UNIVERSITÉ DU QUÉBEC À MONTRÉAL

RÉDUCTION DE LA CONSOMMATION D'ÉNERGIE DANS LES CENTRES DE DONNÉES EN UTILISANT LE FORMALISME DES RÉSEAUX DE FILES D'ATTENTE

MÉMOIRE PRÉSENTÉ COMME EXIGENCE PARTIELLE DE LA MAITRISE EN INFORMATIQUE

PAR FATOUMATA BALDE

OCTOBRE 2018

UNIVERSITÉ DU QUÉBEC À MONTRÉAL Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.07-2011). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

Je tiens à remercier tout d'abord ma directrice de recherches, Professeur Halima Elbiaze, pour sa patience, et surtout pour sa confiance, ses remarques et ses conseils, sa disponibilité et sa bienveillance. Qu'elle trouve ici le témoignage de ma profonde gratitude. Je voudrais également remercier le professeur Bamba Gueye pour sa disponibilité et son efficacité, il n'a ménagé aucun effort pour la réussite de ce mémoire mais aussi les membres du jury pour avoir accepté d'évaluer ce travail et pour toutes leurs remarques et critiques. Je tiens aussi à remercier mon cher mari qui a été très compréhensif. Je te remercie chéri pour l'amour que tu m'as apporté tout au long de ce travail. Je remercie également Marie-Claude Cote pour sa disponibilité et sa réactivité ainsi que Mouhamed Dieye que je tiens sincèrement à remercier pour son aide précieuse:

Merci à vous tous !!!.

DÉDICACES

A la mémoire de mon défunt père.

À la meilleure et merveilleuse maman sur terre,

À l'homme le plus généreux, le plus patient et le plus doux que je n'ai jamais

rencontré, je veux parler de mon cher et tendre époux.

À mes chers enfants : Awa, Sergine, Diadia.

À mes frères et ma sœur. À ma tante et homonyme.

À Bamba Gueye qui n'arrête jamais de me motiver.

À mes beaux parents et à toute ma famille.

À tous mes amis et collègues.

À tous ceux qui, par un mot, m'ont donné la force de continuer

TABLE DES MATIÈRES

LISTE DES FIGURES
LISTE DES TABLEAUXi
LISTE DES ABREVIATIONS
RÉSUMÉ
INTRODUCTION GENERALE
Contexte
Objectif
Organisation
1. CHAPITRE 1: ETAT DE L'ART SUR L'ECONOMIE D'ENERGIE DANS LE CLOUD COMPUTING
1.1. Introduction
1.2. Topologie des centres de données
1.2.1 Topologie tree-based
1.2.2 Topologies récursives1
1.2.3 Comparaison des performances de quelques topologies
1.3. Consommation d'énergie dans les centres de données
1.3.1. Les principales causes des pertes d'énergie2
1.4. Techniques de réduction de la consommation d'énergie dans les centre de données
1.4.1. Approches basées sur les méthodes de virtualisation et de migration de machines virtuelles
1.4.2. Approches existantes pour réduire les besoins en énergie du réseau2
1.4.3. Gestion et analyse du trafic dans le Cloud privé et public2
1.4.4. Économies d'énergie dans les réseaux d'accès2
1.4.5. Economie d'énergie basée sur l'organisation du centre de données 3
1.5. Conclusion
2. Chapitre 2: GESTION DE L'ENERGIE DANS LES CENTRES DE DONNEES (GREENPOD)
2.1. Introduction 3

2.2.	Co	ontexte de la proposition	35
2.3.	Pr	ésentation du modèle	37
2.3	3.1.	Modèle analytique	43
2.3	3.2.	Calcul des probabilités	45
2.3	3.3.	Calcul des métriques de performances	47
2.4.	Ré	ésultats numérique	48
2.5.	Co	onclusion	59
CONCI	LUSI	ON GENERALE	61
Bilan	1:		61
Pers	pectiv	ves:	61
BIBLIC	GR A	APHIE	i

LISTE DES FIGURES

Figure 1-1: Taxonomie des topologies de centre de données [Yang et al. 2013]	6
Figure 1-2: Topologie arborescente à trois niveaux [Kliazovich et al. 2010]	7
Figure 1-3: Topologie Fat Tree [Yang et al. 2013]	
Figure 1-4: La topologie Fat-tree de Facebook [Andreyev 2014]	
Figure 1-5: La topologie Clos Network [Yang et al. 2013]	
Figure 1-6: Construire un DCell1 à partir de DCell0s avec n = 4 [Chirgwin 2015]	
1-7: Construction de Bcube1 à partir de Bcube0 pour n=4 [Chirgwin 2015]	
1-8:Répartition de la consommation d'énergie dans un centre de données [Jing et al.	
2017]	
Figure 2-1: Topologies Fat-tree	38
Figure 2-2: Modèle du système	
Figure 2-3: Comportement d'un serveur du premier groupe	40
Figure 2-4: Comportement d'un serveur des groupes restants	
Figure 2-5: Notre modèle du système M/M/(k^3)/4	
Figure 2-6: La puissance consommée en fonction du nombre de clients dans le	
système	48
Figure 2-7: Consommation d'énergie en fonction du temps de simulation pour un	
système peu sollicité	50
Figure 2-8: Consommation d'énergie en fonction du temps de simulation pour un	
système moyennement sollicité	51
Figure 2-9: Consommation d'énergie et temps d'attente de notre proposition en	
fonction du temps de simulation pour un système moyennement sollicité	52
Figure 2-10: Consommation d'énergie en fonction du temps de simulation pour un	
système très sollicité	53
Figure 2-11: Consommation d'énergie et temps d'attente de notre proposition en	
fonction du temps de simulation pour un système très sollicité	53
Figure 2-12: Temps d'attente des deux propositions en fonction du temps de	
simulation	54
Figure 2-13: Consommation d'énergie des deux propositions en fonction du nombre	,
de serveurs	
Figure 2-14: Consommation d'énergie de notre proposition par rapport à un centre d	le
données par classique	55
Figure 2-15: Consommation d'énergie des Commutateurs par rapport au deux	
propositions	56
Figure 2-16: Consommation totale d'énergie (Commutateurs+serveurs) par rapport	
deux propositions	
Figure 2-17: Consommation totale d'énergie (Commutateurs+serveurs) de notre	
	57

Figure 2-18: Consommation d'énergie des serveurs en fonction de la charge du	
système	. 58
Figure 2-19: Energie consommée (Serveurs + Commutateurs) en fonction de la	
charge	.59

LISTE DES TABLEAUX

Tableau 1-1: Résumé de la performance des topologies [Yang et al. 2013]	16
Tableau 1-2: Résumé de la redondance matérielle [Yang et al. 2013]	17
Tableau 2-1: Nombre de serveurs actifs et les seuils d'activation et de désactivation	à
chaque niveau	43

RÉSUMÉ

Réduire l'énergie consommée dans le Cloud Computing devient une direction de recherche de plus en plus difficile en raison de la croissance exponentielle des services qui y sont hébergés et livrés. De nos jours, la réduction de la consommation d'énergie dans les centres de données a reçu beaucoup d'attentions des chercheurs, l'objectif étant d'augmenter l'efficacité énergétique dans les Cloud. Malgré ces efforts, un travail récent [Baliga et al. 2011] a montré que le Cloud Computing n'est pas toujours une technologie verte.

Les coûts énergétiques élevés liés à l'exploitation d'un centre de données ont conduit à une réflexion sur le fonctionnement efficace d'un centre de données. Les centres de données sont conçus pour supporter la charge de trafic maximale attendue. Il serait bien pour un fournisseur d'adapter de manière dynamique le nombre de ressources offertes en fonction de la charge du trafic actuel. Dans nos propositions, nous présentons un modèle basé sur la théorie des files d'attente permettant de minimiser la consommation d'énergie, mais aussi de garantir des performances en termes de temps d'attente. Nous utilisons des politiques de gestion dynamique qui activent/désactivent un certain nombre de POD dans une topologie Fat-tree k_aray en fonction du trafic. Nous développons un modèle de file d'attente avec des seuils d'activation qui sont proportionnels à la valeur de k pour activer les serveurs et les Commutateurs des autres POD.

Mots clés: Consommation d'énergie; Cloud Computing; Centre de données ; Files d'attente.

LISTE DES ABREVIATIONS

DCN	Data center Network	Réseau de centre de données			
DVFS	Dynamic Voltage	Mise à l'échelle dynamique de la			
	FrequencyScaling	fréquence de tension			
EMS	Energy Management System	Système de gestion d'énergie			
EoR	End of Row	End of Row			
EPA	Environmental Protection Agency	Agence de protection environnementale			
ISP	Infrasutructure service Provider Fournisseur de service				
MBFD	Modified Best Fit Decreasing	Algorithme (MBFD)			
NPA	No Power Aware				
Pod	Performance Optimized Datacenter	Optimisé les performances des DC			
Qos	Quality of service	Qualité de service			
SLA	Service Level Agreement	Accord de niveau de service			
ST	Single Threshold	Seuil unique			
ToR	Top of Rack	Top of Rack			
UADB	Universite Alioune Diop de	Universite Alioune Diop de Bambey			
	Bambey				
VM	Virtual Machine	Machine virtuelle			

INTRODUCTION GENERALE

Contexte

La réduction de la consommation d'énergie dans le « cloud computing » (Informatique nuagique) est actuellement un gros challenge du fait de la croissance exponentielle des services qui y sont hébergés et offerts. Malgré les efforts consentis, de récentes études [Baliga et al. 2011] ont montré que le cloud est énergivore. Dans un article paru dans le journal RTL, les auteurs affirment que la pollution numérique pourrait rapidement dépasser celle de l'aéronautique [Sulmont et al 2016]. Cette pollution représentait selon les auteurs 2% des émissions en dioxyde de carbone et devrait doubler en 2019. En 2014, la consommation d'énergie des centres de données américains est estimée à 70 milliards de kWh, soit environ 1,8% de la consommation totale des États-Unis [Brackbill 2016]. Les coûts énergétiques élevés liés à l'exploitation d'un centre de données ont conduit à une réflexion sur leur fonctionnement efficace. Etant conçus pour supporter un maximum de charge de travail, une bonne topologie et une bonne organisation des centres de données permettrait d'économiser de l'énergie.

Objectif

Dans le cadre de ce mémoire, nous proposons une solution de réduction de la consommation d'énergie dans les centres de données. Notre objectif est dans un premier temps, d'évaluer la quantité d'énergie consommée dans le centre de données depuis les routeurs coeur jusqu'aux serveurs, mais aussi la quantité d'énergie économisée quand les serveurs et Commutateurs non utilisés sont mis hors tension. Nous proposons par la suite un dispositif permettant de faire des économies d'énergie par rapport aux solutions existantes. Nous avons pour cela utilisé le formalisme des réseaux de file d'attente pour modéliser le problème.

Organisation

Nous allons dans un premier temps présenter un état de l'art sur les techniques de réduction de la consommation d'énergie dans les centres de données. Pour y arriver, nous présentons les différentes topologies des centres de données, par la suite, nous montrons les différentes causes des pertes d'énergie dans les centres de données. Un ensemble d'approches permettant de réduire la consommation d'énergie sera par la suite présenté. On va citer des approches basées sur les méthodes de virtualisation et de migration de machines virtuelles, des approches permettant de réduire les besoins en énergie du réseau, des mécanismes permettant d'économiser de l'énergie dans les réseaux d'accès et enfin des mécanismes basée sur l'organisation du centre de données:

La dernière partie présentera le modèle que nous avons proposé pour réduire la consommation d'énergie dans les centres de données en utilisant le formalisme de file d'attente. Dans ce modèle nous utilisons une topologie Fat-tree avec activation et désactivation des groupes de serveurs dépendant de la valeur de K. Cette gestion dynamique des serveurs dépendra de la taille de la file d'attente.

1. CHAPITRE 1: ETAT DE L'ART SUR L'ECONOMIE D'ENERGIE DANS LE CLOUD COMPUTING

1.1. Introduction

Les centres de données à grande échelle constituent le support d'infrastructure de base pour les services du Cloud en constante expansion. Ainsi, les caractéristiques de performance et de fiabilité des centres de données auront un impact significatif sur l'évolutivité de ces services. En particulier, le réseau du centre de données doit être agile et reconfigurable afin de répondre rapidement aux demandes d'applications et aux besoins de service en constante évolution. Des travaux de recherche importants ont été réalisés sur la conception des topologies de réseau de centres de données afin d'améliorer les performances des centres de données. Dans cette partie, nous présentons quelques travaux sur les conceptions et les topologies de réseaux de centres de données publiées récemment. Nous commençons par une discussion sur diverses topologies de réseau de centres de données représentatives, et les comparons à plusieurs propriétés afin de mettre en évidence leurs avantages et leurs inconvénients.

1.2. Topologie des centres de données

La conception de l'infrastructure des centres de données a récemment suscité un vif intérêt dans les milieus universitaire et industriel. Les centres de données doivent s'adapter à la croissance rapide des applications sur Internet comme l'hébergement et la distribution de contenu vidéo (par exemple, YouTube, NetFlix), les réseaux sociaux (p. ex. Facebook, Twitter) et les calculs à grande échelle (p. ex., exploration de données, bioinformatique, indexation). Par exemple, les services en ligne Microsoft Live sont pris en charge par un centre de données basé à Chicago, qui est l'un des plus grands centres de données jamais créés, couvrant plus de 700 000 pieds

carrés. Le Cloud Computing, en particulier est caractérisé comme l'aboutissement de l'intégration des infrastructures informatiques et de données pour fournir une approche évolutive, agile et rentable pour répondre aux besoins informatiques critiques en constante augmentation (en termes de calcul, de stockage et d'applications) dans les entreprises [2] [Greenberg et al. 2009].

Les centres de données massifs fournissant le stockage, forment le noyau de l'infrastructure pour le Cloud [Greenberg et al. 2009]. Il est donc impératif que l'infrastructure du centre de données, y compris la mise en réseau des centres de données, soit bien conçue pour que le déploiement et la maintenance de l'infrastructure soient rentables. Avec la disponibilité des données et la sécurité en jeu, le rôle du centre de données est plus critique que jamais.

La topologie du réseau interconnectant les serveurs a un impact significatif sur l'agilité et la reconfigurabilité de l'infrastructure du centre de données afin de répondre aux demandes de changement d'applications et aux exigences de service. Aujourd'hui, les réseaux des centres de données utilisent principalement des commutateurs Top of Rack (ToR) interconnectés par des commutateurs de End of Row (EoR), eux-mêmes connectés via des commutateurs centraux. Cette approche conduit à une surabondance de bande passante significative sur les liens dans le noyau du réseau [Al-Fares et al. 2008], et a incité plusieurs chercheurs à suggérer d'autres approches pour des architectures de réseau économiques et évolutives. Selon la reconfigurabilité de la topologie après le déploiement du DCN (Data Center Network), il existe des architectures fixes et des architectures flexibles. Les architectures fixes peuvent être classées en deux catégories: les architectures arborescentes telles que Fat tree [Greenberg et al. 2009] et Clos Network [Hamilton et al. 2009], et les topologies récursives comme DCell [Guo et al. 2008] et BCube [Lu et al. 2009]. Les architectures flexibles telles que cThrough [Wang et al. 2010],

Helios [Farrington et al. 2010] et OSA [Chen et al. 2012] permettent la reconfigurabilité de la topologie du réseau.

Dans cette section, nous présentons les topologies des centres de données arborescente. Compte tenu du coût et de la disponibilité, Ethernet haut débit (1 Gbps, 10 Gbps, 100 Gbps) est le plus souvent utilisé dans la construction de réseau interne de centres de données [Kant 2009]. Cependant, bien que la plupart des centres de données utilisent des commutateurs Ethernet pour interconnecter les serveurs, il existe encore de nombreuses façons de mettre en œuvre les interconnexions, ce qui conduit à différentes topologies de réseau de centre de données. Chacune de ces différentes topologies est caractérisée par différentes exigences de ressources, visant à apporter des améliorations à la performance des centres de données. Si les serveurs et les commutateurs sont considérés comme des sommets, les fils comme des arêtes, la topologie de chaque réseau de centre de données peut être représentée sous forme de graphe.

