

Optimisation énergétique des Data-Centers et usages des énergies renouvelables

Marwa Haddad^{ab}, Jean-Marc Nicod^{ab} et Marie-Cécile Péra^{ab}

^aInstitut FEMTO-ST, Université Bourgogne Franche-Comté, CNRS, ENSMM
F-25000 Besançon, FRANCE

^cFCLAB, FR 3538, Université Bourgogne Franche-Comté, CNRS,
rue Thierry Mieg, F-90010 Belfort, FRANCE,

THÉMATIQUES : Énergie (section 63 du CNU) - Informatique (section 27 du CNU)

Résumé

Aujourd'hui, chacun d'entre nous est connecté en permanence à Internet pour communiquer, partager des fichiers, photos, musiques, s'informer, consommer, et ce de manière transparente. Cette facilité apparente cache une réalité, tous les services auxquels nous accédons instantanément tournent en continu sur des ordinateurs regroupés dans des centres de données dédiés, de taille gigantesque, nommées aussi data-centers. La consommation d'énergie occasionnée est source de pollution à travers les émissions de CO₂ dues à la production d'électricité. Ainsi, dans cet article, nous évoquons d'abord les infrastructures actuelles des data-centers et leurs stratégies d'optimisation énergétique, puis les démonstrateurs d'intégration des énergies vertes. Enfin, le projet financé par l'Agence Nationale de la Recherche, DATA-ZERO (www.datazero.org), auquel nous participons, est décrit. Son but est d'étudier comment un data-center peut être alimenté exclusivement par des énergies renouvelables avec optimisation de son dimensionnement, de l'ordonnement des calcul et de l'engagement des sources.

Mots Clés : Optimisation énergétique, Optimisation de ressources, Data-centers, Cloud, Énergie renouvelable, Green Computing, Émissions CO₂.

1 INTRODUCTION

Aujourd'hui, si nous avons la possibilité d'accéder à notre boîte aux lettres électronique depuis notre téléphone, qui n'a pas la capacité à contenir tous nos messages, cela signifie que ces messages sont stockés quelque part sur un disque dur accessible à tout instant. Il est en est de même pour toutes les données que nous consultons sans nous poser de question. En effet, il existe de nombreuses infrastructures matérielles qui reçoivent nos requêtes et nous mettent à disposition les contenus auxquels nous souhaitons accéder. Il en est de même pour nos données. Cela suppose également des infrastructures réseaux complexes pour acheminer l'information et gérer tous les problèmes de contention. Ces dispositifs qui étaient majoritairement déployés au niveau des ressources des entreprises il y a encore quelques années sont maintenant regroupés aujourd'hui dans ce que nous appelons des data-centers.

L'agence de l'environnement et de la maîtrise de l'énergie a déclaré que l'envoi d'un mail avec une pièce jointe revient au coût de fonctionnement d'une ampoule basse consommation de forte puissance pendant 1 heure. Et comme chaque jour, 10 milliards de mails sont envoyés à travers le monde, cela correspond en moyenne à 50 GWh, soit la production électrique de 15 centrales nucléaires fonctionnant pendant une heure [18]. Ainsi, les data-centers consomment énormément d'énergie. Les estimations pour l'année 2010 indiquent qu'ils consomment environ 1,5% de l'électricité totale utilisée dans le monde [11] avec la moitié de la consommation totale attribuée au refroidissement car les ressources des

data-centers produisent de grandes quantités de chaleur.

