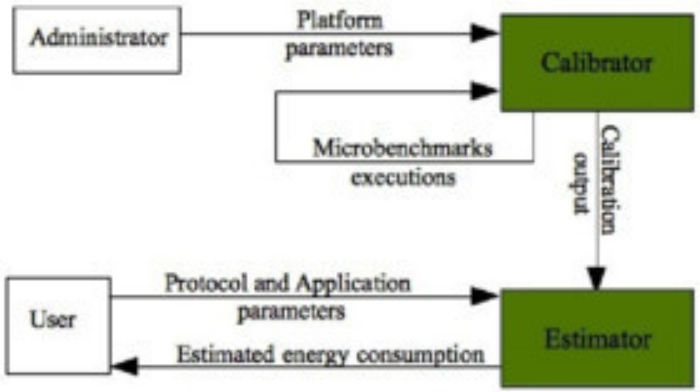
Adaptation dans le réseau

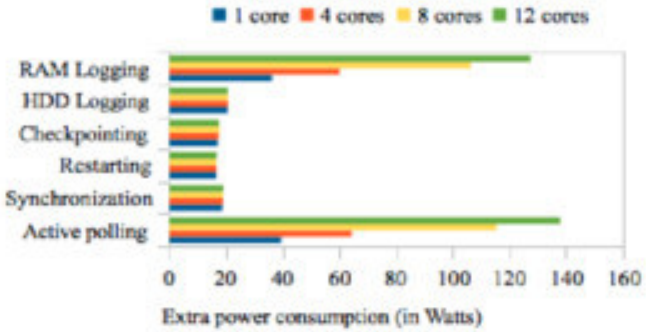
- Idée : redimensionner dynamiquement les capacités du réseau
- ALR (Adaptive Link Rate) : ajuster la bande passante d'un lien réseau en fonction du traffic actuel
- Comment définir les bons seuils ? Risques d'oscillations – incréments (10 Mb, 100 Mb, 1 Gb, 10Gb)
- Difficile : provoque de l'instabilité pour les applications (congestions ?)

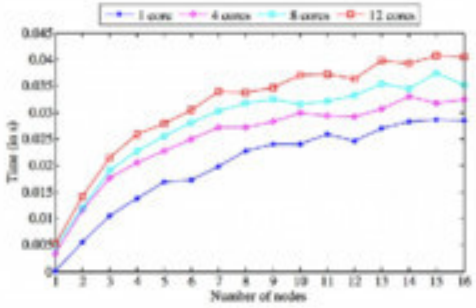


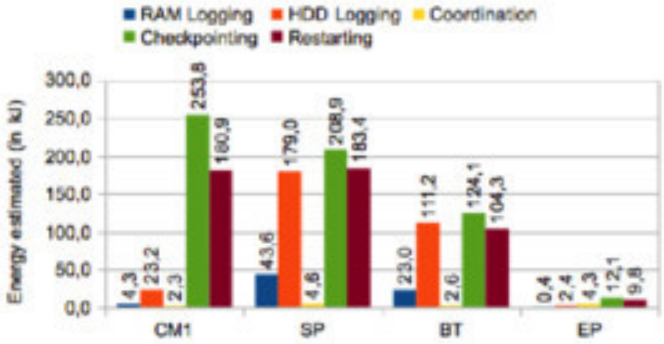
Services dans le HPC

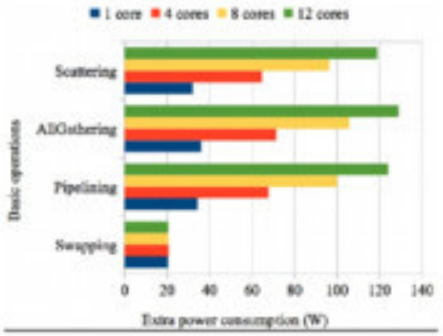
- Exemple services : résilience et diffusion
- 4 étapes : analyse de service, mesures, calibration et estimation
- But : Aider utilisateur à choisir le bon service en fonction des contraintes des applications