La figure 1-1 représente la taxonomie des différents réseaux de centre de données.

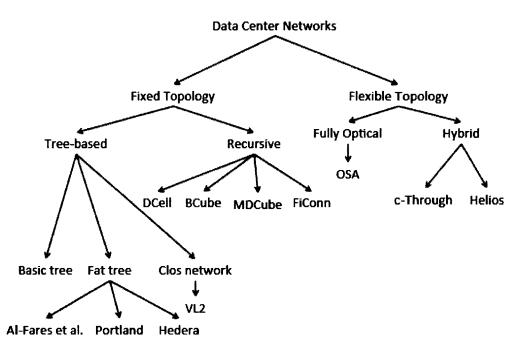


Figure 1-1: Taxonomie des topologies de centre de données [Yang et al. 2013]

Ci-dessous les notations utilisées dans la suite.

n : nombre de ports d'un commutateur dans une architecture.

k : nombre de ports d'un serveur dans une architecture.

N: nombre total de serveurs dans un centre de données.

Il est à noter que n et k peuvent varier en fonction de la position du nœud.

1.2.1 Topologie tree-based

L'architecture arborescente standard et leurs variantes sont largement utilisées dans la conception de réseaux de centres de données [Al-Fares et al. 2008], [Niranjan et al. 2009], [Al-Fareset al. 2010].

❖ Topologie Basic-tree

Les centres de données reposent généralement sur l'architecture hiérarchique à trois niveaux basée sur une structure arborescente, illustrée à la figure 1-2

Dans une topologie à trois niveaux, il existe un niveau cœur qui est la racine de l'arbre, un niveau d'agrégation qui est au milieu. Le troisième niveau est constitué d'un groupe d'hôtes (principalement 16/32) qui est connecté à un commutateur, appelé commutateur ToR (Top of Rack) créant ainsi la couche d'accès à la topologie.

Dans une topologie basic-Tree, les commutateurs de niveau supérieur doivent prendre en charge la communication de données entre un grand nombre de serveurs. Des commutateurs offrant des performances et une fiabilité supérieure sont donc nécessaires dans ces niveaux. Le nombre de serveurs dans une architecture arborescente est limité par le nombre de ports sur les commutateurs.

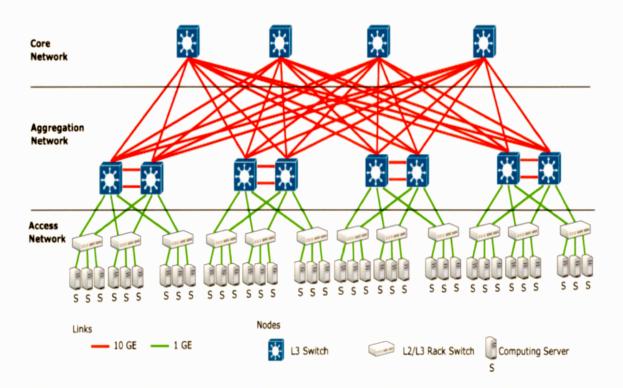


Figure 1-2: Topologie arborescente à trois niveaux [Kliazovich et al. 2010]

❖ Topologie Fat-tree

Fat-tree est une version étendue d'une topologie arborescente [Leiserson 1985]. Fat-tree a été utilisé comme topologie pour les centres de données par plusieurs chercheurs [Al-Fares 2008], [Niranjan et al 2009]. C'est un réseau basé sur un arbre binaire complet.

Parmi les n-ports, le commutateur au niveau edge est connecté à $\frac{n}{2}$ serveurs et à $\frac{n}{2}$ commutateurs au niveau agrégation.

Les $\frac{n}{2}$ commutateurs niveau agrégation, les $\frac{n}{2}$ commutateurs niveau edge et les servers connectés aux commutateurs niveau edge forment ce qu'on appelle Pod. Au niveau cœur, il y a $(\frac{n}{2})^2$ commutateurs, chacun est connecté à chaque Pod parmi les "n"Pod.

La figure 1-3 présente une topologie Fat-Tree avec n=4. Contrairement à la topologie basic Tree, tous les trois niveaux utilisent le même type de commutateur. Les commutateurs de haute performance ne sont pas nécessaires dans les niveaux cœur et agrégation. Le nombre maximum de serveurs dans une topologie Fat-Tree à n-ports est $\frac{n^3}{4}$ [Al-Fares 2008].

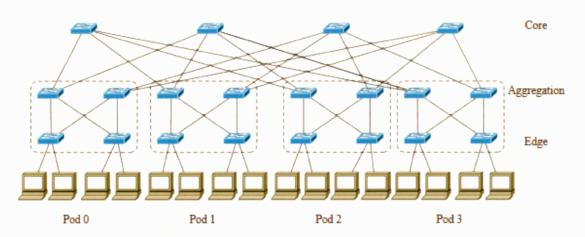


Figure 1-3: Topologie Fat Tree [Yang et al. 2013]

Facebook [Andreyev 2014]a aussi déployé une version de la topologie Fat-Tree afin d'obtenir une bande passante de bissection élevée, un déploiement réseau

rapide et une évolutivité des performances pour rester en phase avec la nature agile des applications exécutées dans les centres de données.

Il se compose de Pod, une unité de réseau standard comme le montre la figure 1-3. La bande passante de liaison ascendante de chaque TOR est 4 fois (4 * 40G = 16 * 10G) la bande passante descendante pour chaque serveur connecté.

Pour mettre en œuvre la connectivité à l'échelle du bâtiment, il a créé quatre «plans» indépendants de commutateurs de colonne vertébrale (commutateur de niveau 3), chacun pouvant être mis à l'échelle jusqu'à 48 périphériques indépendants dans un plan. Border Gateway Protocol (BGP4) est utilisé comme un protocole de contrôle pour le routage tandis qu'un contrôleur centralisé est déployé pour pouvoir remplacer les chemins de routage chaque fois que nécessaire, en adoptant une approche de « contrôle distribué, priorisation centralisée ». Afin d'utiliser tous les chemins disponibles entre deux hôtes, le routage ECMP (EqualCost Multiple Path) avec le hachage basé sur le flux est implémenté.

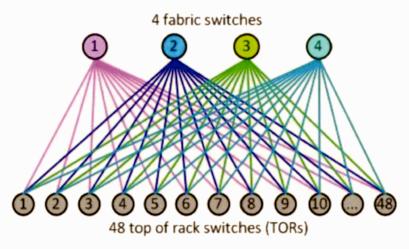


Figure 1-4: La topologie Fat-tree de Facebook [Andreyev 2014]

* Topologie Clos Network:

Greenberg et al. ont proposé une architecture de réseau de centre de données utilisant une topologie Clos Network dans [Hamilton et al. 2009]. Un Clos

Network est également un arbre à plusieurs racines. Un Clos Network se compose généralement de trois niveaux de commutateurs : les commutateurs ToR directement connectés aux serveurs, les commutateurs d'agrégation connectés aux commutateurs ToR et les commutateurs intermédiaires connectés aux commutateurs d'agrégation. Ces trois niveaux de commutateurs sont appelés commutateur "Input", "middle", et "output" dans le Clos original [Dally et Towles 2004]. Le nombre de commutateurs est déterminé par le nombre de ports sur les commutateurs intermédiaires et commutateurs d'agrégation. Si chacun de ces commutateurs a "n" ports, il y aura "n" commutateurs d'agrégation et $\frac{n}{2}$ commutateurs intermédiaire. Il y a exactement un lien entre chaque commutateur intermédiaire et chaque commutateur d'agrégation. Les $\frac{n}{2}$ ports restant sur chaque commutateur d'agrégation est connecté à $\frac{n}{2}$ commutateurs ToR. Chacun des commutateurs ToR est connecté à $\frac{n}{2}$ différents commutateurs d'agrégation et les ports restants sur les commutateurs ToR sont connectés aux serveurs.

Il y a $\frac{n^2}{4}$ commutateurs ToR. Pendant que les commutateurs intermédiaires et les commutateurs d'agrégation doivent avoir le même nombre de ports, le nombre de ports d'un commutateur ToR n'est pas limité. Si " n_{ToR} " ports de chaque commutateur ToR sont connectés aux serveurs, il y aura $\frac{n^2}{4} * n_{ToR}$ serveurs dans le réseau. La figure 1-5 montre une Clos Network avec n=4 et " $n_{ToR}=2$.

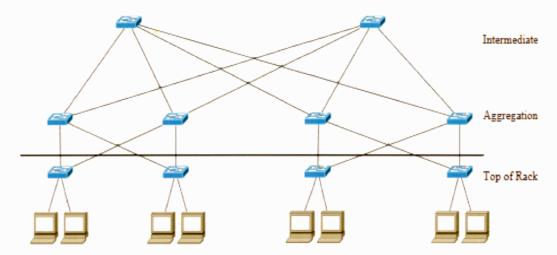


Figure 1-5: La topologie Clos Network [Yang et al. 2013]

1.2.2 Topologies récursives

Les topologies arborescentes peuvent être mises à l'échelle en insérant plusieurs niveaux de commutateurs, tandis que chaque serveur est connecté à un seul des commutateurs de niveau inférieur. Les topologies récursives ont toujours des niveaux comme dans les topologies arborescentes. Cependant, les topologies récursives utilisent des structures de niveau inférieur comme des cellules pour construire des structures de niveau supérieur, et les serveurs dans des topologies récursives peuvent être connectés à des commutateurs de différents niveaux ou même d'autres serveurs. Il existe plusieurs ports réseau sur les serveurs de topologies récursives, ce qui les rend significativement différents des topologies arborescentes. Les graphes, plutôt que les arbres multi-racines, sont adaptés pour représenter des architectures récursives. Certaines topologies récursives représentatives seront abordées dans les sections suivantes.

Topologie DCell:

L'élément le plus fondamental d'un DCell, appelé DCell₀, consiste à avoir n serveurs et un commutateur à n-port. Chaque serveur dans un DCell₀ est connecté au commutateur dans le même DCell₀. Soit DCell_k un DCell de niveau-k. La première étape consiste à construire un DCell¹ à partir de plusieurs DCell₀.

Chaque DCell₁ a n + 1 DCell₀, et chaque serveur de chaque DCell₀ dans un DCell₁ est connecté à un serveur dans un autre DCell₀, respectivement. En conséquence, les DCell₀ sont connectés les uns aux autres, avec exactement un lien entre chaque paire de DCell₀. Une procédure similaire est utilisée pour construire un DCell_k à partir de plusieurs DCell_{k-1}. Dans un DCell_k, chaque serveur aura finalement*k* + 1liens : le premier lien ou le lien de niveau 0 connecté à un commutateur lors de la formation d'un DCell₀, et un lien de niveau i connecté à un serveur dans le même DCell₁ mais à un DCell₁₋₁ différent.

Supposons que chaque DCell_{k-1}ait des serveurs t_{k-1} , alors un DCell_k sera constitué de t_k DCell_{k-1}, et par conséquent de $t_{k-1} \times t_k$.

Évidemment, nous avons $t_{k=1} \times (t_{k-1} + 1)$. La figure 1-6 montre un DCell₁ avec n = 4. On peut voir que le nombre de serveurs dans un DCell croît de façon exponentielle, et le nombre total de niveaux dans un DCell est limité par le nombre de carte d'interface réseau sur les serveurs qu'il contient. DCell peut évoluer vers un très grand nombre de serveurs en utilisant des commutateurs et des serveurs avec très peu de ports. Par exemple, lorsque n = 6, k = 3, un DCell entièrement construit peut comprendre plus de 3 millions de serveurs [Guo et al. 2008].

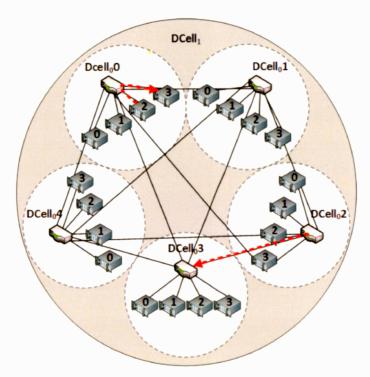


Figure 1-6: Construire un DCell1 à partir de DCell0s avec n = 4 [Chirgwin 2015]

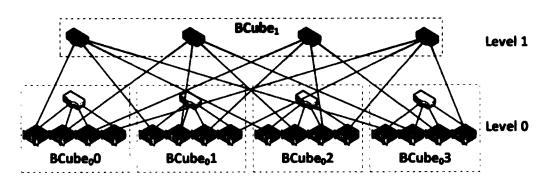
* Topologie Bcube

BCube [Guo et al. 2009] est une topologie récursive spécialement conçue pour les centres de données modulaires embarqués. La construction d'un cluster de centre de données dans un conteneur d'expédition de 20 ou 40 pieds le rend très portable. Comme les demandes changent dans différents centres de données, le cluster entier peut être facilement déplacé. Bien que le temps de déploiement soit considérablement plus court, l'inconvénient de cet environnement est qu'en raison des contraintes opérationnelles et d'espace, une fois déployé sur le terrain, il est difficile de desservir le cluster. L'élément le plus fondamental d'un BCube, nommé BCube₀, est également le même que celui d'un DCell₀: n serveurs connectés à un commutateur n-port.

La principale différence entre BCube et DCell réside dans la façon dont ils évoluent. BCube utilise plus de commutateurs lors de la construction d'architectures de plus haut niveau. Lors de la construction d'un BCube₁, "n" commutateurs supplémentaires

sont utilisés, se connectant à exactement un serveur dans chaque BCube₀. Par conséquent, un BCube₁ contient "n" BCube₀s et "n" commutateurs supplémentaires (Si les commutateurs des BCube₀ sont pris en compte, il y a "2n" commutateurs dans un BCube₁). Plus généralement, un BCube_k est construit à partir de "n" commutateurs BCube_{k-1}s et "nk" extra n-port. Ces commutateurs supplémentaires sont connectés à exactement un serveur dans chaque BCube_{k-1}. Dans un BCube de niveau-k, chaque niveau requiert "nk" commutateurs (dont chacun est un commutateur n-port). BCube utilise plus de commutateurs lors de la construction de structures de niveau supérieur, alors que DCell n'utilise que des commutateurs n-port de niveau 0.

Les serveurs doivent cependant avoir "k+1" d'interface réseau. La figure 1-7 montre un BCube₁ avec n=4. Tout comme le DCell, le nombre de niveaux dans BCube dépend du nombre de ports sur les serveurs. Le nombre de serveurs dans BCube croît exponentiellement avec les niveaux, beaucoup plus lentement que DCell. Par exemple, lorsque n=6, k=3, un BCube entièrement construit peut contenir 1296 serveurs. Considérant que BCube est conçu pour des centres de données basés sur des conteneurs, une telle évolutivité est suffisante [Guo et al. 2009].



1-7: Construction de Bcube1 à partir de Bcube0 pour n=4 [Chirgwin 2015]

1.2.3 Comparaison des performances de quelques topologies

Le tableau 1-1 présente les comparaisons de certaines métriques de performance de différentes topologies.

***** Bande passante

Les quatre premières lignes du tableau 1-1 montrent la bande passante que les topologies peuvent offrir sous différents modèles de trafic. "One-to-one" signifie la bande passante maximale que la topologie peut offrir lorsqu'un serveur arbitraire envoie des données à un autre serveur arbitraire, et ainsi de suite.

La bande passante "All to All" signifie que chaque serveur établit un flux vers tous les autres serveurs. Chaque type de modèle de trafic est significatif pour différentes applications. Par exemple, un trafic one-to-severals survient lorsque le système de fichiers fait des répliques, le trafic one-to-all se produit lors de la mise à jour de certains logiciels sur tous les serveurs et le trafic all-to-all est très courant dans MapReduce[Yang et al. 2013]. Dans le tableau 1-1, les largeurs de bande sont exprimées en nombre de liens, ce qui implique que chaque lien dans un centre de données a la même bande passante. Il est montré dans le tableau 1-1 que les bandes passantes one-to-one, one-to-severals et one-to-all sont en fait limitées par le nombre de ports sur les serveurs, ou le degré de serveurs. La topologie Basic tree offre la plus petite bande passante all-to-all, limitée par le nombre de ports du commutateur à la racine. Comme fat-tree introduit des commutateurs redondants, il peut offrir autant de chemins disjoints que le nombre de serveurs.