Pour réduire cette énorme consommation électrique, l'une des initiatives est de mettre au point des mesures efficaces pour réduire à la fois les besoins en matière de calcul et de refroidissement. En ce sens, Kant [10] envisageait une évolution des data-centers physiques privés vers des infrastructures externalisées, virtualisées et géographiquement distribuées fournissant le même niveau de contrôle et d'isolement que les infrastructures existantes. De même, Uddin et al. [20] ont proposé une ligne directrice pour les gestionnaires de data-center afin de concevoir et de mettre en œuvre correctement la virtualisation dans les data-centers. Certaines de ces stratégies visent des mesures telles que la réduction des achats fréquents de matériels, la réduction des coûts d'énergie et de refroidissement, la consolidation de la charge de travail et la réduction des serveurs physiques. En outre, ils ont souligné l'importance de la mise en œuvre de métriques vertes et ont proposé une méthodologie pour choisir des mesures appropriées pour mesurer les performances des data-centers en termes d'efficacité énergétique, d'économies, d'initiatives vertes et d'émissions de CO₂. Les grandes entreprises investissent également dans cette direction. Par exemple, Apple a construit une centrale solaire de 40 MW pour son data-center en Caroline du Nord [1]. McGraw-Hill a récemment complété une matrice solaire de 14 MW pour son data-center [14].

L'article est organisé de la manière suivante. Nous présentons dans un premier temps (paragraphe 2) les différents projets de recherches dont l'objet était l'étude de la problématique énergétique des data-centers. Dans un

deuxième temps (paragraphe 3), nous décrivons les propositions d'amélioration dans ce domaine apportées par ces projets. Le paragraphe 4 introduit le projet ANR DATA-ZERO sur lequel nous travaillons.

2 APERÇU SUR LES DATA-CENTERS

Un data-center se présente comme une structure ou un ensemble de structures abritant des équipements hébergeant les systèmes d'information d'entreprises ou de collectivités, des serveurs, des baies de stockage et des outils de télécommunications. Les requêtes adressées à un data-center sont de différentes natures qui vont de la consultation d'information au déclenchement de traitement plus ou moins important comme les traitements journaliers ou des traitements de calcul type SaaS. Pour cela, ces centres maintiennent un haut niveau de disponibilité et de sécurité.

2.1 DISTRIBUTION DU SYSTÈME ÉLECTRIQUE

Le système de distribution d'énergie du data-center est l'équipement chargé de fournir l'énergie électrique aux charges du système (équipements informatiques et mécaniques). Étant donné que le réseau public peut avoir des micro coupures ou des pannes plus longues qui peuvent causer le dysfonctionnement du data-center, il est essentiel d'assurer une alimentation appropriée. Cet équipement doit ainsi répondre à la qualité d'énergie demandée et à la sécurité d'approvisionnement. La figure 1 montre un schéma de l'infrastructure d'alimentation d'un data-center standard.

En cas de pannes majeures du réseau public, il existe un générateur diesel de secours. L'onduleur, qui peut utiliser différentes technologies de stockage comme des batteries, permet d'attendre le démarrage des générateurs diesel. Enfin, les unités de distribution de l'alimentation (PDU) et les unités d'alimentation (PSU) sont responsables de la distribution et de l'adaptation de la puissance des serveurs.

Le système de distribution d'énergie est différent pour chaque type de data-center en tenant compte des spécificités de l'installation. Cependant, il peut être classé en quatre catégories principales ou « niveaux » (*tier en anglais*). D'une manière générale, le niveau I correspond à la structure de base et comporte des composantes de capacité et des voies de distribution non redondantes. Le niveau II est assez semblable au niveau I, incluant les redondances dans les composantes de capacité, mais pas dans les chemins de distribution de l'alimentation. Les niveaux III et IV présentent des redondances à la fois en termes de composantes de capacité et de voies de distribution. Le niveau IV est la seule infrastructure tolérante à toutes les pannes de sorte que toute défaillance d'un composant n'aura pas d'impact sur le matériel informatique [19].

La figure 2 présente une illustration de la topologie électrique pour le système de distribution d'énergie pour le niveau I.