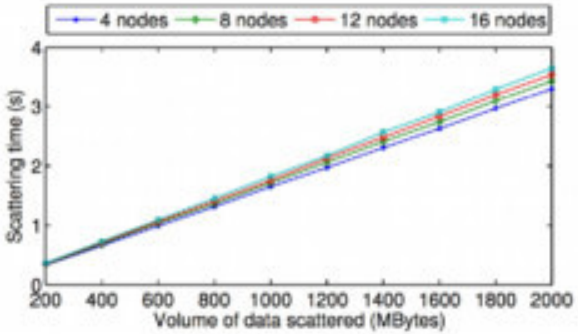


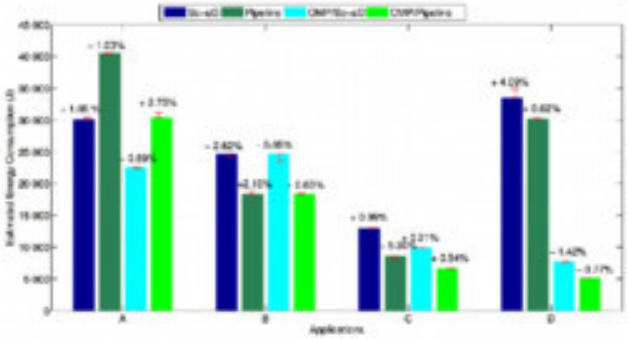










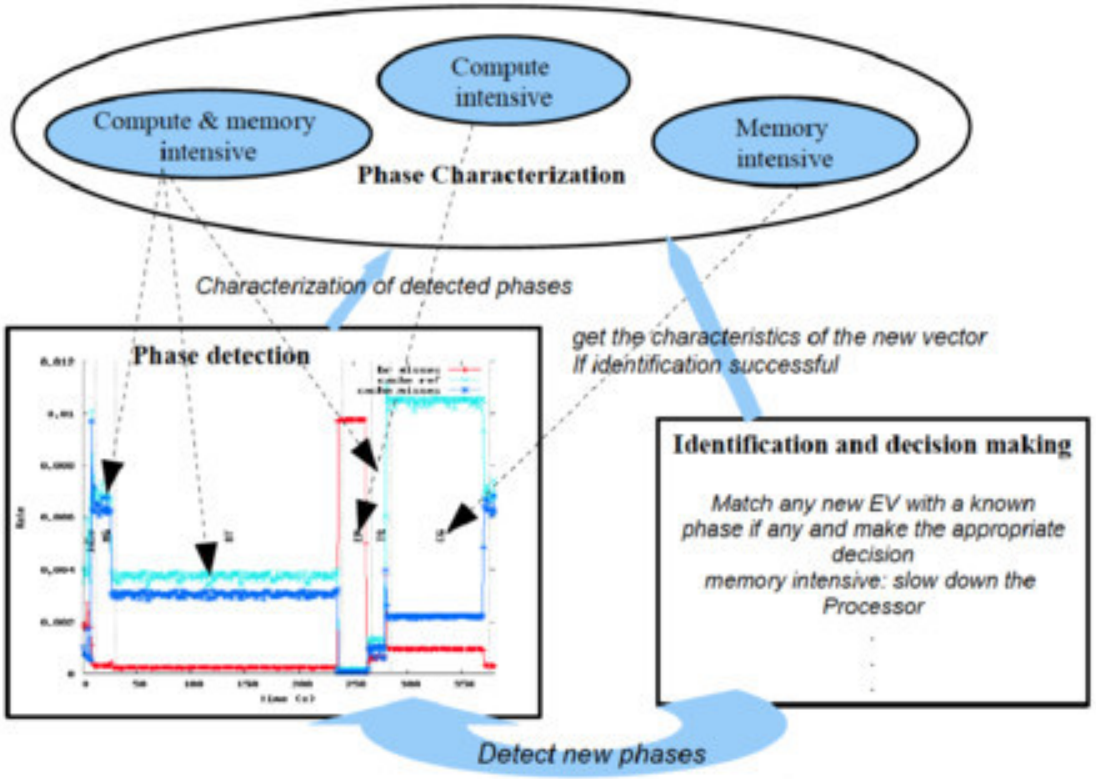


Redimensionnement autonomique ?