La largeur de bissection est le nombre minimal de connexions qu'il faut retirer pour avoir deux parties égales non connectées (mesure la capacité du réseau à transmettre des messages simultanément). Plus elle est grande, meilleure est la capacité de tolérance aux pannes qu'une topologie aura. Basic tree a la plus petite largeur de bissection, ce qui signifie qu'il est le plus vulnérable de tous. Une dégradation progressive des performances implique que lorsque de plus en plus de composants tombent en panne dans le centre de données, les performances diminuent lentement sans chutes dramatiques. Lors de la construction d'un nouveau centre de données, il est habituel qu'une topologie partielle soit construite en premier et que d'autres

composants soient ajoutés en fonction des besoins futurs. Ce réseau partiel peut être vu comme un réseau avec beaucoup d'échecs. Construire un réseau partiel avec une topologie de dégradation gracieuse signifie que les gens peuvent obtenir les performances attendues avec moins de composants. [Guo et al. 2009] donne une comparaison entre fat-tree, DCell et BCube sur cette métrique avec des résultats de simulation.

Tableau 1-1: Résumé de la performance des topologies [Yang et al. 2013]

	Tree-Based A	Architecture	Recursive Architecture			
	Basic Tree	Fat Tree	Clos Network	Dcell	Bcube	
One-to-one	1	1	1	k + 1	k + 1	
One-to-several	1	1	1	k + 1	k + 1	
One-to-all	1	1	1	k + 1	k + 1	
All-to-one	n	N	$\frac{2N}{n_{ToR}}$	$>\frac{N}{2^k}$ $\frac{n}{n-1}(N-1)$		
BisectionWidth	$\frac{n}{2}$	$\frac{N}{2}$	$\frac{N}{n_{ToR}}$	$> \frac{N}{4\log_n N}$	$\frac{N}{2}$	

* Redondance matérielle

Le tableau 1-2 montre la redondance matérielle offerte par différentes topologies. Avant de discuter du tableau 1-2, certaines définitions doivent être clarifiées :

Définition 1 (Node-Disjoint Paths)[Yang et al. 2013] : Le minimum du nombre total de chemins qui ne partagent aucun nœud intermédiaire commun entre des serveurs arbitraires.

Définition 2 (chemins disjoints)[Yang et al. 2013] : nombre minimal de chemins qui ne partagent aucun nœud edge commun entre des serveurs arbitraires.

Définition 3 (tolérance f-fault) [Duato et al. 2003]: Un réseau est f-fault tolérant si pour tout composant défectueux du réseau, le réseau est toujours connecté.

Définition 4 (Niveau de redondance)[Duato et al. 2003] : Un réseau a un niveau de redondance égal à r si et seulement si, après avoir supprimé tout ensemble de liens r,

il reste connecté, et qu'il existe un ensemble de liens r + 1 tels qu'après les avoir supprimés, le réseau n'est plus connecté.

Tableau 1-2: Résumé de la redondance matérielle [Yang et al. 2013]

		Tree-Based Architecture		Recursive Architecture			
			Basic Tree	Fat Tree	Clos	DCell	BCube
Node-disjoint Paths			1	1	1	k+1	k+1
Edge-disjoint Paths			1	1	1	k+1	k+1
	Comm utateur s	Edge/ToR	0	0	0		
Redondancy		Aggregation	0	$\frac{n}{2}-1$	1	. k	k
Level		Core/Intermediate	0	$\frac{n}{4}-1$	$\frac{n}{2}-1$		
	Server		_	-	-	k	k

Dans le tableau 1-2, les niveaux de redondance des différents composants sont indiqués respectivement, ce qui reflète l'influence des différentes défaillances. Les nombres de chemins de nœuds disjoints et de chemins de bord-disjoints sont les mêmes dans chacune des topologies. En effet, il existe exactement un lien entre chaque paire de nœuds directement connectés (serveur à serveur, serveur à commutateur, commutateur à commutateur) dans ces topologies. Si des liens redondants sont ajoutés rendant la topologie multi-graphes, il y aura plus de chemins edge-disjoints que de chemins node-disjoints. Le nombre de chemins node-disjoints montre la capacité d'un système à gérer les défaillances de nœuds, tandis que le nombre de chemins edge-disjoints montre la capacité à gérer les échecs de liaison. Une défaillance de nœud peut avoir un impact plus grave qu'un échec de liaison, car elle entraîne généralement plusieurs échecs de liaison.

Les topologies tree-based ont moins de chemins node-disjoint et de chemins edge-disjoint par rapport à DCell et BCube. La principale raison à cela est que les topologies arborescentes placent moins de ports sur les serveurs. Chaque serveur est connecté à un commutateur de périphérique avec une connexion unique dans les topologies arborescentes. Moins de ports sur les serveurs, rend les connexions entre les serveurs et les commutateurs vulnérables aux défaillances. Par exemple, dans les topologies arborescentes, si un commutateur Edge échoue, tous les serveurs qui lui sont connectés seront séparés du reste du centre de données. Cela explique également pourquoi il n'y a pas de redondance des commutateurs edge. DCell et BCube, cependant, ont plus de chemins node-disjoints et les chemins edge-disjoints à cause des ports multiples sur leurs serveurs, ce qui entraîne un niveau de redondance plus élevé. À moins que chacun des chemins node-disjoint ait un échec, deux serveurs resteront toujours connectés. Il est à noter que le niveau de redondance de DCell ou BCube est déterminé par le nombre de ports sur les serveurs. Ce tableau ne montre pas la différence entre le taux de défaillance des différents composants.

Pourquoi nous avons porté notre choix sur la topologie Fat-Tree?

- Fat-Tree a la même bande passante au niveau de toutes les bissections
- Chaque couche a la même bande passante agrégée
- Peut être construit en utilisant des dispositifs avec une capacité uniforme
- Chaque port supporte la même vitesse que l'hôte final
- Tous les appareils peuvent transmettre à la vitesse de la ligne si les paquets sont uniformément distribués le long des chemins disponibles
- > Grande évolutivité: le commutateur k-port prend en charge $\frac{k^3}{4}$ serveurs

1.3. Consommation d'énergie dans les centres de données

Un centre de données, appelé aussi centre Un centre de données, appelé aussi centre de traitement de données est un site physique sur lequel se trouve regrouper des équipements constituant du système d'information de l'entreprise (ordinateurs centraux, serveurs rangés dans des baies, baies de stockage, équipement réseaux et de télécommunication, etc.)¹. Ils doivent être construits pour supporter les demandes croissantes de stockage des données et de calcul. Selon [Da COSTA 2015], en 2014 déjà, les centres de calcul classiques pouvaient contenir plusieurs dizaines de milliers de serveurs sur des surfaces de l'ordre de dix mille mètres carrés. Cette taille permettait d'organiser de manière efficace la gestion des fluides (refroidissement, électricité). Selon la même source, pour un centre de données classique d'environ 40 000 serveurs utilisés par 500 000 services, la consommation électrique représente environ 10 mégawatts. Cette consommation est telle que le prix d'achat des serveurs est dépassé en quelques années par le prix de leur consommation électrique.

Le directeur général Jim Witham de GaNSystems, leader des dispositifs de gestion de l'alimentation à base de nitrure de gallium (GaN), spécialisés dans la conversion de puissance, les semi-conducteurs et les transistors, affirme dans un article² paru en Février 2018, qu'en 2017, environ 8 millions de centres de données dans le monde (allant des petits serveurs aux grands centres de données) géraient la charge des données. Ils avaient besoin pour cela de 416,2 térawattheures d'électricité. Cela représente 2% de la consommation mondiale totale et devrait atteindre 5% de la consommation mondiale d'ici 2020.

Il est alors important de mettre en place des mécanismes permettant d'économiser de l'énergie pour l'ensemble des composants du Cloud.

Pour une bonne maitrise de l'énergie dans le Cloud Computing, on s'intéresse généralement à ses différents composants tels que les centre de données, les serveurs et le système de refroidissement. Selon un rapport de EPA [EPA 2007], les serveurs consomment 80% de l'énergie totale de l'infrastructures informatique et 40% de l'énergie totale des centres de données. Ils constituent l'une des cibles prioritaires dans

¹https://fr.wikipedia.org/wiki/Centre de donn%C3%A9es

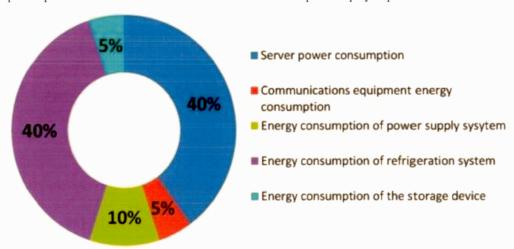
²https://www.datacenterknowledge.com/industry-perspectives/achieving-data-center-energy-efficiency

le dispositif de mise en place d'une politique d'économie d'énergie dans les centres de données [Schappi et al. 2010]. L'énergie consommée par le système de refroidissement est importante, mais elle est proportionnelle à la consommation d'énergie de l'infrastructure informatique.

Des systèmes de refroidissement gratuits utilisés par les grandes compagnies telles que Google, Facebook, ebay ... sont intéressants pour réduire la consommation d'énergie liée au système de refroidissement. Ces approches réduisent la température dans les centres de données en utilisant de l'air ou de l'eau naturellement fraiche au lieu de la réfrigération mécanique. Et cela contribue à diminuer de manière drastique la consommation d'énergie liée au système de refroidissement. Ce gain en énergie pour le refroidissement peut atteindre les 100% dans certaines conditions climatiques.

1.3.1. Les principales causes des pertes d'énergie

Il existe plusieurs facteurs liés à la perte d'énergie dans les centres de données du Cloud Computing. Les statistiques relatives à la consommation d'énergie des centres de données provenant de plusieurs enquêtes sont résumées à la figure 1-8, couvrant principalement le coût des ressources informatiques et physiques.



1-8:Répartition de la consommation d'énergie dans un centre de données [Jing et al. 2017]

1.3.1.1. Les serveurs

Comme indiqué dans la précédente partie, les serveurs font partie des éléments les plus énergivores dans les centres de données. Les principales raisons sont leurs sous-utilisations.

L'augmentation de la taille des datacenters a entrainé une augmentation considérable du nombre de serveurs dans les centres de données. Beaucoup d'études ont montré que les serveurs ne sont pas utilisés de manière efficiente³, [Schappi et al. 2010]. L'*Uptime Institute*, dans son rapport de 2014 [Stansberry 2014], évalue à 20% les serveurs obsolètes, trop vieux ou initialisés. Quant à Google [Barroso et U. Hölzle, 2007], une étude interne a montré qu'en 2007 une grande majorité de leurs serveurs étaient utilisés à moins de 50% de leur capacité. Ces serveurs contiennent pour la plupart des informations qui nécessitent leur disponibilité 24h/24 et 7 jours/7. Ces informations peuvent être des fois des fichiers attachés à certains mails datant de plusieurs années. La disponibilité de ces serveurs entraine une grande consommation électrique et produit de la chaleur qu'il faut extraire.

A cela s'ajoute les pertes d'énergie des serveurs pendant leur état de repos. Dans [Naone 2009], il est indiqué qu'un serveur inactif consomme environ 70% de sa puissance maximale s'il est non utilisé. Les éléments qui sont les principaux responsables de la consommation d'énergie au niveau des serveurs sont le CPU, la mémoire vive, le disque de stockage, les blocs d'alimentation et les bus de connecteur d'extension [Beloglav 2013].

Si les CPUs peuvent accepter de fonctionner en utilisant les modes de faible puissance, les autres composants ne peuvent être complètement ou partiellement mis hors tension. L'adoption des architectures multi-cœurs entraine une diminution de la consommation d'énergie, par contre l'utilisation des puces sur les mémoires DRAM

³https://www.nrdc.org/sites/default/files/data-center-efficiency-assessment-IP.pdf 2014

entraine une augmentation de la consommation d'énergie [Djouhra, Thèse de doctorat 2016]. Ces pertes d'énergie des serveurs sont considérées comme l'une des causes majeures de l'inefficience énergétique des centres de données.

1.3.1.2. Le système de refroidissement

Comme on peut le voir sur la figure 1-8, le refroidissement des centres de données est l'une des causes liées à la perte d'énergie dans les centres de données. Cela est du au fait que les serveurs sont des appareils électroniques de grandes tailles qui produisent énormément de chaleurs. Il est essentiel de maintenir des conditions de température et d'humidité appropriées pour s'assurer du bon fonctionnement des serveurs et de tous les équipements informatiques du centre de données. Ces conditions sont assurées par des systèmes de climatisation qui sont énergivores.

1.4. Techniques de réduction de la consommation d'énergie dans les centres de données

1.4.1. Approches basées sur les méthodes de virtualisation et de migration de machines virtuelles

Les centres de données ont connu ces dernières années une ascension fulgurante. De plus en plus, beaucoup de services sont proposés via le Cloud Computing. Cela pousse les fournisseurs à adopter des technologies et des politiques plus efficaces dans la gestion de l'infrastructure de ces centres de données. L'efficacité énergétique est devenue un enjeu majeur car ces centres de données ont tendance à consommer beaucoup d'énergies. Plusieurs solutions sont proposées dans la littérature pour réduire les impacts liés à l'environnement. Aujourd'hui, la question de l'écologie est omniprésente dans tous les domaines et l'informatique aussi se met au vert.

Le Cloud offre plusieurs techniques permettant la réduction de la consommation d'énergie. De la virtualisation à la consolidation en passant par la migration de

machines virtuelles, plusieurs possibilités s'offrent pour réduire la consommation d'énergie dans les centres de données. La gestion dynamique des composants du centre de données est aussi un moyen permettant de gérer de manière efficace la consommation d'énergie.

Nous présentons dans cette partie un état de l'art sur la gestion de la consommation d'énergie dans les centres de données.

La technologie de virtualisation fait partie des événements ayant contribué à l'événement du Cloud Computing. Elle permet d'instancier plusieurs machines dites "virtuelles" sur une seule machine physique. Ainsi, cette technologie peut permettre de réduire le nombre de machines physiques en cours d'exécution dans un centre de données et participer à la réduction de la consommation d'énergie. Lorsque des machines physiques sont sur-utilisées ou très peu utilisées, les machines virtuelles qui y sont instanciées peuvent être déplacées vers d'autres machines physiques. Cette migration des machines virtuelles peut ainsi participer à l'équilibrage de charge et à la réduction de la consommation d'énergie et peut se faire à chaud ou à froid. On parle dans ce cas de consolidation. La consolidation est une approche efficace pour parvenir à une meilleure efficacité énergétique, elle contribue à améliorer l'utilisation des ressources par de nombreuses migrations de machines virtuelles résidents sur plusieurs serveurs "sous-utilisés" vers un seul serveur [Djouhra, Thèse doctorat 2016].

Plusieurs auteurs ont travaillé sur la réduction de la consommation d'énergie dans le Cloud Computing.

Dans [Beloglazov et al. 2012], les auteurs ont utilisé la technique de migration de machines virtuelles. Les auteurs ont mis l'accent sur la consommation d'énergie du processeur. La mémoire vive, les interfaces réseaux et autres n'ont pas été pris en compte. Des seuils supérieur et inférieur d'utilisation du CPU des serveurs ont été

définis dans leur proposition. Lorsque l'utilisation du processeur de l'hôte est en dessous du seuil inférieur, toutes les machines virtuelles doivent être transférées et la machine hôte est mise en veille afin d'économiser de l'énergie. Par contre, si l'utilisation du processeur dépasse le seuil supérieur, certaines machines virtuelles sont migrés vers d'autres machines physiques pour réduire l'utilisation du CPU. La deuxième phase consiste en la mise en place de la machine virtuelle sélectionnée. L'algorithme MBFD (Modified Best Fit Decreasing) a été utilisé. Plusieurs métriques de performances ont été considérées et les résultats obtenus par les auteurs ont montré que la consommation d'énergie pouvait être réduite jusqu'à 77 et 53% par rapport respectivement à NPA (No Power Aware) et DVFS (Dynamic Voltage Frequency Scaling) avec 5,4% des violations de SLA(Accord de niveau de service). NPA suppose que tous les hôtes fonctionnement à 100% d'utilisation du CPU et consomment une puissance maximale tout le temps, tandis que la technique DVFS permet d'augmenter ou de réduire le voltage et la fréquence du processeur.