Dans cette structure, l'alimentation est normalement transmise en courant alternatif triphasé (AC). Cependant, comme les ordinateurs sont alimentés en courant continu (DC), ils sont équipés d'un redresseur électronique

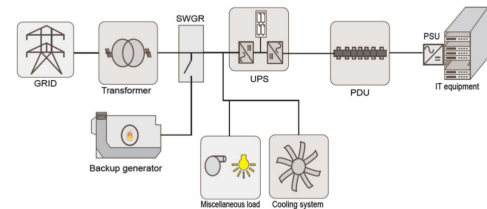


FIGURE 1 – Distribution du système électrique d'un data-center (source : [7])

(convertisseur AC/DC). Il est également possible de noter que seules les charges critiques sont connectées à la sortie de l'onduleur pour éviter tout affaissement transitoire en cas de panne secteur tant que le générateur diesel n'a pas démarré. De plus, cette figure illustre la topologie électrique du système de distribution d'énergie pour le niveau III. Notons que les principales différences entre le niveau I (ou II) et le niveau III est le niveau des redondances dans le système. Dans le niveau III, les composants de capacité considèrent le niveau de redondance numéro 1 pour éviter la déconnexion des charges informatiques en cas de défaillance d'un composant. Ainsi, les redondances sont prises en compte par les onduleurs ainsi que les associations moteur/générateur (M/G). De plus, la connexion au réseau est renforcée en doublant les liens réseaux avec l'extérieur et le nombre de générateurs diesel de secours. En conséquence, les charges informatiques peuvent être alimentées par deux sources d'alimentation indépendantes.

3 OPTIMISATION DES DATA-CENTERS

L'industrie et les chercheurs étudient actuellement de nombreuses méthodes pour réduire l'énergie utilisée dans les data-centers. Ils se sont surtout intéressés aux économies d'énergie à réaliser pour le refroidissement. Patterson [15] a examiné l'empreinte énergétique complète d'un data-center ainsi que l'impact qu'une température ambiante accrue aurait sur chaque composant de l'infrastructure. Cette analyse indique qu'il existe une température optimale pour le fonctionnement de chaque data-center. Ses recherches ont ainsi augmenté le seuil de température admissible.

En outre, les chercheurs Zhang et al. [23] étudient d'autres concepts pour augmenter l'efficacité énergétique dans le système de refroidissement en utilisant un flux d'air variable (*free cooling* en anglais).

D'autres chercheurs étudient de nouvelles mesures liées à l'alimentation électrique. Certains d'entre eux disposent d'une documentation disponible et ont des résultats bien connus, comme dans le cas de Urgaonkar et al. [21] qui ont utilisé l'UPS (uninterruptible power supply) comme élément de stockage électrique combiné à des algorithmes de charge de travail afin de réduire les coûts découlant de l'utilisation des onduleurs ou l'utilisation des sources d'énergies renouvelables (SER) pour minimiser le coût d'exploitation d'un data-center. En outre, les responsables des data-centers devraient utiliser les composants les plus économiques afin de réduire la consommation d'énergie et le coût opérationnel de ses installations.