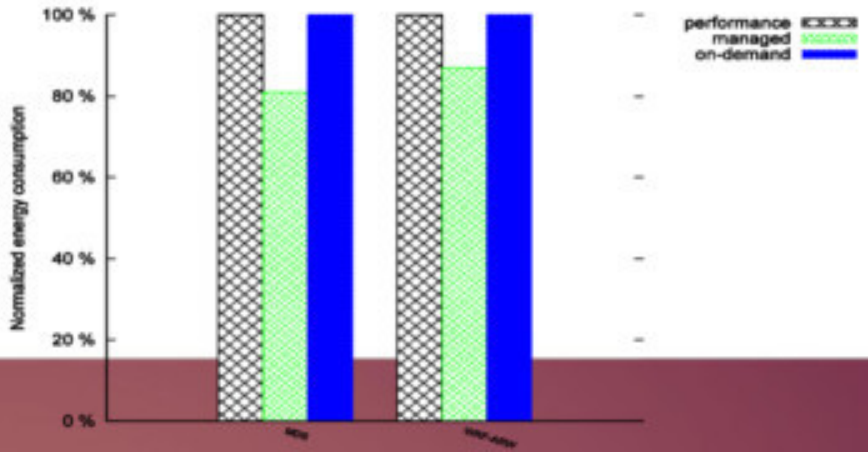
- Applications ont une utilisation non régulière des ressources
- Mais applications trop complexes à comprendre
- Besoin de frameworks qui appliquent les bons (multi-)leviers aux bons moments : mieux que DVFS/*Gouverneur on demand*

Approche :

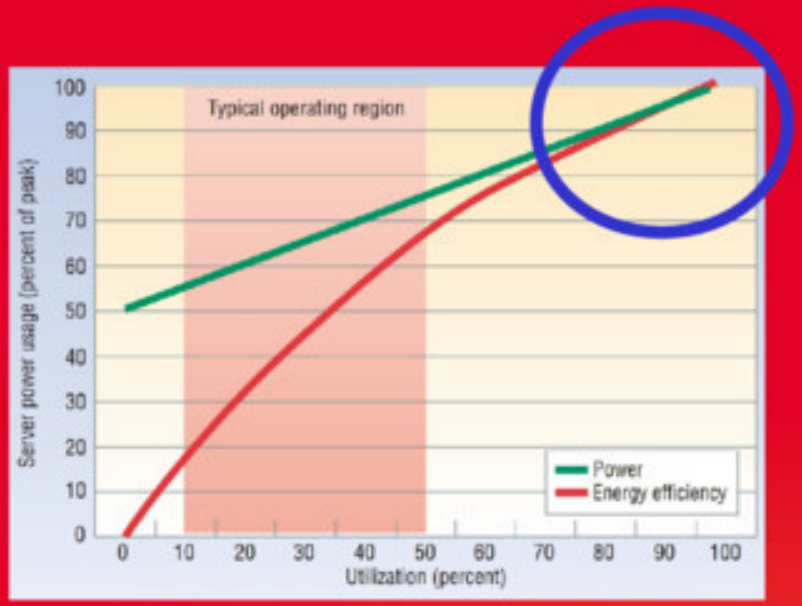
- Analyse ADN live de l'exécution des applications et services
- Détection des phases
- Caractérisation des applications
- Applications de leviers verts



Phase label	Possible reconfiguration decisions
compute intensive	switch off memory banks; send disks to sleep; scale the processor up; put NICs into LPI mode
memory intensive	scale the processor down; decrease disks or send them to sleep; switch on memory banks
mixed	switch on memory banks; scale the processor up send disks to sleep; put NICs into LPI mode
communication intensive	switch off memory banks; scale the processor down switch on disks
I/O intensive	switch on memory banks; scale the processor down; increase disks, increase disks (if needed)