Les auteurs [Beloglazov et Buyya 2010] proposent une gestion des ressources énergétiques dans le Cloud Computing en utilisant la technique de migration à chaud avec une activation/désactivation des nœuds utilisant la technique DVFS. L'algorithme est appliqué à une infrastructure de Cloud constituée de nœuds physiques hétérogènes ayant chacun un CPU avec des processeurs performants et disposant de plusieurs machines virtuelles. L'architecture proposée est constituée de deux gestionnaires sur chaque nœud physique et d'un dispatcheur. Lorsqu'une requête est émise, le dispatcheur envoie la demande au gestionnaire global. Le gestionnaire global donne les informations sur l'utilisation de la ressource, les machines virtuelles qui doivent être allouées et l'émission des ordres de migration. Le gestionnaire des machines virtuelles effectue les redimensionnements, les commandes et les migrations selon les ordres reçus. La migration est effectuée lorsque les ressources sont très peu utilisées ou que leur utilisation atteint 100%. Cette migration s'effectue en deux étapes que sont la sélection de la machine à migrer et la mise en place de

cette dernière. Plusieurs politiques sont simulées: NPA, DVFS, ST (Single Threshold), MM (Deux seuils) par le simulateur CloudSim et les résultats expérimentaux montrent la souplesse de la politique MM [Beloglazov et Buyya 2010].

Une méthode pour détecter la sur-utilisation et la sous-utilisation des processeurs des serveurs a été proposée dans [Sekhar et Jeba 2013]. Dans le cas d'une sur-utilisation du CPU, certaines machines virtuelles sont migrées vers d'autres hôtes afin de réduire l'utilisation du serveur sous un seuil déterminé. En cas de sous-utilisation, toutes les machines virtuelles sont migrées vers d'autres hôtes et cette dernière est éteinte à la fin pour réduire la consommation énergétique. Pour l'allocation des machines virtuelles, les auteurs on utilisé des heuristiques. Le simulateur CloudSim a été utilisé pour implémenter la politique de migration et les heuristiques. Les résultats obtenus sont meilleurs que ceux précédemment obtenus dans la littérature.

Dans [Belbekkouche et al. 2012], les études ont montré que lorsque les nœuds et les liens virtuels sont conjointement intégrés, ils procurent un meilleur mappage avec un moindre coût en un temps réduit comparé aux deux approches d'étape d'intégration [Su et al. 2012]. Les auteurs de [Su et al. 2012] ont proposé une intégration nommée "EA-VNE" (Energy-Aware Virtual Network Embedding) qui est basé sur les contraintes du CPU et de la bande passante, néanmoins, leur algorithme ne supporte pas le partage de chemin. De manière similaire, une allocation de ressource prenant en compte l'hétérogénéité des réseaux virtuels et physiques a été proposée dans [Nogueira et al. 2011]. Leurs contraintes d'élimination pour un réseau virtuel intégré sont basées sur l'emplacement, la charge et le nombre de CPU, la quantité de mémoire vive disponible etc.

L'intégration suggérée dans [Yu et al. 2008] permet le partage de chemin et la migration de lien. Cependant, pour une meilleure intégration comme suggérée dans [Nogueira et al. 2011], leur approche [Yu et al. 2008] devrait corréler leur intégration de lien avec le précédent mappage de nœud. Fondamentalement, la plupart de ces

algorithmes ne prennent pas en compte le fait que les ressources appartenant à la fois au réseau physique et virtuel peuvent être hétérogènes.

L'avènement de l'Internet des objets(IoT), avec le développement rapide de la technologie informatique en nuage ainsi que la consommation d'énergie dans les centres de données causent beaucoup d'inquiétude. Le réseau défini par logiciel (SDN) offre une solution réalisable au problème rigide de l'architecture réseau traditionnelle. Des études [Liao et Wang 2018] beaucoup plus récentes ont montré que l'utilisation du SDN pour contrôler l'information de l'ensemble du réseau avec une gestion dynamique de l'équipement peut économiser environ la moitié de l'énergie consommée dans les centres de données.

1.4.2. Approches existantes pour réduire les besoins en énergie du réseau

Des travaux précédents comme [Pries et al. 2011], [Gupta et Singh 2003] et [Zhang et al. 2008] proposent de faire des économies d'énergie au niveau de la couche réseau. Les auteurs de [Zhang et al. 2008] observent qu'un gain d'énergie peut être obtenu soit en reconfigurant de manière dynamique la topologie Internet via le routage, ou plus tôt pendant la phase de la conception du réseau. Nedevschi et al. dans [Nedevschi et al. 2008] proposent de mettre les composants du réseau en veille pendant les temps d'inactivité (ou dans un mode d'économie d'énergie) et ainsi retarder les lots de trafic sur les buffers situés aux bords du réseau. En retardant les paquets, ils peuvent réorganiser le trafic vers des rafales plus cohérentes, et donc produire une organisation plus efficace en énergie. Les auteurs affirment qu'en général, un dispositif qui est utilisé à une fréquence moindre peut permettre une réduction considérable de la consommation d'énergie.

Néanmoins, ces solutions peuvent augmenter le délai et les pourcentages de perte de paquets si un composant donné du réseau est éteint et qu'un lot de paquets arrivent sur ses interfaces. En raison du comportement des flux dans le réseau filaire et les

contraintes de QoS rigoureuses, des rotations des nœuds en mode On/Off, peuvent être très difficile dans le contexte filaire. Il est important de remarquer que les protocoles Internet devraient être modifiés si nous prenons en compte les études comme [Gupta et Singh 2003], [Zhang et al. 2008] et [Pries et al. 2011]. En effet, Gupta et al. dans [Gupta et Singh 2003], décrivent les problèmes qui peuvent se produire sur deux importants protocoles comme OSPF et IBGP lorsque nous mettons des interfaces de commutateurs et de routeurs en mode veille.

Pour s'attaquer au problème de la réduction de la consommation d'énergie, les auteurs de [Pries et al. 2011] ont proposé d'ajouter aux couches de pile un Système de gestion de l'énergie (EMS) qui a pour rôle de surveiller de manière proactive la consommation instantanée d'énergie de chaque lien pilote et routeur, les informations sur les chemins, les taux de trafic et l'état de la QoS des flux. Selon ces informations, le EMS est capable de recommander un ensemble de chemins qui optimisent les QoS des utilisateurs avec la plus basse consommation d'énergie. Par conséquent, un ensemble de nœuds ou certains liens peuvent être éteints afin de réduire la consommation d'énergie. Notez que l'EMS conçu s'exécute sur la couche IP supérieure [Pries et al. 2011]. Néanmoins, une grande quantité de données est collectée par les EMS avant la découverte d'un chemin donné qui diminue la consommation d'énergie avec un certain niveau de QoS.

Malgré un gain de 10 à 20% de consommation d'énergie, les résultats obtenus montrent que le délai mesuré et les paquets perdus augmentent lorsqu'un ensemble de routeurs sont mis hors tension. L'augmentation du délai mesuré ainsi que les pics de perte sont une sérieuse préoccupation pour les applications en temps réel. Un pic de perte momentané est également enregistré depuis des liens donnés pouvant être désactivés alors que certains paquets sont encore stockés dans la file d'attente du routeur [Pries et al. 2011].

Dans un but similaire, Berl et al. [Berl et al. 2010] suggèrent d'optimiser ou de redévelopper les protocoles réseau afin d'obtenir plus d'efficacité énergétique. En

outre, ils suggèrent de réduire le coût énergétique des logiciels des centres de données qui exécutent les applications du Cloud. Une proposition intéressante est de prendre en compte la consommation d'énergie comme paramètre clé d'un fonctionnement optimal lorsque les ingénieurs font le plan de la réplication du contenu et des algorithmes de diffusion pour les services du Cloud.

1.4.3. Gestion et analyse du trafic dans le Cloud privé et public

Baliga et al. [Baliga et al. 2011], en considérant trois différents modèles de services de Cloud comme le logiciel, le stockage et le traitement, étudie l'énergie consommée dans la commutation et la transmission, ainsi que le traitement et le stockage des données selon qu'on ait un Cloud public ou privé. Ces auteurs présentent une étude sur l'énergie utilisée pour transmettre des données sur le Cloud. Ils ont constaté que le transport de données entre le Cloud et l'ordinateur personnel peut consommer des quantités d'énergie encore plus importantes que leur stockage.

En outre, en prenant en compte le stockage comme service du Cloud, ils soutiennent que le stockage dans le Cloud public peut consommer dans l'ordre de trois à quatre fois plus de puissance que le stockage dans le Cloud privé. Ils estiment la consommation d'énergie par bit de la transmission et la commutation pour un Cloud public à environ 2.7 μJ/b [Baliga et al. 2011]. La raison principale est due au fait que les clients des Cloud publics utilisent ses services via Internet. En réalité, les paquets appartenant à un flux donné traversent en moyenne 12 à 14 sauts entre la source et la destination [Kliazovich et al. 2010], [Chase et al. 2001], et donc l'énergie consommée pour transporter ces paquets augmente.

La conclusion la plus importante notée dans [Baliga et al. 2011] est que lorsqu'on veut comparer la consommation d'énergie du Cloud basé sur le services (logiciels, stockage, traitement) avec un ordinateur personnel, nous devrions inclure la consommation d'énergie requise pour transporter les données du client vers les

ressources du Cloud et vice versa. Il est obligatoire de prendre en compte le transport, si les clients utilisent les services du Cloud depuis un Cloud public accessible à travers Internet.

La tendance générale observée dans [Baliga et al. 2011] est que le Cloud privé est plus efficace que les Cloud publics. Cette constatation est liée aux économies d'énergies dans le transport, et c'est donc particulièrement important de calibrer l'énergie consommée dans le transport de données entre les clients et le Cloud afin de progresser vers des Cloud plus économes en énergie.

1.4.4. Économies d'énergie dans les réseaux d'accès

Compte tenu de la grande hétérogénéité des équipements, l'amélioration de l'efficacité énergétique du futur Internet constitue un enjeu majeur. En effet, les auteurs de [Bolla et al. 2011] ont montré qu'en ce qui concerne une configuration typique de réseau ISP/telco, les équipements d'accès (par exemple, les modems, les passerelles domestiques, les équipements de réseau à large bande sans fil, les multiplexeurs d'accès aux liens des abonnés numériques, etc.) consomment environ 70% de l'exigence globale en énergie du réseau, alors que la consommation d'énergie du réseau de transport/noyau est de 30%. La raison principale est due au fait que les nœuds d'accès sont estimés approximativement à 94% du réseau d'accès, et la consommation d'énergie approximative par appareil est de 10000, 6000 et 1600 Watt / h respectivement pour le noyau, le transport et le réseau d'accès [Bolla et al. 2011]. Dans [Gomma et al. 2011], les auteurs proposent le Broadband Hitch-Hiking (BH²), qui permet de rendre "vert" le réseau d'accès. L'objectif est de diminuer l'énergie consommée dans les réseaux d'accès côté utilisateur ainsi que sur les modems DSLAM du fournisseur de services Internet (ISP). Ils suggèrent d'agréger le trafic de plusieurs utilisateurs dans quelques routeurs qui sont disponibles à proximité de la résidence des clients et donc laisser les routeurs restants en mode veille. Cela se fait en donnant la possibilité aux point d'accès sans fil (AP) de se connecter à d'autres AP,

puis envoyer leur trafic pour le routage. Ce faisant, le routeur local d'un client donné peut entrer en mode veille.

Plusieurs problèmes devraient être abordés avant un véritable déploiement de BH². Tout d'abord, il est obligatoire de comprendre comment la coopération demandée entre les propriétaires des passerelles (routeurs) dans la même zone géographique sera effectuée. Aussi, si l'échange de routage, avant d'arriver aux modems ISP DSLAM peut être possible en dépit des fournisseurs du client. Enfin, il est nécessaire de prendre en compte les problèmes de sécurité et de confidentialité, car le trafic des clients sera acheminé via les AP voisins. Par exemple, un AP malveillant peut détourner le trafic ou modifier les données qui doivent être transmises et, par conséquent, protéger l'intégrité des données est nécessaire pour un déploiement réussi de BH².

Nonobstant le fait que les dispositifs d'accès ont de petites exigences en matière d'énergie, l'amélioration de l'efficacité énergétique de l'accès au noyau et au périphérique réseau est obligatoire car l'énergie consommée par les routeurs devient plus importante avec l'augmentation du taux d'accès des utilisateurs [Bolla et al. 2011]. Simplement, la consommation globale d'énergie augmentera au fur et à mesure que la capacité du réseau augmente. En revanche, l'énergie consommée par les périphériques réseau appartenant aux réseaux d'accès est en grande partie indépendante du volume de trafic [Baliga et al. 2011].

Les tendances actuelles (expansion des services basés sur Internet) conduit à une croissance exponentielle du trafic IP global. Cette expansion a été une grande préoccupation tout au long du travail novateur mené par Gupta et al. [Gupta et Singh 2003] et est aujourd'hui une question primordiale pour le futur Internet [Bola et al. 2011]. Par conséquent, l'énergie consommée dans le réseau de transport (Métropolitaine et noyau) devrait être réduite autant que possible puisque les équipements requis pour acheminer le trafic Internet augmentent [Zhang et al. 2008].

1.4.5. Economie d'énergie basée sur l'organisation du centre de données

Heller et al. [Heller et al. 2010] ont publié un document portant sur le compromis entre l'efficacité énergétique et la résilience. Pendant le fonctionnement normal du réseau, les commutateurs supplémentaires utilisés pour les chemins de sauvegarde sont désactivés et seront uniquement activés en cas de panne de réseau ou de charge élevée. Le mécanisme proposé est mis en œuvre dans un banc d'essai où OpenFlow est utilisé pour la gestion des commutateurs. Cependant, ils n'arrêtent que les commutateurs et non les serveurs. Pries et Al. [Pries et al. 2011] montrent que ceux-ci ne consomment qu'entre 5% et 10% de la consommation globale d'énergie.

Kliazovich et al. [Kliazovich et al. 2010] ont développé un environnement de simulation pour calculer la consommation d'énergie de plusieurs architectures de centres de données. En plus de montrer la part de la consommation d'énergie du réseau et du serveur, ils indiquent la quantité d'énergie pouvant être sauvegardée en utilisant le voltage et l'échelle de fréquence dynamique ou la gestion dynamique de l'alimentation. L'un des premiers articles présentant une gestion dynamique des ressources selon la charge proposée est présenté par Chase et al. [Chase et al. 2001]. Ils proposent une architecture où les clusters de serveurs sont redimensionnés dynamiquement conformément aux SLA négociés.

Une approche plus détaillée est présentée par Chen et al. [Chen et al. 2005]. Trois solutions sont proposées pour réduire la consommation d'énergie des serveurs dans un centre de données. Pour la première solution, le comportement de la charge de travail d'un avenir proche est prévu tandis que la seconde est une solution réactive, utilisant des retours d'informations périodiques du système d'exécution. La troisième solution proposée est une solution hybride utilisant une combinaison de prédiction et de rétroaction périodique.