Le dernier domaine où la société de centre de données

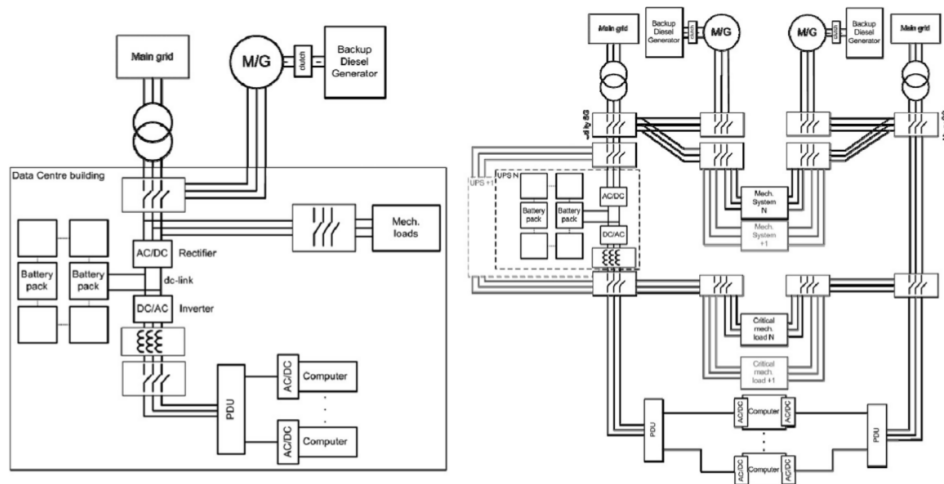


FIGURE 2 – Schéma de la distribution électrique des data-centers : niveau I et niveau III respectivement (source : [7])

fait de gros efforts pour optimiser ou augmenter l'efficacité énergétique est la gestion de la charge de travail de données, où des techniques bien connues telles que la consolidation et l'hétérogénéité permettent de réaliser des économies d'énergie considérables. De nos jours, les efforts visant à améliorer la gestion de la charge de travail de données combinent les techniques commentées avec les caractéristiques de l'infrastructure afin d'obtenir le maximum de bénéfices souhaités. En tant que politique complémentaire, les data-centers qui fonctionnent en partie avec des énergies renouvelables peuvent planifier leur charge de travail (si possible) en fonction de la disponibilité de telles énergies [6, 12].

4 MISE EN ŒUVRE DES RESSOURCES RENOUVELABLES

Récemment, les chercheurs portent plus d'intérêt pour l'utilisation des énergies renouvelables pour fournir une partie de la consommation globale d'énergie des data-centers. Cependant, alors que les énergies renouvelables sont disponibles par intermittence, la demande des data-centers doit être satisfaite, même lorsque l'énergie verte n'est pas disponible.

Un système de transfert de puissance gère en toute sécurité les sources d'énergie en isolant l'électricité de différentes sources et en s'assurant que le data-center reçoive suffisamment de puissance. Stewart et Shen [17] ont décrit un programme de recherche pour la gestion des énergies renouvelables dans les data-centers, concluant que prendre en compte la charge des tâches des data-centers permet de réduire encore leur dépendance à l'égard des énergies non renouvelables. Malkamäki et Ovaska [13] ont étudié l'énergie solaire et le potentiel de refroidissement dans les data-centers européens. Ils ont aussi mis en évidence les relations de base entre l'énergie solaire, la température de l'air et les besoins ultérieurs de refroidissement. Malheureusement, les endroits où le potentiel de production d'énergie solaire est élevé sont moins favorables au refroidissement en raison de sa température ambiante élevée. Arlit et al. [2] ont introduit une méthode pour exploiter un data-center avec de l'énergie renouvelable qui

minimise la dépendance à la puissance du réseau tout en minimisant le coût d'investissement. Ils y parviennent en intégrant la demande des data-centers et la disponibilité des ressources pendant leur exploitation. Ils ont conçu et géré la charge du data-center afin d'utiliser les énergies renouvelables sur site, principalement avec des PV, et ainsi compenser entièrement l'utilisation d'énergie non renouvelable du réseau. Plus précisément, ils ont décrit comment combiner l'utilisation des énergies renouvelables avec la planification dynamique de la charge de travail informatique et les techniques de gestion intégrées pour améliorer l'utilisation globale du data-center, tout en permettant à la demande d'être adaptées en fonction de la disponibilité des ressources.