EcoInfo

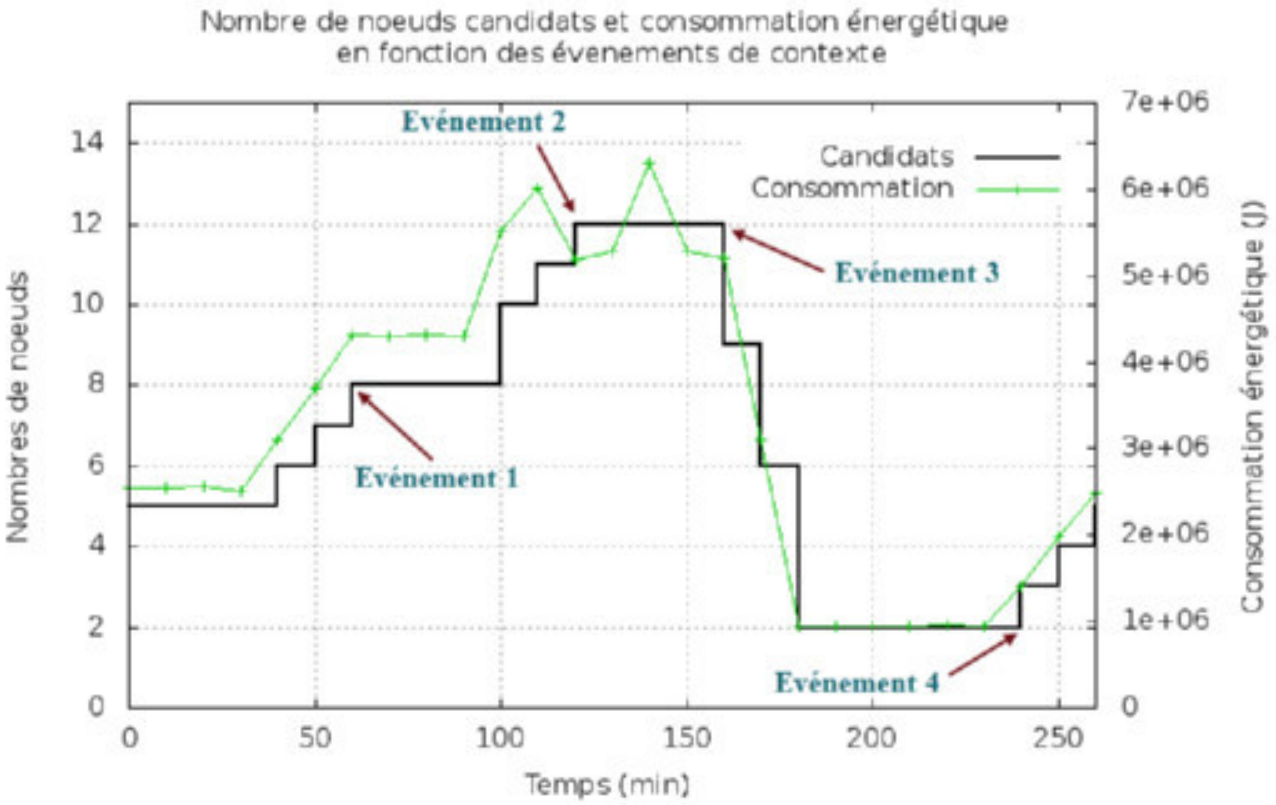


Consolider / aggréger

Inria
INVENTEURS DU MONDE NUMÉRIQUE

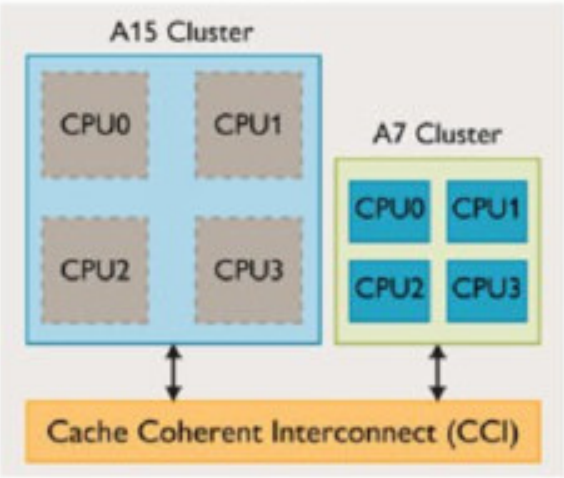
Energie et les gestionnaires

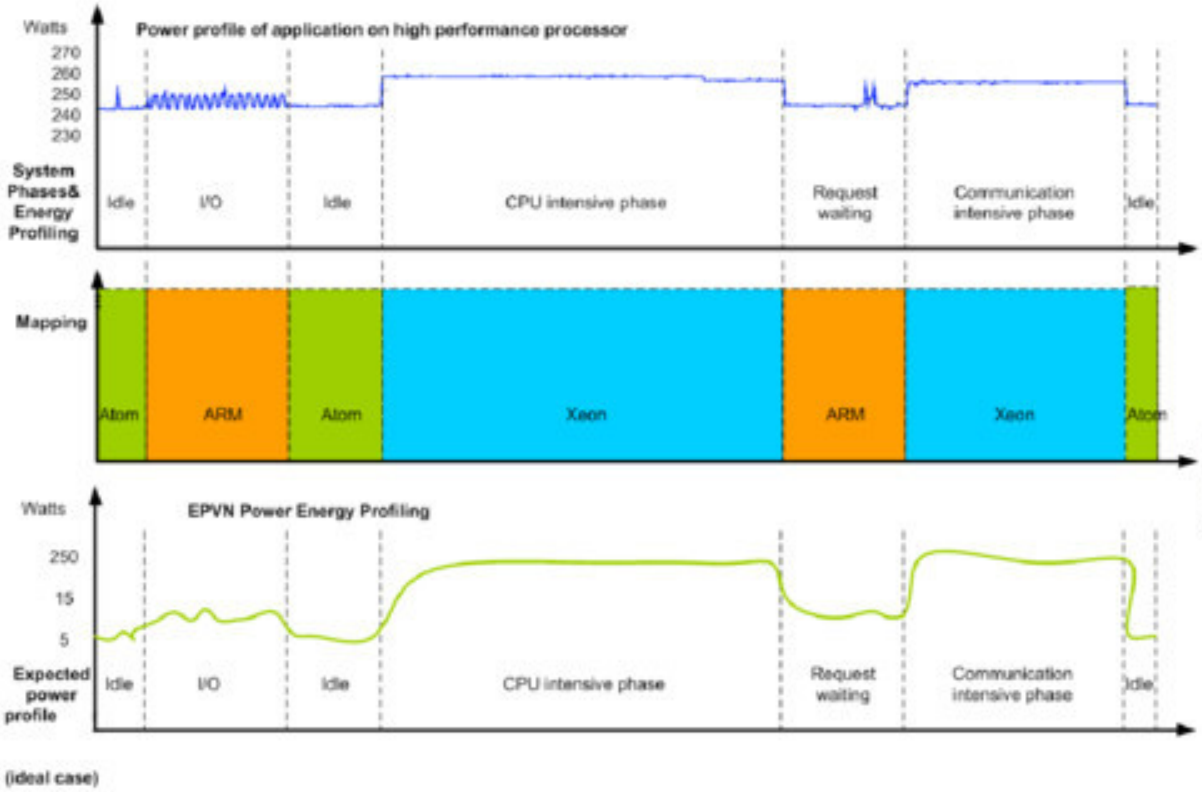
- Même si les nuages ne sont pas encore verts, certains principes des clouds peuvent nous aider