Les auteurs de [Schwartz et al. 2012] évaluent les économies d'énergie en adaptant de manière dynamique le nombre de serveurs selon la charge actuelle en utilisant des seuils sur la file d'attente pour activer des serveurs supplémentaires et sur le nombre

total de requêtes dans le système pour désactiver les serveurs. En outre, ils utilisent une véritable approche d'optimisation multi-objectifs pour trouver un optima entre le temps d'attente et la consommation d'énergie. Pour l'architecture du centre de données, ils utilisent une topologie Three-tier avec deux groupes de serveurs. Le premier groupe de serveurs est toujours actif. Si le nombre de requêtes dans la file d'attente dépasse le seuil fixé, le deuxième groupe de serveurs est activé. Ces derniers le resteront jusqu'à ce que le nombre total de requêtes dans le système dépassent le second seuil fixé. Christian et al cherchent à garantir une qualité de service (un temps de réponse acceptable) tout en économisant de l'énergie. Les résultats obtenus montrent que des configurations existent, de sorte que la consommation d'énergie peut être considérablement réduite tout en conservant un temps d'attente moyen acceptable. Ainsi, ils garantissent des accords de niveau de service à l'utilisateur final tout en conservant environ 40% d'énergie.

DAN et al dans [Liao et al. 2015] abordent le problème de la consommation d'énergie en tenant compte des performances du système dans un cadre d'optimisation général. Plutôt que de considérer un seul groupe de sauvegarde, les auteurs de [Liao et al. 2015] proposent trois groupes de serveurs dont les deux sont des groupes de sauvegarde. Compte tenu des considérations précédentes, à savoir le compromis entre la consommation d'énergie et la performance, et l'attribution correcte de groupes de sauvegarde, ils étudient la relation entre la consommation totale d'énergie et le temps moyen d'attente du service dans une méthode de gestion dynamique élaborée. Étant donné que tous les serveurs sont divisés en trois groupes, les auteurs modélisent un centre de données avec des arrivées simple suivant une distribution poissonienne, un tampon de capacité finie et un temps de service exponentiel. Ils évaluent les économies d'énergie en adaptant dynamiquement les groupes de serveurs en fonction de la charge actuelle en utilisant des seuils d'activation sur la taille de la file d'attente pour démarrer un certain groupe de serveurs. Afin de trouver les seuils d'activation optimum, ils formulent le problème d'optimisation de la minimisation de l'énergie

selon les exigences de performance données en termes de temps moyen d'attente du service.

1.5. Conclusion

Dans ce chapitre, nous avons passé en revu les travaux relatif à la consommation d'énergie dans les centres données. Nous avons commencé par présenter quelques topologies des centres de données, par la suite, nous avons évoqué la consommation d'énergie avec les principales causes des pertes d'énergie dans les centres de données. Quelques travaux permettant de réduire la consommation d'énergie dans les centres de données sont aussi exposés dans cette partie.

Le chapitre suivant présentera notre approche et nos propositions permettant de rendre plus écologique les centres de données.

2. Chapitre 2: GESTION DE L'ENERGIE DANS LES CENTRES DE DONNEES (GREENPOD)

2.1. Introduction

La demande croissante de services Internet et de Cloud Computing a considérablement augmenté le nombre de serveurs et de Commutateurs opérant dans les grands centres de données. Cela entraîne également une augmentation énorme de la consommation d'énergie dans ces centres de données massifs. Malgré que les centres de données soient mis en place à des fins et applications différentes, la structure de base est similaire. Le fournisseur de services vise à réaliser des revenus élevés tout en garantissant les accords de niveau de service (SLA).

Cependant, le niveau de charge moyen d'un centre de données est d'environ 60% de la charge de pointe [Benson et al. 2010].

Même si plusieurs serveurs ne sont pas sous charge, ils consomment encore environ 65% de la consommation maximale [Greenberg et al. 2009].

Pour réduire la consommation d'énergie, il est important de développer un moyen efficace d'éteindre les serveurs inutilisés pendant les heures de charge de travail faible [Mazzucco et al. 2010], [Grid 2010].

Certains auteurs ont travaillé sur la gestion dynamique dans les centres de données et suggèrent qu'un seul groupe de serveurs de sauvegarde soit activé / désactivé dynamiquement [Gandhi et al. 2010], Christian et al ont montré qu'ils ont un gain de 40% en consommation d'énergie par rapport au centre de donnée par défaut dans une topologie three-tiers avec un groupe de serveur de sauvegarde dynamique [Schwartz et al. 2012]. Dan et al montre qu'un nombre approprié de groupes de sauvegarde (deux groupes) peut avoir un meilleur résultat [Liao et al. 2015]. Dans notre étude nous avons utilisé la topologie Fat-tree K-aray et nos résultats montrent que nous

économisons beaucoup plus d'énergie que Dan et al avec un nombre de groupes de sauvegarde plus important.

Cependant, le coût d'installation et le temps de configuration des serveurs de commutation ne peuvent pas être négligés [Gandhi et al. 2010].

Dans ce document nous asseyons de trouver un compromis entre la réduction de la consommation d'énergie et un temps d'attente acceptable. Les serveurs étant divisés en un nombre de groupes dépendant de la valeur de k, la décision de les activer ou de les désactiver est basée sur un nombre de seuils qui doivent être correctement dimensionnés en fonction du niveau de charge exprimé par la longueur de la file d'attente.

2.2. Contexte de la proposition

La consommation d'énergie des centres de données ne cesse d'augmenter ces dernières années. Le chercheur suédois et expert en analyse du cycle de vie chez Huawei, Anders Andrae [Andrea 2017], prédit que d'ici 2025, les centres de données représenteront la plus grande part des TIC dans la production mondiale d'électricité (33%), suivis des smartphones (15%) des réseaux (10%) et de la télévision (9%).

En 2010, leur consommation représentait environ 1,3% de la consommation totale d'électricité dans le monde [Koomey 2011] et leur consommation d'énergie peut être triplée de 2010 à 2020 si leur efficacité énergétique reste au niveau de 2010 [Shehabi et al. 2016]. Les deux principaux consommateurs d'énergie des centres de données étant son équipement informatique plus particulièrement les serveurs, et ses systèmes de refroidissement [Dayarathna et al. 2016], les technologies visant à réduire la consommation énergétique de ces deux composants dans les centres de données sont en développement rapide ces dernières années [Ni et Bai 2017].

Plusieurs auteurs proposent de nouvelles architectures pour les centres de données dont le déploiement est moins couteux et offre plus de résilience [Baliga et al. 2011] -

[Fischer et al. 2013]. Des techniques existent et visent principalement à réduire la consommation d'énergie des serveurs [Mardani et al. 2015]. Actuellement, les fournisseurs de services dans le Cloud trouvent d'autres moyens de réduire la quantité de carbone de leur infrastructure [Mardani et al. 2015]. Les grands fournisseurs du Cloud tels que Google, Amazon, Microsoft et IBM espèrent alimenter leurs centres de données en utilisant des sources d'énergie renouvelables [Sing et al. 2017]. Les futurs centres de données Cloud sont tenus de fournir des services avec un minimum d'émissions de carbone et de chaleur sous forme de gaz à effet de serre [Buyya et Sing 2018].

Des travaux ont étudié des cas dynamiques de gestion des serveurs dans le Cloud [Gandhi et al. 2010], [Ghamkhari et Mohsenian-Rad 2013], [Guenter et al. 2011] qui activent et désactivent les serveurs en fonction de la charge actuelle.

Notre proposition est comparée aux travaux de Dan et al. [Liao et al. 2015]. En effet, dans leurs propositions, ils divisent le nombre de serveurs en trois groupes. Le premier est toujours actif et les deux autres sont utilisés dynamiquement en fonction des seuils d'activation basés sur la taille de la file d'attente. Ils travaillent avec une architecture tree-tiers. Dans le cas d'un grand centre de données, constitués de milliers de serveurs, il peut arriver régulièrement qu'on active avec leur proposition, un groupe de serveur pour traiter quelques requêtes. Le nombre de serveur d'un groupe dans de tels centre de données peut être très grand. Dans certains cas, quelques serveurs du groupe peuvent régler le problème.

Les travaux de Dan et al [Liao et al. 2015] sont similaire à ceux de Schwartz et al [Schwartz et al. 2012], ils divisent le nombre de serveurs en trois groupes; le premier est toujours actif et les deux autres sont utilisés dynamiquement en fonction des seuils d'activation basé sur la taille de la file d'attente. Ils travaillent avec une architecture tree-tiers.

Dans nos travaux, Nous utilisons des modèles mathématiques et de simulation pour réduire la consommation d'énergie dans les centres de données. Nous utilisons la topologie fat-tree et nous démarrons avec le quart des serveurs et Commutateurs c'est à dire $\frac{k}{4} * \left(\frac{k}{2}\right)^2$ serveurs et $\frac{k}{4} * k$ Commutateurs.

Les autres groupes (un groupe de serveurs constitue le nombre de serveurs dans un POD) sont activés et désactivés en fonction des seuils d'activation basés sur la taille de la file d'attente. Pour éviter des pertes de requêtes, on vérifie si les serveurs ont terminé leur traitement en cas de désactivation, au cas échéant on les désactive sinon on replace les requêtes en tête de file avant de désactiver les serveurs. Nous montrons par la suite que nous économisons plus d'énergie que la proposition de Dan et al. mais aussi nous avons un gain énorme en énergie par rapport à un centre de données par défaut.

2.3. Présentation du modèle

Le modèle de centre de données que nous considérons est un K-array Fat-Tree et nous démarrons initialement le quart des POD. En fonction de la demande, nous activerons les autres POD mis initialement hors tension et si la demande diminue, nous les désactiverons avec les ports des Commutateurs Core correspondants. Pour parvenir à cela, nous allons fixer des seuils d'activation et de désactivation.

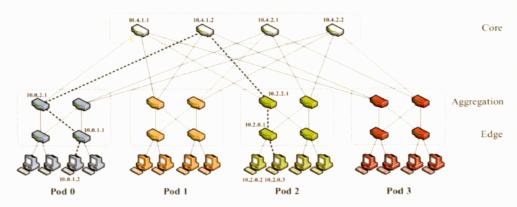


Figure 2-1: Topologies Fat-tree

Contrairement aux propositions [Schwartz et al. 2012] et [Liao et al. 2015], nous utiliserons la topologie Fat_Tree permettant d'économiser plus d'énergie par rapport aux autres topologies. Nous ne limiterons pas non plus le nombre de groupes de serveurs à ajouter à un ou deux, ce nombre dépendra de k, c'est-à-dire du nombre de POD, mais aussi les Commutateurs non utilisés seront éteints. Notre objectif est d'évaluer la quantité d'énergie consommée dans le centre de données depuis les routeurs core jusqu'aux serveurs et la quantité d'énergie économisée quand les serveurs et Commutateurs non utilisés sont mis hors tension. Initialement, seuls les commutateurs core et le quart des PODs sont actifs, ce qui permet d'économiser aussi bien de l'énergie au niveau des serveurs mais aussi au niveau des Commutateurs.

Considérons le modèle de la figure 2-2 suivante constitués de serveurs homogènes.

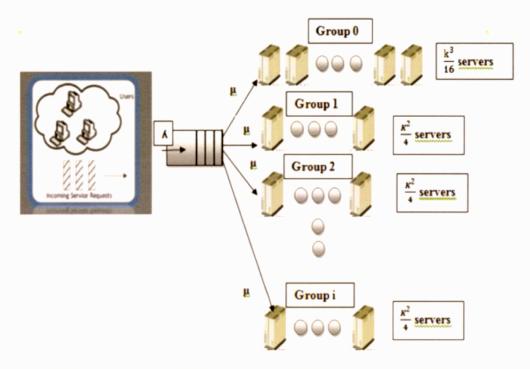
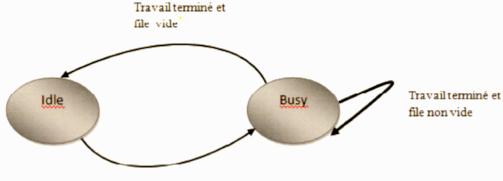


Figure 2-2: Modèle du système

Nous supposons que les requêtes arrivent suivant une distribution poisonienne de taux $\mathbf{\Lambda}$ et que chaque serveur exécute une seule requête à la fois avec un temps de service suivant une distribution exponentielle de taux $\mathbf{\mu}$. Le système peut être modélisé comme une file $M/M/\frac{K^3}{4}$, avec $\frac{K^3}{4}$, le nombre total de serveurs dans une topologie fat-tree k-aray.

Quand une requête arrive dans le système, elle est allouée à un serveur à l'état idle parmi les $\frac{K^3}{16}$ constituant le quart des serveurs qui restent toujours allumés quelque soit le nombre de requêtes dans le système. Si les serveurs sont occupés, la requête est mise en attente jusqu'à ce qu'un serveur soit disponible. Le comportement d'un serveur de ce groupe est illustré dans la figure 2-3.



Nouvelle requête allouée au serveur

Figure 2-3: Comportement d'un serveur du premier groupe

Ce premier groupe de serveurs se comporte comme un centre de données par défaut. Soit P une variable aléatoire qui détermine le nombre de requêtes dans le système et P_j la probabilité qu'il y ait j requête dans le système. Initialement les $\frac{k}{2}$ Commutateurs edge et agrégations dans chaque POD sont sous tension ainsi que les $\left(\frac{k}{2}\right)^2$ Commutateurs core.

Si $j < \frac{k^3}{4}$ requêtes, les j serveurs sont occupés et consomment e_{busy} et les $\frac{k^3}{4}$ – j serveurs sont inoccupés et consomment e_{idle} . La borne supérieure de la consommation d'énergie est :

$$E_{max} = \left[\sum_{j=0}^{\frac{k^3}{4}} P_j * \left(j * e_{busy} + \left(\frac{k^3}{4} - j \right) e_{idle} \right) + e_{busy} * \sum_{j=\frac{k^3}{4}+1}^{\infty} P_j \right] + \frac{k^2}{2} \left(e_{busyS_e} + e_{busyS_a} \right) + \frac{k^2}{2} e_{busyS_c}$$

Avec

 $e_{\it busy}$ l'énergie du serveur en traitement

 e_{idle} l'énergie du serveur à l'état idle

 e_{busyS_a} l'énergie du commutateur aggreagtion

 $e_{busyS-e}$ l'énergie du commutateur edge

e_{busyS} cl'énergie du commutateur core

La borne minimale de la consommation d'énergie correspond au cas où les serveurs inoccupés sont automatiquement éteints et consomment e_{off} .

 E_{\min}

$$\sum_{j=0}^{\frac{k^{3}}{4}} P_{j} * \left(j * e_{busy} + \left(\frac{k^{3}}{4} - j\right) e_{off}\right) + e_{busy} * \sum_{j=\frac{k^{3}}{2}+1}^{\infty} p_{j} + \left(nbP * \frac{k}{2}\right) \left(e_{busySe} + e_{busySa}\right) + \left(\frac{k^{2}}{2} - nbP * \frac{k}{2}\right) + \frac{k^{2}}{8} * e_{busySe} + e_{busySe}$$

Avec nbP=nombre de POD actifs.

Le comportement d'un serveur parmi ceux qui sont réservés est illustré à la figure 2-4.

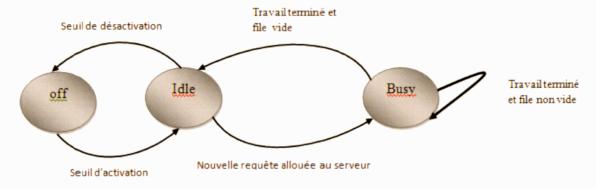


Figure 2-4: Comportement d'un serveur des groupes restants

Dans notre approche, nous avons i seuils d'activation et i seuils de désactivation. Le système est divisé en k groupes de serveurs (POD). Initialement, les $\frac{k}{4}$ groupes de serveurs sont désactivés et nous avons un nombre de serveurs de base actifs S_{a0} égal à $\frac{K}{4}*\left(\frac{k}{2}\right)^2$ (nombre de POD actifs * nombre de serveurs dans chaque POD) ce qui donne $\frac{k^3}{16}$ serveurs actifs. Le ième groupe de serveur sera activé si le nombre de

requêtes dans la file dépasse le seuil θ_i , avec $\theta_i \in [S_{a_{l-1}}, S_{a_{l-1}} + \left(\frac{k}{2}\right)^2]$ et $S_{a_{l-1}}$, le nombre de serveurs actifs au niveau i-1. Autrement dit dès qu'il y a $S_{a_{l-1}} + \theta_i$ requêtes dans le système, le groupe i est activé. Ce groupe de serveurs restera actif jusqu'à ce que le nombre total de requêtes dans la file soit égal à $S_{a_{l-1}}$.