Pour réduire la consommation et recycler l'énergie thermique autrement gaspillée, de nombreux chercheurs ont étudié l'introduction de technologies différentes pour les data-centers. L'une d'entre elles est la Cogénération de chaleur et d'électricité (CHP) qui pour rendre le data-center plus rentable et énergétiquement efficace. L'utilisation de cette technologie répond à l'ensemble des besoins énergétiques des data-centers, avec des unités d'absorption utilisées pour récupérer la chaleur déchargée par un moteur thermique ou une pile à combustible. En outre, une application plus large de la cogénération réduirait la demande d'électricité des centrales et réduirait la congestion des infrastructures de transport et de distribution d'électricité [4]. Guizzi et al. [8] ont présenté une analyse comparative entre un data-center conventionnel en utilisant une cogénération pour produire de l'électricité couplée avec une machine d'absorption pour fournir du froid. Plus tard, Guizzi et Manno [9] ont discuté de l'évaluation économique et énergétique d'un système de cogénération, pour un data-center adapté à une charge IT de 100 kW, basé sur un reformeur à vapeur à membrane de gaz naturel produisant un flux d'hydrogène pur pour la production d'électricité dans une membrane électrolyte polymère pile à combustible. La chaleur a été récupérée à la fois de l'unité de reformage et de la pile à combustible afin de répondre aux besoins d'un immeuble de bureaux situé à proximité. Les simulations ont démontré que 47% de réduction des

coûts pourraient être atteints lorsque l'énergie thermique provenant du système de cogénération serait récupérée. En outre, le système innovant de gestion de l'énergie peut également offrir des économies substantielles d'un point de vue environnemental. Ebay [22] a récemment commuté un data-center dans l'Utah entièrement alimenté par des piles à combustible de 6 MW alimentées par du gaz naturel reformé. Il est à noter que ce type d'approvisionnement énergétique rend l'infrastructure plus fiable vis à vis des pannes du réseau.

Même si l'utilisation des énergies renouvelables sur place dans les data-centers réels est encore à un stade précoce, certaines entreprises ont mis en œuvre différentes solutions d'énergie verte dans leur portefeuille. La mise en œuvre de l'énergie solaire dans les data-centers n'a pas été largement utilisée, car il faut une très grande surface de panneaux photo-voltaïques pour produire même une fraction de l'énergie requise par ces infrastructures de haute densité énergétique. Néanmoins, il existe déjà des installations réussies. Des panneaux solaires de 100 kW occupant 730 m^2 sont situés sur le toit du data-center sur le campus d'Emerson dans le Missouri. De la même façon et pour tester le potentiel d'utilisation de l'énergie solaire photo-voltaïque pour les data-centers, Intel a installé 10 kW d'électricité dans un data-center au Nouveau-Mexique et Goiri et al. [5] a développé Parasol, un prototype de data-center vert qui comprend un petit conteneur, un ensemble de panneaux solaires, un bloc de batterie électrique et des convertisseurs. De même, des chercheurs de l'Université du Massachusetts ont construit Blink [16], un cluster de cartes mères d'ordinateurs portables alimenté par deux micro-éoliennes et deux panneaux solaires. Un petit data-center en Illinois est devenu le premier data-center à 100% éolien aux États-Unis en passant ses besoins énergétiques quotidiens avec une éolienne de 500 kW [3].

DATAZERO est un projet financé par l'ANR, qui vise à concevoir un centres de calcul en s'appuyant sur les technologies les plus récentes en termes de sources d'énergie renouvelables, d'équipements électriques, de serveurs informatiques et de systèmes de gestion du cloud. Ce projet vise à réduire de façon spectaculaire la consommation d'énergie des futurs data-centers en utilisant les sources renouvelables suivantes : les cellules photo-voltaïques, les éoliennes, les piles à combustibles.

5 LE PROJET ANR DATAZERO

Dans DATAZERO, la question adressée est la façon de gérer l'électricité et les flux de services afin de fournir des services aux clients de manière robuste et efficace au sein de data-centers fonctionnant avec plusieurs sources d'énergie. Pour répondre à cette question, nous avons identifié sept enjeux scientifiques :

1. Faire coïncider les contraintes de demande et d'enveloppe sur les plans électriques et informatiques.
2. Dimensionner correctement l'équipement.
3. Contrôler de façon optimale les convertisseurs électriques.
4. Planifier et gérer la charge informatique.
5. Prendre en compte la gestion thermique.

6. Étudier la complexité du problème d'optimisation
7. Développer un outil de simulation.