- Ordonnanceurs verts : où placer les taches de calcul et de stockage ? En lien avec extinction des ressources inutiles

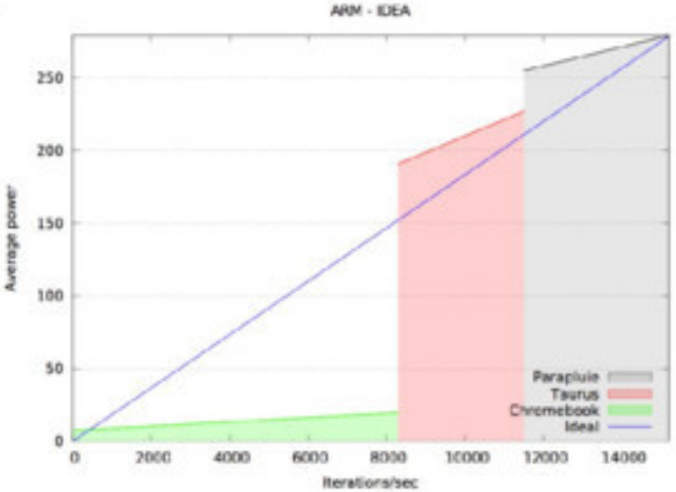


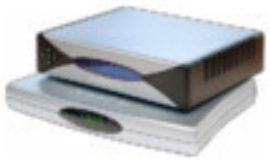
Energie et les gestionnaires

- Exemple des architectures hybrides : avec processeurs basse consommation, co-processeurs, GPUs...
- Placer les tâches aux bons endroits
- On retrouve de la proportionnalité énergétique !