Exemple : Le deuxième groupe de serveurs sera activé si le nombre de requêtes dans la file dépasse le seuil $\theta_1 avec$ $\theta_1 \in \left[\frac{k^3}{16}, \frac{k^3}{16} + \left(\frac{k}{2}\right)^2\right]$ c'est-à-dire dès qu'il y ait $\frac{k^3}{8} + \theta_1$ requêtes dans le système. Ce groupe de serveurs sera désactivé si le nombre de requêtes en attente dans la file est inférieur ou égale à $\frac{k^3}{16}$.

Le tableau suivant présente pour chaque niveau, son seuil d'activation et de désactivation ainsi que le nombre de serveurs actifs.

Tableau 2-1: Nombre de serveurs actifs et les seuils d'activation et de désactivation à chaque niveau

Niveau	Nombre de serveurs actifs	Seuil d'activation θ_i (Nombre de requêtes)		Seuil	de
	S _i			désactivation	
				(Nombre	de
				requêtes)	
		En attente (θ_i)	Système($S_{i-1} + \theta_i$)	En attente	
0	$S_0 = \frac{\kappa^3}{16}$	-	-	-	
1	$S_{1}=S_{0}+\frac{K^{2}}{4}=\frac{K^{3}}{16}+\frac{K^{2}}{4}$	$\theta_1 \in [S_0, S_0 + \frac{K^2}{4}]$	$S_0 + \theta_1$	S ₀	
2	$S_2 = \frac{K^3}{16} + \frac{K^2}{2}$	$\theta_2 \in \left[S_1, S_1 + \frac{K^2}{4} \right]$	$S_1 + \theta_2$	S ₁	
3	$S_3 = \frac{K^3}{16} + \frac{3K^2}{4}$	$\theta_3 \in \left[S_2, S_2 + \frac{K^2}{4} \right]$	$S_2 + \theta_3$	S ₂	
•••			•••		
I	$S_{i} = \frac{K^{3}}{16} + i \frac{K^{2}}{4}$	$\theta_i \in [S_{i-1}, S_{i-1}]_{a_{i-1}} + \frac{K^2}{4}]$	$S_{i-1} + \theta_i$	S _{i-1}	

2.3.1. Modèle analytique

Dans cette partie, nous utilisons le modèle de la théorie des files d'attentes pour la réduction de la consommation d'énergie dans les centres de données.

Définition des files d'attente : Les files d'attente sont essentiellement caractérisées par 3 paramètres notés sous la forme A/S/N (notation de Kendall) où:

A Représente la loi associée aux arrivées dans la file.

S Loi associée au temps de service des serveurs.

N Nombre de serveurs.

Quand il n'y a pas d'autres paramètres, les autres caractéristiques sont par défaut:

- Capacité maximale de la file d'attente infinie.
- Nombre d'usagers infini.
- Discipline de service FIFO.

Les lois sont classées parmi:

M Loi exponentielle (Markovienne).

D Loi déterministe.

G Loi générale quelconque.

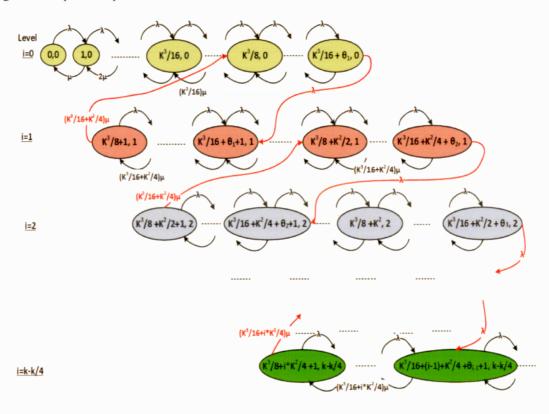


Figure 2-5: Notre modèle du système M/M/(k^3)/4

Pour représenter l'espace d'état de notre système, notre modèle est considéré comme un ensemble de systèmes S_i^j où j est le nombre de requêtes dans le système.

$$i = \begin{cases} 0 \text{ si seuls les serveurs de base } S_{a_0} \text{ sont activ\'es} \\ 1,2, \dots 3 * \frac{k}{4} \text{ si les serveurs reserv\'es sont activ\'es} \end{cases}$$

Le système active le groupe de serveurs i s'il y a au moins θ_i requêtes dans la file d'attente c'est-à-dire s'il y a $S_{a_{i-1}}+\theta_i$ requêtes dans le système. Ce groupe i de serveurs est désactivé si le nombre de requêtes en attente est égale à $S_{a_{i-1}}$.

Soit P une variable aléatoire, P_j^i , la probabilité qu'il ait j requêtes dans le système à l'état i.

Nous commençons par formuler les premières équations d'état du système S_0^j qui correspond à l'état du système où aucun groupe de serveurs réservés n'est encore activé.

$$\begin{split} j\mu P_{j}^{0} &= \lambda P_{j-1}^{0} &\qquad \qquad 0 < j < S_{a_{0}} \\ S_{a_{0}} \mu P_{j}^{0} &= \lambda P_{j-1}^{0} &\qquad \qquad S_{a_{0}} \leq j < 2S_{a_{0}} \\ S_{a_{0}} \mu P_{j}^{0} + S_{a_{1}} \mu P_{2S_{a_{0}}}^{1} &= \lambda P_{j-1}^{0} &\qquad 2S_{a_{0}} \leq j < S_{a_{0}} + \theta_{1} \end{split}$$

Les équations des états des systèmes S_i^j (état du système ou les groupes de serveurs i sont activés) sont les suivantes :

$$\begin{aligned} Pour \ i \leftarrow 1 \ \grave{a} \ 3 * \frac{k}{4} \\ S_{a_i} \mu P_j^i &= \lambda P_{j-1}^i + \lambda P_{Sa_{i-1}}^{i-1} + \theta_i \\ S_{a_i} \mu P_j^i &= \lambda P_{j-1}^i \\ S_{a_i} \mu P_j^i &= \lambda P_{j-1}^i \\ S_{a_i} \mu P_j^i + S_{a_{i+1}} \mu P_{2Sa_i}^1 &= \lambda P_{j-1}^i \\ \frac{k^3}{4} \mu P_j^{\frac{k}{2}} &= \lambda P_{j-1}^{\frac{k}{2}} \\ j \geq \frac{k^3}{2} \end{aligned}$$

2.3.2. Calcul des probabilités

$$P_{j}^{0} = \left(\frac{\lambda}{\mu}\right)^{j} \frac{1}{j!} P_{0}^{0} \qquad 0 < j \le S_{a_{0}}$$

$$P_{j}^{0} = \frac{\left(\frac{\lambda}{\mu}\right)^{j}}{S_{a_{0}}^{j-S_{a_{0}}}S_{a_{0}}!}P_{0}^{0}$$

$$S_{a_{0}} < j \le 2S_{a_{0}}$$

$$P_{j}^{0} = \left(\frac{\lambda}{S_{a_{0}}\mu}\right)^{j-2S_{a_{0}}} * \frac{\left(\frac{\lambda}{\mu}\right)^{2S_{a_{0}}}}{S_{a_{0}}!S_{a_{0}}} P_{0}^{0} - \sum_{m=0}^{j-2S_{a_{0}}-1} \frac{S_{a_{1}}}{S_{a_{0}}} \left(\frac{\lambda}{S_{a_{0}}\mu}\right)^{m} P_{2S_{a_{0}}+1}^{1} \qquad 2S_{a_{0}} < j \leq S_{a_{0}} + \theta_{1}$$

$$\Rightarrow P_{j}^{0} = \left(\frac{\lambda}{S_{a_{0}}\mu}\right)^{j-2S_{a_{0}}} * \frac{\left(\frac{\lambda}{\mu}\right)^{2S_{a_{0}}}}{S_{a_{0}}!S_{a_{0}}^{S_{a_{0}}}} P_{0}^{0} - \frac{S_{a_{1}}}{S_{a_{0}}} * \frac{\left(1 - \frac{\lambda}{S_{a_{0}}\mu}\right)^{j-2S_{a_{0}}}}{1 - \frac{\lambda}{S_{a_{0}}\mu}} P_{2S_{a_{0}}+1}^{1}$$

$$2S_{a_0} < j \leq S_{a_0} + \theta_1$$

Avec
$$\lambda P_{S_{a_0}+\theta_1}^0 = S_{a_1} \mu P_{2S_{a_0}+1}^1$$

$$\Rightarrow P_{2S_{\alpha_{0}}+1}^{1} = \frac{\left(\frac{\lambda}{S_{a_{0}}\mu}\right)^{\theta_{1}-S_{\alpha_{0}}} * \frac{\left(\frac{\lambda}{\mu}\right)^{2S_{\alpha_{0}}}}{S_{a_{0}}!S_{a_{0}}^{S_{\alpha_{0}}}}}{\frac{S_{a_{0}}!S_{a_{0}}^{S_{\alpha_{0}}}}{1-\frac{\lambda}{S_{\alpha_{0}}\mu}} + \frac{S_{a_{1}}\mu}{\lambda}}{P_{0}^{0}}$$

Pour
$$i \leftarrow 1 \grave{a} 3 * \frac{k}{4}$$

$$P_{2S_{a_{i-1}}+1}^{i} = \frac{\left(\frac{\lambda}{S_{a_{i-1}}\mu}\right)^{\delta_{i}-S_{a_{i-1}}}}{\frac{S_{a_{i}}\mu}{\lambda} + \frac{S_{a_{i}}}{S_{a_{i-1}}} * \sum_{m=0}^{\delta_{i}-S_{a_{i-1}}-1} \left(\frac{\lambda}{S_{a_{i-1}}\mu}\right)^{m} P_{2S_{a_{i-1}}}^{i-1}$$

$$P_{j}^{i} = \sum_{m=0}^{j-2S_{a_{i-1}}} \left(\frac{\lambda}{S_{a_{i}}\mu}\right)^{m} P_{2S_{a_{i-1}}+1}^{i} \qquad 2S_{a_{i-1}} < j \le S_{a_{i-1}} + \theta_{i}$$

$$P_{j}^{i} = \left(\frac{\lambda}{S_{a_{i}}\mu}\right)^{j-(S_{a_{i-1}}+\theta_{i}+1)} P_{S_{a_{i-1}}+\theta_{1}+1}^{i} + \sum_{m=0}^{j-1-(S_{a_{i-1}}+\theta_{i}+1)} \left(\frac{\lambda}{S_{a_{i}}\mu}\right)^{m} * \frac{1}{S_{a_{i}}\mu} \left(S_{a_{i}}\mu P_{S_{a_{i-1}}+\theta_{i}+1}^{i} - \lambda P_{S_{a_{i-1}}+\theta_{i}}^{i} - \lambda P_{S_{a_{i-1}}+\theta_{i}}^{i}\right)$$

$$S_{a_{i-1}} + \theta_i < j \le 2S_{a_j}$$

$$P_{j}^{\frac{k}{2}} = \left(\frac{\lambda}{\frac{k^{3}}{4}\mu}\right)^{j-2S_{a_{\frac{3*k}{4}}}+1} P_{2S_{a_{\frac{3*k}{4}}}+1}^{\frac{k}{2}} \qquad j \ge \frac{k^{3}}{2}$$

Avec
$$\sum_{j=0}^{S_{a_{i-1}}+\theta_i} P_j^0 + \sum_{i=1}^{\frac{k}{2}} \sum_{j=2S_{a_{i-1}}+\theta_i}^{S_{a_{i-1}}+\theta_i} P_j^i + \sum_{j=2S_{a_{\frac{3-k}{4}}-1}}^{\infty} P_j^{\frac{3-k}{4}} = 1$$

2.3.3. Calcul des métriques de performances

Le taux d'utilisation $\rho = \frac{\lambda}{\frac{k^3}{4}\mu}$

Le nombre de clients en attente est:

$$\Omega = \sum_{i=0}^{\frac{k}{4}-1} \left(\sum_{j=S_{a_i}}^{S_{a_j}+\theta_{i+1}} (j-S_{a_i}) P_j^i \right) + \sum_{j=S_{a_{\frac{3^*k}{4}}}}^{\infty} \left(j-S_{a_{\frac{3^*k}{4}}} \right) P_j^{\frac{3^*k}{4}}$$

D'après la formule de Little, le temps d'attente moyen $T_a = \frac{\Omega}{\lambda}$

L'énergie consommée au niveau des serveurs et au niveau des Commutateurs edge agrégation et core est représenté dans la formule suivante :

$$E = \left\{ \sum_{j=0}^{S_{a_0}} P_j^0 \left(j e_{busy} + (S_{a_0} - j) e_{idle} + (S_{a_{\frac{k}{2}}} - S_{a_0}) e_{off} \right) + \sum_{j=S_{a_0}}^{2S_{a_0}} P_j^0 \left(S_{a_0} e_{busy} + (S_{a_{\frac{k}{2}}} - S_{a_0}) e_{off} \right) + \sum_{j=S_{a_0}}^{2S_{a_0}} P_j^0 \left(S_{a_0} e_{busy} + (S_{a_{\frac{k}{2}}} - S_{a_0}) e_{off} \right) + \sum_{j=S_{a_0}}^{2S_{a_0}} P_j^0 \left(S_{a_0} e_{busy} + (S_{a_{\frac{k}{2}}} - S_{a_0}) e_{off} \right) + P_{j>\frac{k^3}{2}}^{3S} S_{a_{\frac{3k}{4}}} e_{busy} + (S_{a_{\frac{k}{2}}} - S_{a_0}) e_{off} \right) + P_{j>\frac{k^3}{2}}^{3S} S_{a_{\frac{3k}{4}}} e_{busy}$$

$$+ nbP * \left(\frac{k}{2} e_{busyS_e} + \frac{k}{2} e_{busyS_e} \right) + \frac{k^2}{4} e_{busyS_e}$$

Cette quantité d'énergie correspond à l'énergie consommée depuis les Commutateurs core jusqu'aux serveurs.

2.4. Résultats numérique

Nous avons commencé par le modèle analytique pour montrer que notre proposition est meilleure que celle de Dan et al. Nous avons utilisé k=8 ie 128 serveurs avec un nombre de requêtes dans le système variant entre 32 et 400 avec une charge du système $\rho=0,7$.



Figure 2-6: La puissance consommée en fonction du nombre de clients dans le système

La figure 2-6 montre que l'énergie augmente en fonction du nombre de clients dans le système puis devient stable si le nombre total de serveurs est atteint.

Dans la simulation, nous avons considéré un système composé de 432 serveurs (k=12 pour une topologie Fat-Tree) ayant chacun un temps de service de 0.2 seconde, la puissance consommée par un serveur occupé est e_{buv} =150w, celle d'un serveur à l'état idle est e_{ulle} =240w et celle d'un serveur éteint est e_{off} =0w. Pour notre proposition et celle de DAN et al, nous avons subdivisé les serveurs en groupes. Au début de la simulation, nous activons un groupe de serveurs donné. En fonction de la taille de la file d'attente, nous activons ou désactivons un nouveau groupe de serveurs. Pour la proposition de DAN et al, les serveurs sont subdivisés en trois groupes de n, m1 et m2 serveurs. n, m1 et m2 étant des entiers. Dès que les seuils d'activation Θ_1 et θ₂ sont atteints, on active le groupe suivant. Nous avons choisi ces seuils d'activation et de désactivation en rapport avec leur proposition ($n \le \Theta_1 \le n + m_1$ et $m_1 \le \Theta_2 \le n + m_1$ m₂+m₁). On désactive un groupe si le nombre de clients dans le système passe en dessous du nombre de serveurs activés. Pour notre proposition, nous avons choisi d'activer un nouveau groupe de serveurs dès que la taille de la queue dépasse le nombre de serveurs après activation. Nous désactivons un groupe dès que la taille de la queue est inferieure au nombre de serveurs activés. Lors de la désactivation, notre algorithme récupère l'ensemble des requêtes en cours d'exécution sur les serveurs à désactiver et les replace en tête de file avec leurs temps d'arrivée dans le système pour éviter de les perdre ou les obliger à reprendre leurs demandes. Les résultats obtenus montrent que dans toutes les configurations (système peu sollicité, moyennement sollicité ou trop sollicité), nous économisons plus d'énergie tout en gardant un temps de réponse acceptable par rapport aux autres propositions.