Les principaux objectifs de DATAZERO sont des data-centers de taille moyenne (jusqu'à 1000 m^2 et environ 1 MW) où la charge informatique peut être gérée via Virtualisation ou Cloud orchestrator couramment rencontrés dans les entreprises et les institutions publiques.

L'exécution de ces différents travaux se fera en collaboration entre les différents partenaires : les laboratoires LAPLACE (Toulouse), FCLAB (Belfort), IRIT (Toulouse), FEMTO-ST (Besançon) et la société EATON (Grenoble).

6 CONCLUSION

La croissance de la demande des data-centers au cours des dernières années a entraîné une augmentation de leur puissance et donc de leur consommation électrique.

L'impact de ces infrastructures sur le plan énergétique mondiale n'est donc pas sans valeur. Ainsi, de nombreux chercheurs se sont concentrés sur le développement de connaissances, d'outils et de normes systématiques afin de réduire la consommation des data-centers et d'intégrer les énergies renouvelables dans leur portefeuille énergétique. Récemment, plusieurs métriques ont été introduites par les organismes de standardisation et de professionnels et les établissements de recherche pour évaluer l'efficacité énergétique dans les data-centers. Cependant, il est crucial de développer une méthodologie et des normes communes pour un calcul approprié des coefficients d'efficacité énergétique les plus pertinents. Au fur et à mesure que les technologies de l'information continuent d'améliorer leurs vitesses de calcul et leurs capacités année après année, elles consomment davantage d'énergie et la dissipation de chaleur dans les data-centers continue d'augmenter.

Tous ces phénomènes ne sont pas bien caractérisés et donc d'autres efforts doivent être faits, par exemple pour caractériser le potentiel de réutilisation de la chaleur à différentes températures de travail et de refroidissement en fonction de la fiabilité de l'équipement informatique, ou de mieux étudier l'alimentation des data-center par des énergies renouvelables. Toutefois, en raison de la nature complexe et spécifique du système de ces modèles et des difficultés à obtenir des variables d'entrée, les principaux défis sont maintenant de développer des modèles efficaces capables d'estimer dynamiquement la consommation d'énergie et les flux de masse dans les data-centers afin d'améliorer leur caractérisation et leur conception. En outre, l'introduction de stratégies d'efficacité énergétique dans le portefeuille de data-center pourrait éviter les effets négatifs sur le climat.

REMERCIEMENT

Ce travail est soutenu par le projet ANR DATAZERO (contrat n° ANR-15-CE25-0012) et par le Labex ACTION (contrat n° ANR-11-LA BX-01-01).

RÉFÉRENCES

- [1] Apple. Apple facilities, environmental footprint report. *Apple and the Environment*, 2013.