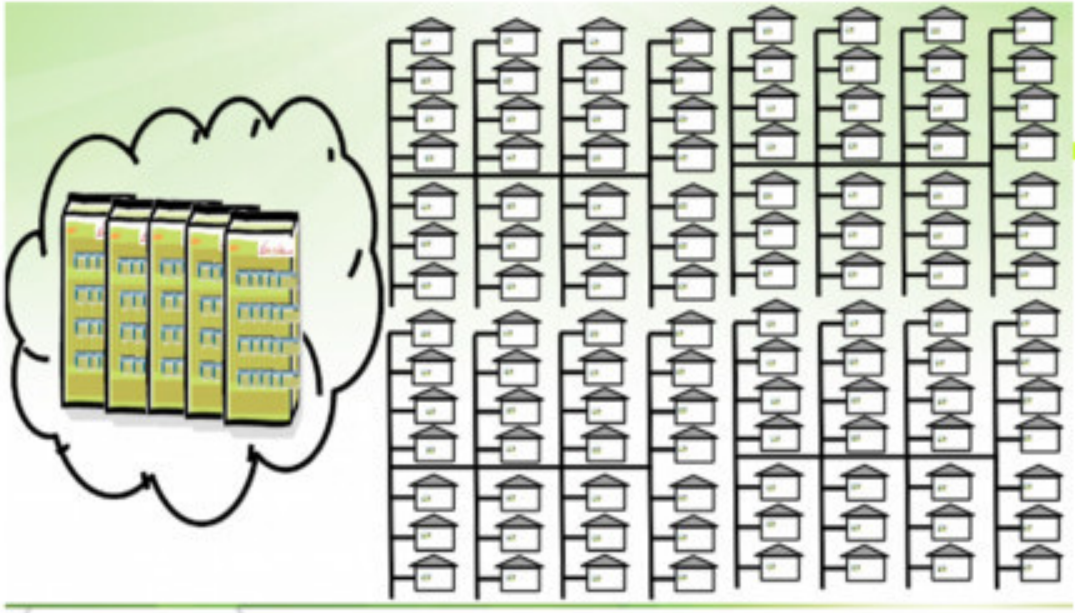


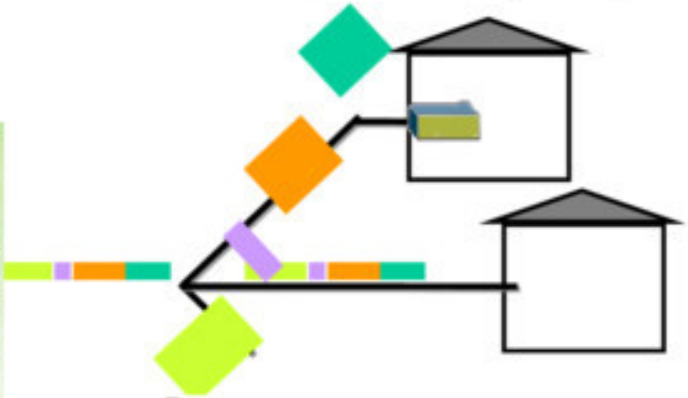


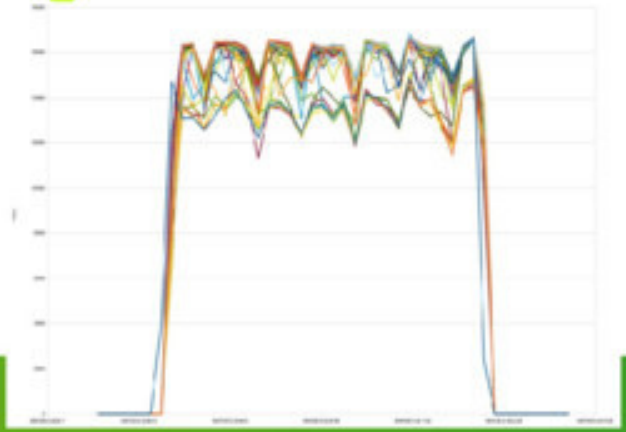


Virtual Home Gateway

- Passerelles à la maison : services multiples (données, voix, multimédia, tv, jeux…) → seulement de 15 à 30W
- Mais toujours allumées et grand facteur d'échelle
- Approche : virtualiser les services maison, les relocaliser chez l'opérateur réseau et ne garder à la maison qu'une “simple” prise quasiment passive à très faible consommation (<1 W)
- Challenge : assurer la qualité de service (latence, services..)
- Démontrer les gains en énergie potentiels combinés à la facilité de gestion et partage de services







Conclusions

- Le chemin est encore long… mais la recherche de l'efficacité énergétique change la manière dont nous utilisons et concevons les applications, les services, les protocoles, les équipements….

- More4Same, Same4Less, More4Less

- Sortir les resultats de recherche des labos et les mettre sur les produits du marché

- Sensibiliser, sensibiliser, agir….

- Remettre humain dans la boucle / QoS

EcoInfo

Inria
INVENTEURS DU MONDE NUMÉRIQUE

Laurent.lefevre@inria.fr

Merci à : D. Balouek, F. Cappello, E. Caron, R. Carpa, G. DaCosta, M. Dias de Asuncao, M. Diouri, J.-P. Gelas, O. Gluck, J.C. Mignot, A.-C. Orgerie, G. Tsafack, J.-M. Pierson, L. Pouilloux, F. Rossigneux, P. Stolf, V. Villebonnet

Laurent.lefevre@inria.fr