Pour un système relativement peu sollicité (*Probabilité des inter-arrivées* =0.45), nous avons noté une consommation moyenne d'énergie pour notre proposition de 24167.61 watt. Pour la proposition de DAN et al, cette consommation moyenne est

estimée à 29567.61 watt, tandis que pour un centre de données par défaut où on démarre l'ensemble des serveurs, la consommation moyenne est estimée à 72767.61 watt. Ainsi, par rapport au centre de données par défaut, nous économisons 66.79% d'énergie, tandis que DAN et al économisent 59.37%. La figure 2-7, montre la consommation d'énergie en fonction du temps de simulation. Les systèmes étant peu sollicité, aucun nouveau groupe de serveurs n'est activé.

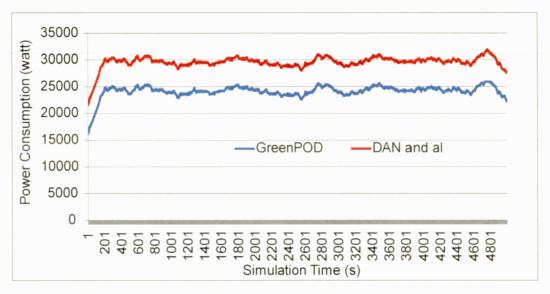


Figure 2-7: Consommation d'énergie en fonction du temps de simulation pour un système peu sollicité

Pour un système un peu plus sollicité (*Probabilité des inter-arrivées* =0.65), notre consommation moyenne d'énergie est estimée à 29242.66 watt. Pour la proposition de DAN et al, cette consommation moyenne est estimée à 33189.05 watt, tandis que pour le centre de données par défaut, elle est égale à 76389.05 watt. Donc par rapport à un centre de données par défaut, nous économisons 61.72% d'énergie, tandis que DAN et al économisent 56.55% d'énergie. La figure 2-8 présente les consommations d'énergie en fonction du temps de simulation, et nous remarquons des activations et désactivations pour notre proposition. Dan et al ayant démarré plus de serveurs, le système n'a pas eu besoin de démarrer un nouveau groupe de serveurs.

La figure 2-9 présente la consommation d'énergie ainsi que le temps d'attente des clients en fonction du temps de simulation. Nous pouvons remarquer qu'à chaque fois que nous activons un groupe de serveurs, le temps d'attente diminue légèrement. Lors des différentes désactivations, nous remarquons une augmentation brusque du temps d'attente. Cela s'explique par le fait que les clients qui étaient entrain d'être servis au niveau des serveurs à désactiver sont replacés en tête de file avec leurs temps d'arrivées initiaux dans le système qui sont forcément inférieurs aux autres clients en attente. Leurs réels temps d'attente est donc égale au temps courant moins leurs temps d'arrivées initiaux dans le système. Ayant repris leurs services, ces requêtes ont plus attendus que les autres, mais cela est plus avantageux que de les abandonner ou de les replacer en queue de file. Dès que ce groupe de requêtes terminent leurs services, le temps d'attente redevient normal. Cette hypothèse n'a pas était évoquée dans les travaux de DAN et al.



Figure 2-8: Consommation d'énergie en fonction du temps de simulation pour un système moyennement sollicité

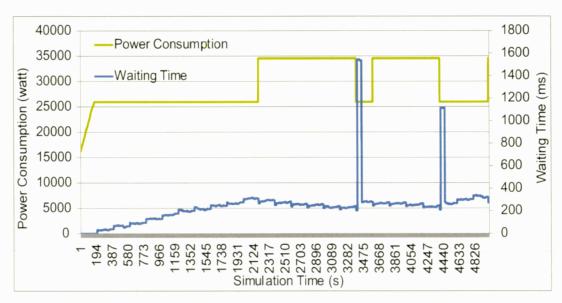


Figure 2-9: Consommation d'énergie et temps d'attente de notre proposition en fonction du temps de simulation pour un système moyennement sollicité

Pour un système très sollicité (*Probabilité des inter-arrivées* =0.95), la consommation moyenne d'énergie pour notre proposition est de 42652.39 watt. Pour la proposition de DAN et al, cette consommation moyenne est estimée à 45447.26 watt, tandis que pour un centre de données par défaut, elle est estimée à 81637.16 watt. Par rapport au centre de données par défaut, nous économisons 47.75% d'énergie, tandis que DAN et al économisent 44.33%. La figure 2-10 présente les consommations d'énergie en fonction du temps de simulation. Les systèmes étant très sollicités, dans toutes les deux propositions, il y a eu des activations et désactivations de groupes de serveurs.

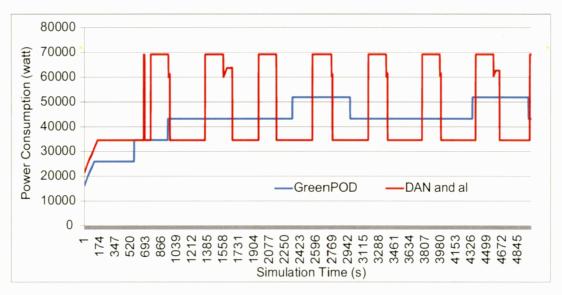


Figure 2-10: Consommation d'énergie en fonction du temps de simulation pour un système très sollicité

La figure 2-11 présente les consommations d'énergie et temps d'attente de notre proposition pour un système très sollicité. Nous faisons la même remarque sur la variation du temps d'attente à chaque désactivation d'un groupe de serveurs.

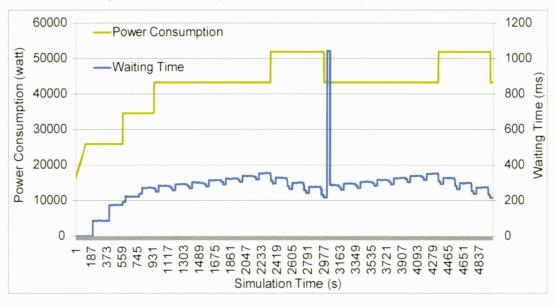
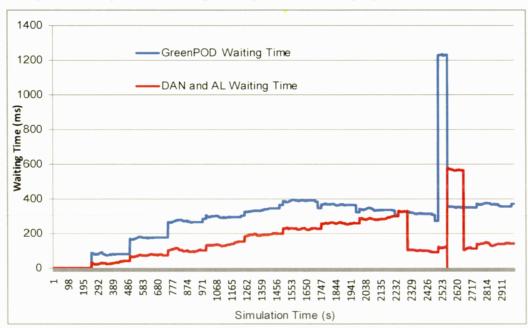


Figure 2-11: Consommation d'énergie et temps d'attente de notre proposition en fonction du temps de simulation pour un système très sollicité



La figure 2-12 représente le temps de réponse des deux propositions.

Figure 2-12: Temps d'attente des deux propositions en fonction du temps de simulation

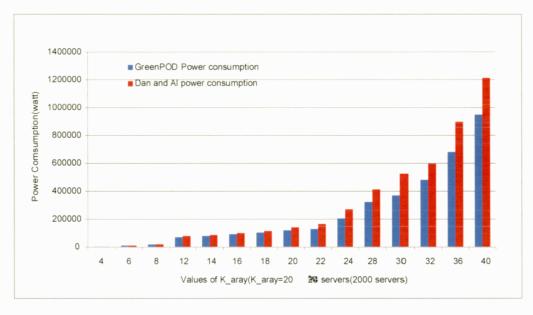


Figure 2-13: Consommation d'énergie des deux propositions en fonction du nombre de serveurs

La figure 2-13 représente la consommation d'énergie en fonction du nombre de serveurs dans chaque modèle. Nous remarquons que plus le nombre de serveurs augmente plus le gain en énergie devient significatif. Dans la simulation, nous avons augmenté le temps de simulation en fonction de la valeur de k_aray. Cela nous permet d'avoir des résultats plu précis.

La figure 2-14 suivante nous permet de savoir le gain en énergie par rapport à un centre de données par défaut.



Figure 2-14: Consommation d'énergie de notre proposition par rapport à un centre de données par classique

Les résultats de la figure 2-14 montre que pour un k_aray=[16-80], nous arrivons à obtenir un gain en énergie variant entre 59% et 74%.

Les résultats ci-dessous montrent la consommation d'énergie au niveau des serveurs. Cependant nous avons calculé la consommation d'énergie au niveau des Commutateurs. Dans notre simulation nous avons fait varié le k_aray entre [4-36] avec un temps de simulation inferieur au résultat ci-dessus et nous avons évalué la

consommation d'énergie au niveau des Commutateurs et des serveurs dans le pire cas c'est-à-dire que nous supposons que les Commutateurs allumés sont en service ie ils consomment une puissance $e_{busyS_e} = e_{busyS_e} = 300$ w et ceux qui sont éteints consomme une puissance $e_{off} = 0$ w.

La figure 2-15 montre la consommation d'énergie au niveau des Commutateurs par rapport aux deux propositions.

Nous pouvons remarquer que plus le nombre de serveurs est important, plus le gain en énergie est important par rapport à la proposition de Dan et al.

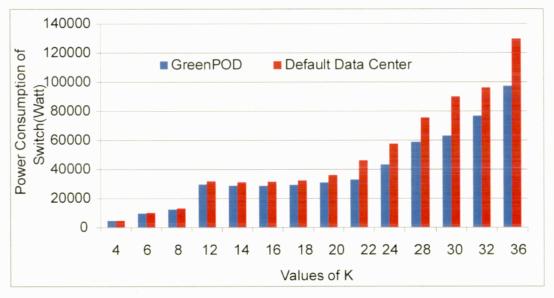


Figure 2-15: Consommation d'énergie des Commutateurs par rapport au deux propositions

La figure 2-16 montre la consommation totale d'énergie (Switch et serveurs) par rapport au deux propositions.

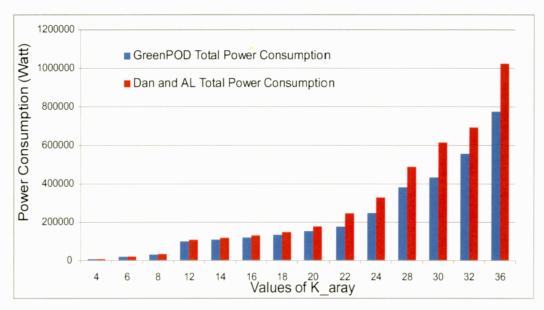


Figure 2-16: Consommation totale d'énergie (Commutateurs+serveurs) par rapport au deux propositions

Enfin la figure 2-17 montre la consommation totale d'énergie de notre proposition par rapport à un centre de données classique.

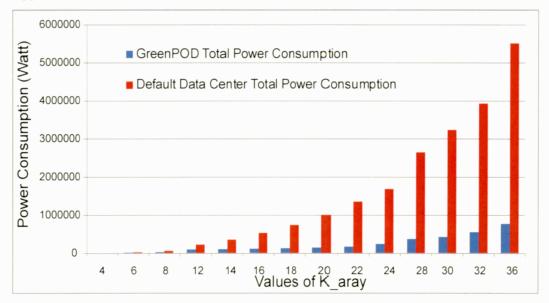


Figure 2-17: Consommation totale d'énergie (Commutateurs+serveurs) de notre proposition par rapport à un centre de données par classique

Les trois courbes suivantes présentent l'énergie consommée en fonction de la charge du système. La figure 2-18 présente la consommation d'énergie des serveurs en fonction de la charge du système. Par exemple, pour une charge de 40% avec un temps de simulation de 2000 milli seconde GreenPOD consomme 472984.5 watt contre 618784.5 watt pour Dan et al.

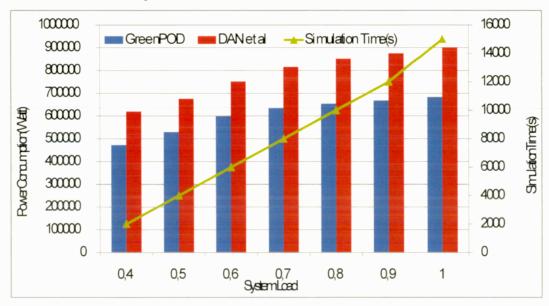
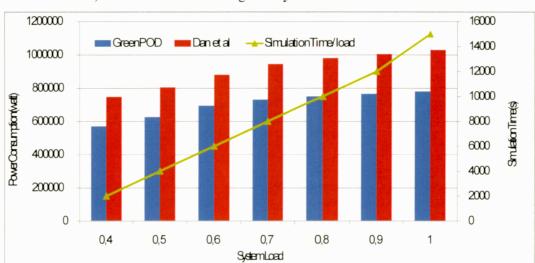


Figure 2-18: Consommation d'énergie des serveurs en fonction de la charge du système



La figure 2-19 présente l'énergie consommée par l'ensemble (serveurs + Commutateurs) en fonction de la charge du système.

Figure 2-19: Energie consommée (Serveurs + Commutateurs) en fonction de la charge

2.5. Conclusion

Dans ce mémoire, nous avons proposé un modèle de fonctionnement avec différents seuils d'activation pour réduire la consommation d'énergie dans un centre de données avec la topologie Fat-tree. Pour cela nous avons ajusté le nombre de serveurs et Commutateurs actifs dans le système tout en éteignant les autres. La décision de savoir si les serveurs doivent être activés est basée sur le nombre de requêtes dans la file. Nous avons comparé notre modèle avec celui de Dan et al et les résultats montrent que nous économisons beaucoup plus d'énergie par rapport à eux avec un temps de réponse acceptable.

Ainsi, nous pouvons garantir les accords de niveau de service à l'utilisateur final tout en économisant environ 74% de l'énergie avec l'adaptation des commutateurs et serveurs.

Dans nos travaux futurs, nous étudierons les centres de données hétérogènes pour rendre notre modèle plus pratique. Nous allons analyser notre modèle dans un

environnement plus complexe avec des classes de clients différentes et des temps de services différents.

CONCLUSION GENERALE

Bilan:

Dans ce mémoire, nos travaux de recherche ont porté sur la réduction de la consommation d'énergie dans les centres de données en utilisant le formalisme de file d'attente M/M/C. Dans un premier temps, nous avons réalisé une présentation générale des architectures de centre de données. De plus, nous avons énuméré des causes de perte d'énergie et des techniques utilisées pour palier à cette perte. Nous avons proposé une contribution dont l'objectif est d'apporter des solutions en utilisant la gestion dynamique des serveurs pour avoir un gain en énergie tout en ayant un temps de réponse acceptable. Nous avons comparé notre solution avec d'autres et nous avons montré que la notre était meilleure. Pour chaque charge du système utilisé (peu sollicitée 45%, moyennement sollicitée 65% ou très sollicités 95%), notre solution reste toujours meilleure en gain d'énergie.

Perspectives:

En se basant sur les travaux et résultats obtenus dans le cadre de ce mémoire, nous envisageons de définir des contraintes plus spécifiques concernant l'hétérogénéité des serveurs. Nous prévoyons aussi des simulations basées sur des SLA différents ou bien des taux d'arrivés différents. Nous pourrons aussi tester le modèle avec une file M/G/1 avec des classes de clients différentes.

BIBLIOGRAPHIE

[Al-Fares et al. 2008] M. Al-Fares, A. Loukissas, and A. Vahdat, "A scalable, commodity data center network architecture," in Proceedings of the ACM SIGCOMM 2008 conference on Data communication. ACM, 2008, pp. 63–74.

[Andrea 2017] Dr. Anders S.G. Andrae (Huawei), "Total Consumer Power Consumption Forecast"at the Nordic Digital Business Summit, Helsinki, Finland, October 5, 2017.