- [2] Martin Arlitt, Cullen Bash, Sergey Blagodurov, Yuan Chen, Tom Christian, Daniel Gmach, Chris Hyser, Niru Kumari, Zhenhua Liu, Manish Marwah, et al. Towards the design and operation of net-zero energy data centers. In *Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 2012 13th IEEE Intersociety Conference on*, pages 552–561. IEEE, 2012.
- [3] Thomas Brunschwiler, Gerhard Ingmar Meijer, Stephan Paredes, Werner Escher, and Bruno Michel. Direct waste heat utilization from liquid-cooled supercomputers. In *2010 14th International Heat Transfer Conference*, pages 429–440. American Society of Mechanical Engineers, 2010.
- [4] Ken Darrow and Bruce A Hedman. *Opportunities for combined heat and power in data centers*. ICF International, 2009.
- [5] Íñigo Goiri, William Katsak, Kien Le, Thu D. Nguyen, and Ricardo Bianchini. Parasol and greenswitch : Managing datacenters powered by renewable energy. *SIGARCH Comput. Archit. News*, 41(1) :51–64, March 2013.
- [6] Íñigo Goiri, Josep Ll Berral, J Oriol Fitó, Ferran Julià, Ramon Nou, Jordi Guitart, Ricard Gavaldà, and Jordi Torres. Energy-efficient and multifaceted resource management for profit-driven virtualized data centers. *Future Generation Computer Systems*, 28(5) :718–731, 2012.
- [7] Íñigo Goiri, Md E Haque, Kien Le, Ryan Beauchea, Thu D Nguyen, Jordi Guitart, Jordi Torres, and Ricardo Bianchini. Matching renewable energy supply and demand in green datacenters. *Ad Hoc Networks*, 25 :520–534, 2015.
- [8] Giuseppe Leo Guizzi and Michele Manno. Fuel cell-based cogeneration system covering data centers’ energy needs. *Energy*, 41(1) :56–64, 2012.
- [9] Giuseppe Leo Guizzi and Michele Manno. Fuel cell-based cogeneration system covering data centers’ energy needs. *Energy*, 41(1) :56–64, 2012.
- [10] Krishna Kant. Data center evolution : A tutorial on state of the art, issues, and challenges. *Computer Networks*, 53(17) :2939–2965, 2009.
- [11] Jonathan Koomey. Growth in data center electricity use 2005 to 2010. *A report by Analytical Press, completed at the request of The New York Times*, 9, 2011.
- [12] Chao Li, Ruijin Zhou, and Tao Li. Enabling distributed generation powered sustainable high-performance data center. In *High Performance Computer Architecture (HPCA2013), 2013 IEEE 19th International Symposium on*, pages 35–46. IEEE, 2013.
- [13] Tuomo Malkamäki and Seppo J Ovaska. Solar energy and free cooling potential in european data centers. *Procedia Computer Science*, 10 :1004–1009, 2012.
- [14] Rich Miller. Huge solar array will support nj data center. *datacenterknowledge*, 2011. <http://www.datacenterknowledge.com/archives/2011/06/14/huge-solar-array-will-support-nj-data-center/>.
- [15] Michael K Patterson. The effect of data center temperature on energy efficiency. In *Thermal and Thermomechanical Phenomena in Electronic Systems, 2008. ITherm 2008. 11th Intersociety Conference on*, pages 1167–1174. IEEE, 2008.
- [16] Navin Sharma, Sean Barker, David Irwin, and Prashant Shenoy. Blink : Managing server clusters on intermittent power. *SIGARCH Comput. Archit. News*, 39(1) :185–198, March 2011.
- [17] Christopher Stewart and Kai Shen. Some joules are more precious than others : Managing renewable energy in the datacenter. In *Proceedings of the workshop on power aware computing and systems*, pages 15–19. IEEE, 2009.
- [18] Coline Tison and Laurent Lichenstein. Internet : la pollution cachée. <https://www.youtube.com/watch?v=75mx9pRjYlg&t=1402s>, oct 2016. Camicas Production avec la participation de France Télévision.
- [19] W Pitt Turner IV, JH PE, PE Seader, and KJ Brill. Tier classifications define site infrastructure performance. *Uptime Institute*, 17, 2006.
- [20] Mueen Uddin, Asadullah Shah, and Raed Alsaqour. Implementation of virtualization in data centers to increase proficiency and performance. *Journal of Theoretical & Applied Information Technology*, 53(2), 2013.
- [21] Rahul Uргаonkar, Bhuvan Uргаonkar, Michael J Neely, and Anand Sivasubramaniam. Optimal power cost management using stored energy in data centers. In *Proceedings of the ACM SIGMETRICS joint international conference on Measurement and modeling of computer systems*, pages 221–232. ACM, 2011.
- [22] Connie Yang. Ebay tech blog. <http://tech.ebay.com/>.
- [23] Hainan Zhang, Shuangquan Shao, Hongbo Xu, Hui-ming Zou, and Changqing Tian. Free cooling of data centers : A review. *Renewable and Sustainable Energy Reviews*, 35 :171–182, 2014.