[Andreyev 2014], A. Andreyev, "Introducing data center fabric, the next-generation facebook data center network", Novembre 2014, [Online], https://code.facebook.com/posts/360346274145943/

[Baliga et al. 2011] J. Baliga, R. W. A. Ayre, K. Hinton, and R. S. Tucker, "Green Cloud Computing: Balancing Energy in Processing, Storage and Transport," Proceedings of the IEEE, vol. 99, no. 1, pp. 149–167, Jan. 2011.

[Barroso et U. Hölzle, 2007] L. A. Barroso et UrsHölzle. "The Case for Energy-Proportional Computing". In :*IEEE Computer* 40 (2007). url: http://www.computer.org/portal/site/computer/index.jsp?ageID=computer_level1&path=computer/homepa ge/Dec07&file=feature.xml&xsl=article.xsl.

[Belbekkouche et al. 2012] A. Belbekkouche, M. Hasan, and A. Karmouch, "Resource discovery and allocation in network virtualization," *IEEE Communications Surveys & Tutorials*, no. 99, pp. 1–15, June 2012

[Beloglazov et Buyya, 2010] A. Beloglazov and R. Buyya, "Energy efficient management in virtualized cloud data centers", in Cluster, cloud and Grid Computing (CCGrid), 10th IEEE/ACM International Conference, 'Melbourne, Autrialia), p. 826-831, 2010.

[Beloglazov et al. 2012] A. Beloglazov, J. Abawajy, and R. Buyya, "Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing", Future Generation Computer Systems, vol. 28, no 5, p. 755-768, 2012.

[Beloglav 2013] A. Beloglav, Energy-Efficient Management of Virtual Machines in Data Centers for Cloud Computing. PhD thesis, Departement of Computing and Information Systems, The university of MELBOURNE, April 2013.

[Berl et al. 2010] A. Berl, E. Gelenbe, M. Di Girolamo, G. Giuliani, H. De Meer, M. Q. Dang, and K. Pentikousis, "Energy-Efficient Cloud Computing," *TheComputer Journal*, vol. 53, no. 7, pp. 1045–1051, Sept. 2010.

[Bolla et al. 2011] R. Bolla, R. Bruschi, F. Davoli, and F. Cucchietti, "Energy efficiency in the future Internet: A survey of existing approaches and trends in energy-aware fixed network infrastructures," *Communications Surveys & Tutorials, IEEE*, vol. 13, no. 2, pp. 223–244, 2011.

[Brackbill 2016]N. Brackbill, "The Energy of the Cloud," Physics 240, Stanford University Fall 2016.

[Buyya et Sing 2018] R. Buyya and S. Singh. "Sustainable Cloud Computing: Foundations and Future Directions." Business Technology & Digital Transformation Strategies, Cutter Consortium, Vol. 21, no. 6, Pages 1-9,2018.

[Chase et al. 2001] J. S. Chase, D. C. Anderson, P. N. Thakar, A. M. Vahdat, and R. P. Doyle, "Managing energy and server resources in hosting centers," in *Proceedings of the 18th ACM Symposium on Operating SystemsPrinciples*, Chateau Lake Louise, Banff, Canada, October 2001.

[Chen et al. 2005] Y. Chen, A. Das, W. Qin, A. Sivasubramaniam, Q. Wang, and N. Gautam, "Managing server energy and operational costs in hosting centers," *ACM SIGMETRICS Performance Evaluation Review*, vol. 33, no. 1, pp. 303–314, 2005.

[Chen et al. 2012] K. Chen, A. Singla, A. Singh, K. Ramachandran, L. Xu, Y. Zhang, X. Wen, and Y. Chen, "Osa: An optical switching architecture for data center networks with unprecedented flexibility," in Proceedings of the 9th USENIX conference on Networked Systems Design and Implementation. USENIX Association, 2012, pp. 18–18.

[Chirgwin 2015] <u>Richard Chirgwin</u>: Which data centre network topology's best? Depends on what you want to break: https://www.theregister.co.uk/2015/10/13/which_data_centre_topology_is_best_depends on what you want to break/

[Da COSTA 2015] G. Da COSTA, "Evaluation et optimisation de performance énergétique des centres de calcul", Thèse en vu de l'obtention de l'habilitation de l'université de Toulouse, présentée et soutenue le 12 Novembre 2015.

[Dally et Towles 2004] W. Dally and B. Towles, Principles and practices of interconnection networks. Morgan Kaufmann, 2004.

[Dayarathna et al. 2016]M. Dayarathna, Y. Wen and F. R., "Data Center Energy Consumption Modeling: A Survey," IEEE Communications Surveys and Tutorials, vol. 18, no. 1, pp. 732-794, 2016.

[Djouhra, Thèse de doctorat 2016] D.Djouhra, Optimisation des performances des data centers des cloud sous contrainte d'énergie consommée, thèse pour le diplôme de doctorat en sciences, présentée et soutenue le 21/04/2016.

[Duato et al. 2003] J. Duato, S. Yalamanchili, and L. Ni, Interconnection networks: An engineering approach. Morgan Kaufmann, 2003.

[EPA 2007] Report to congress on server and data center energy effciency, environmental protection agency,2007.www.energystar.gov/ia/partners/prod_development/downloads/EPA_Da tacenter Report Congress Final1.pdf.

[Farrington et al. 2010] N. Farrington, G. Porter, S. Radhakrishnan, H. Bazzaz, V. Subramanya, Y. Fainman, G. Papen, and A. Vahdat, "Helios: a hybrid electrical/optical switch architecture for modular data centers," in ACM SIGCOMM Computer Communication Review, vol. 40, no. 4. ACM, 2010, pp. 339–350.

[Fischer et al. 2013] A. Fischer, J. F. Botero, M. T. Beck, H. D. Meer, and X. Hesselbach, "Virtual network embedding: A survey," *IEEE Communications Surveys and Tutorials*, vol. 15, no. 4, pp. 1888–1906, 2013.

[Gandhi et al. 2010] A. Gandhi, V. Gupta, M. Harchol-Balter, M. Kozuch. Optimality analysis of energy-performance trade-off for server farm management. In: Proceedings of the 28th Performance, 2010.

[Ghamkhari et Mohsenian-Rad 2013] M. Ghamkhari, H. Mohsenian-Rad. Energy and Performance Management of Green Data Centers: A Profit Maximization Approach. IEEE Transactions on Smart Grid, Vol. 4, No. 2, JUNE 2013, pp.1017-1025

[Ghribi 2014] C. Ghribi, "Energy efficient resource allocation in cloud computing environments. Net-working and Internet Architecture. Institut National des Télécommunications, 2014.

[Gomma et al. 2011] E. Goma, M. Canini, A. L. Toledo, N. Laoutaris, D. Kosti'c, P. Rodriguez, R. Stanojevi'c, and P. Y. Valentin, "Insomnia in the access: or how to curb access network related energy consumption," in *Proceedings of the ACM SIGCOMM 2011 conference*, ser. SIGCOMM '11. New York, NY, USA: ACM, Aug. 2011, pp. 338–349.

[Greenberg et al. 2009] A. Greenberg, J. Hamilton, D. A. Maltz, and P. Patel, "The cost of a cloud: research problems in data center networks," IGCOMMComput. Commun. Rev., vol. 39, no. 1, pp. 68–73, 2009.

[Grevet 2009] N. Grevet, "Le Cloud computing: évolution ou révolution? pourquoi, quand, comment et surtout faut-il prendre les risques?", Mémoire de recherche, Aout 2009.

[Grid 2010] International Conference on Grid Computing (Grid 2010) Energy Efficient Grids, Clouds and Clusters Workshop (E2GC2-2010), pp.313-320,Brussels, Belgium, October 2010

[Guenter et al. 2011] B.Guenter, N,Jain, C.Williams, "Managing cost, performance and reliability tradeoffs for energy-aware server provisioning", IEEEInfocom2011, pp.1332-1340.

[Guo et al. 2008] C. Guo, H. Wu, K. Tan, L. Shi, Y. Zhang, and S. Lu, "DCell: A scalable and fault-tolerant network structure for data centers," ACM SIGCOMM Computer Communication Review, vol. 38, no. 4, pp. 75–86, 2008

[Guo et al. 2009] C. Guo, G. Lu, D. Li, H. Wu, X. Zhang, Y. Shi, C. Tian, Y. Zhang, and S. Lu, "Bcube: A high performance, server-centric network architecture for modular data centers," ACM SIGCOMM Computer Communication Review, vol. 39, no. 4, pp. 63–74, 2009.

[Gupta et Singh 2003] M. Gupta and S. Singh, "Greening of the internet," in SIGCOMM '03: Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications. Karlsruhe, Germany: ACM, 2003, pp. 19–26.

[Hamilton et al. 2009] A. Greenberg, J. R. Hamilton, N. Jain, S. Kandula, C. Kim, P. Lahiri, D. A. Maltz, P. Patel, and S. Sengupta, "Vl2: a scalable and flexible data center network," SIGCOMM Comput. Commun. Rev., vol. 39, no. 4, pp. 51–62, 2009.

[Heller et al. 2010] B. Heller, S. Seetharaman, P. Mahadevan, Y. Yiakoumis, P. Sharma, S. Banerjee, and N. McKeown, "Elastic tree: Saving energy in data center networks," in 7th USENIX Symposium on Networked System Design and Implementation (NSDI), San Jose, CA, USA, April 2010, pp. 249–264.

[Kant 2009] K. Kant, "Data center evolution: A tutorial on state of the art, issues, and challenges," Computer Networks, vol. 53, no. 17, pp. 2939–2965, 2009.

[Kliazovich et al. 2010] D. Kliazovich, P. Bounvry, Y. Audzevich, and S. U. Khan, "Greencloud: A packet-level simulator of energy-aware cloud computing data centers," in *IEEE Globecom*, Miami, FL, USA, December 2010.

[Koomey 2011] J. G. Koomey, "Growth in data center electricity use 2005 to 2010," Oakland, CA, Analytics Press, 2011.

[Leiserson 1985] C. E. Leiserson, "Fat-trees: universal networks for hardware-efficientsupercomputing," IEEE Trans. Comput., vol. 34, no. 10, pp. 892–901, 1985.

[Liao et al. 2015] D.Liao, Ke Li, Gang Sun, Vishal Anand, Yu Gong, Zhi Tan « Energy and Performance Management in Large Data Centers A Queuing Theory Perspective » Computing, Networking and Communications (ICNC), 2015 International Conference on (2012) IEEE.

[Liao et Wang 2018] Q. Liao, Z. Wang," Energy Consumption Optimization Scheme of Cloud Data Center Based on SDN", 8th International Congress of Information and Communication Technology (ICICT- 2018), Procedia Computer Science 131 (2018) 1318–1327.

[Mardani et al. 2015] Abbas Mardani, Ahmad Jusoh, Edmundas Kazimieras Zavadskas, Fausto Cavallaro, and Zainab Khalifah. "Sustainable and renewable energy: An Overview of the Application of Multiple Vriteria Decision Making Techniques and Approaches", Sustainability, vol. 7, no. 10, pp.13947-13984, 2015.

[Mazzucco et al. 2010] M.Mazzucco, D.Dyachuk, and R. Deters, « Maximizing cloud providers revenues via energy aware allocation policies », IEE International Conference on cloud Computing, Miami, Florida, July 2010, pp. 131-138.

[Naone 2009] E. Naone. Conjuring clouds. Technology Review, 112(4):54-56, 2009.

[Nedevschi et al. 2008] S. Nedevschi, L. Popa, G. Iannaccone, S. Ratnasamy, and D. Wetherall, "Reducing network energy consumption via sleeping and rate adaptation," in *Proceedings of the 5th USENIX Symposium onNetworked Systems Design and Implementation*, ser. NSDI'08. Berkeley, CA, USA: USENIX Association, 2008, pp. 323–336. [Online]. Available: http://dl.acm.org/citation.cfm?id=1387589.1387612.

[Ni et Bai 2017] J. Ni and X. Bai, "A review of air conditioning energy performance in data centers," Renewable and Sustainable Energy Reviews, vol. 67, pp.625-640, 2017.

[Niranjan et al. 2009] R. Niranjan Mysore, A. Pamboris, N. Farrington, N. Huang, P. Miri, S. Radhakrishnan, V. Subramanya, and A. Vahdat, "Portland: a scalable fault-tolerant layer 2 data center network fabric," ACM SIGCOMM Computer Communication Review, vol. 39, no. 4, pp. 39–50, 2009.

[Nogueira et al. 2011] J. Nogueira, M. Melo, J. Carapinha, and S. Sargento, "Virtual network mapping into heterogeneous substrate networks," *IEEE Symposium on Computers and Communication*, pp. 438–444, June 2011.

[Pries et al. 2011] R. Pries, M. Jarschel, D. Schlosser, M. Klopf, and P. Tran-Gia, "Power Consumption Analysis of Data Center Architectures," in *GreenNets*, Colmar, France, Oct. 2011.

[Schappi et al. 2010] B. Schappi, T. Bogner, B. P. S. Weeren, F. Bellosa, A. Anglade, and B. Harrison, "Guide d'acahat et de gestion des équipements et des infrastructures pour des serveurs sobres en énergie", 2010. Austrian energy Agency, Université de Karlruche, Sun Micro system.

[Schwartz et al. 2012] C. Schwartz, R. Pries, P. Tran-Gia, "A Queuing Analysis of an Energy-Saving Mechanism in Data Centers", ICOIN 2012, pp 70-75, 2012.

[Sekhar et Jeba 2013] J. Sekhar and G. Jeba, "Energy efficient vm live migration in cloud data centers", International Journal of Computer Science and Network (IJCSN), vol. 2, no. 2, p. 71-75, 2013.

[Shehabi et al. 2016] A. Shehabi, S. J. Smith, N. Horner, I. Azevedo, R. Brown, J. Koomey, E. Masanet, D. Sartor, M. Herrlin and W. Lintner, "United States DataCenter Energy Usage Report," Lawrence Berkeley National Laboratory, Berkeley, California, 2016.

[Sing et al. 2017] Sukhpal Singh, Inderveer Chana and Rajkumar Buyya. "STAR: SLA-aware autonomic management of cloud resources." IEEE Transactions on Cloud Computing(2017)

[Stansberry 2014] M. Stansberry. 2014 Data Center Industry Survey. Rapp. tech. Uptime Institute, 2014.

[Su et al. 2012] S. Su, Z. Zhang, X. Cheng, Y. Wang, Y. Luo, and J. Wang, "Energy aware virtual network embedding through consolidation," in *INFOCOMWorkshops*, 2012, pp. 127–132.

[Sulmont et al. 2016], <u>Rémi Sulmont</u>et Loïc Farge, La "pollution numérique" pourrait rapidement dépasser celle de l'aéronautique, mis à jour le 12/02/2016, [En ligne], https://www.rtl.fr/actu/debats-societe/la-pollution-numerique-pourrait-rapidement-depasser-celle-de-l-aeronautique-7781822692.

[Wang et al. 2010] G. Wang, D. Andersen, M. Kaminsky, K. Papagiannaki, T. Ng, M. Kozuch, and M. Ryan, "c-through: Part-time optics in data centers," in ACM

SIGCOMM Computer Communication Review, vol. 40, no. 4. ACM, 2010, pp. 327–338.

[Yang 2013], Liu Yang, Jogesh K. Muppala and MalathiVeeraraghavan, "A Survey of Data Center Network Architectures.", 2013.

[Yu et al. 2008] M. Yu, Y. Yi, J. Rexford, and M. Chiang, "Rethinking virtual network embedding: substrate support for path splitting and migration," *SIGCOMMComput. Commun. Rev.*, vol. 38, no. 2, pp. 17–29, Mar. 2008. consumption," in *Proceedings of theACM SIGCOMM 2011 conference*, ser. SIGCOMM '11. New York, NY, USA: ACM, Aug. 2011, pp. 338–349.

[Zhang et al. 2008] G.-Q. Zhang, Q.-F. Yang, S.-Q. Cheng, and T. Zhou, "Evolution of the Internet and Its Cores," *New Journal of Physics*, vol. 10, Dec. 2008.