



HAL
open science

Pour une meilleure efficacité énergétique dans un système Smart Grid - Cloud

Mohammed Anis Benblidia

► **To cite this version:**

Mohammed Anis Benblidia. Pour une meilleure efficacité énergétique dans un système Smart Grid - Cloud. Réseaux et télécommunications [cs.NI]. Université de Technologie de Troyes, 2021. Français. NNT : 2021TROY0019 . tel-03810656

HAL Id: tel-03810656

<https://theses.hal.science/tel-03810656v1>

Submitted on 11 Oct 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Thèse
de doctorat
de l'UTT

Mohammed Anis BENBLIDIA

**Pour une meilleure efficacité
énergétique dans un système
*Smart Grid - Cloud***

Champ disciplinaire :
Sciences pour l'Ingénieur

2021TROY0019

Année 2021

THESE

pour l'obtention du grade de

DOCTEUR

de l'UNIVERSITE DE TECHNOLOGIE DE TROYES

en SCIENCES POUR L'INGENIEUR

Spécialité : SYSTEMES SOCIOTECHNIQUES

présentée et soutenue par

Mohammed Anis BENBLIDIA

le 25 juin 2021

**Pour une meilleure efficacité énergétique
dans un système *Smart Grid - Cloud***

JURY

M. É. CHÂTELET	PROFESSEUR DES UNIVERSITES	Président
M. Y. GHAMRI-DOUDANE	PROFESSEUR DES UNIVERSITES	Rapporteur
Mme L. MOKDAD	PROFESSEURE DES UNIVERSITES	Rapporteuse
M. S. AKNINE	PROFESSEUR DES UNIVERSITES	Examineur
M. B. BRIK	MAITRE DE CONFERENCES	Examineur
M. M. ESSEGHIR	MAITRE DE CONFERENCES	Directeur de thèse
Mme L. MERGHEM-BOULAHIA	PROFESSEURE DES UNIVERSITES	Directrice de thèse

À

Mes très chers parents Nacera et Mohamed

Mon épouse bien aimée Nourelhouda

Mes frères Zeneddine, Mehdi, Kadi, Faiza et Djawed

Remerciements

Je tiens tout d'abord à remercier Leila Merghem-Boulaïhia et Moez Esseghir, mes directeurs de thèse, pour la confiance qu'ils ont eue en moi en me proposant de travailler sur ce sujet de thèse, ainsi que pour leur patience.

Mes vifs remerciements vont à Lynda Mokdad et Yacine Ghamri-Doudane, d'avoir accepté de rapporter mon travail. Je les remercie pour la lecture approfondie et l'évaluation de cette thèse. Je remercie également Samir Aknine et Eric Chatelet qui ont accepté d'évaluer ce travail en tant qu'examineurs.

Je tiens également à remercier Bouziane Brik, pour ses précieux conseils qui m'ont permis, à chaque fois, de mieux comprendre, appréhender et maîtriser le travail que l'on demande à un doctorant.

Table des matières

1	Introduction	15
1.1	Transition numérique et augmentation du trafic	16
1.2	Numérisation et Cloud Computing	17
1.3	L'énergie et les data centers du Cloud	17
1.4	Le passage aux réseaux électriques intelligents	19
1.5	Les contributions de cette thèse	20
1.6	Plan du manuscrit	21
2	Smart Grid : Concepts fondamentaux et architecture	25
2.1	Introduction	26
2.2	Vers un réseau électrique intelligent	26
2.3	Plus d'interaction pour plus d'efficacité	28
2.4	Architecture de communication du système smart grid	30
2.4.1	Home Area Network (HAN)	30
2.4.2	Neighborhood Area Network (NAN)	30
2.4.3	Wide Area Network (WAN)	31
2.5	Les principales caractéristiques d'un smart grid	32
2.5.1	La communication bidirectionnelle	32
2.5.2	Les Ressources d'Energie Distribuées (DERs)	32
2.5.3	La Gestion de l'équilibre du système électrique à travers les mécanismes de la réponse à la demande	33
2.5.4	La tarification intelligente	33
2.6	Conclusion	34
3	Gestion d'énergie des data centers du Cloud	35
3.1	Introduction	36
3.2	Cloud computing	36
3.2.1	Principes fondamentaux	36
3.2.2	Le rôle des data centers dans le cloud computing	39
3.3	Data centers et énergie	40

3.3.1	Composants d'un data center	40
3.3.2	Principales métriques d'efficacité énergétique des data centers	42
3.4	Interaction des data centers du cloud avec le smart grid	43
3.5	Conclusion	45
4	Gestion d'énergie des data centers : Solutions existantes et besoins	47
4.1	Introduction	48
4.2	Techniques d'efficacité énergétique des équipements non-IT d'un data center	48
4.3	Techniques d'efficacité énergétique des serveurs d'un data center	48
4.4	Etat de l'art sur la gestion d'énergie des data centers dans un système Smart Grid – Cloud	49
4.4.1	Architecture Smart Grid – Cloud	50
4.4.2	Architecture Smart Microgrid - Cloud	50
4.5	Les outils d'aide à la décision dans un système Smart Grid-Cloud	51
4.5.1	Système multi-agents et la théorie des jeux	52
4.5.2	Optimisation convexe	53
4.5.3	Processus de décision Markovien	53
4.5.4	La logique floue	54
4.6	Discussion et conclusion	54
5	Un jeu non-coopératif pour une répartition équitable de l'énergie entre les data centers	57
5.1	Introduction et motivations	58
5.2	Architecture et modélisation du problème	58
5.2.1	Architecture du système	58
5.2.2	Modélisation de notre jeu non-coopératif	60
5.2.3	Solution optimale du jeu	62
5.3	Résultats des simulations	62
5.3.1	Paramètres des simulations	63
5.3.2	Évaluation des performances de notre approche	63
5.4	Conclusion	66
6	Répartition efficace de l'énergie entre les data centers alimentés par des énergies renouvelables	67
6.1	Introduction et motivations	68
6.2	Architecture et modélisation du problème	69
6.2.1	Architecture du système	69
6.2.2	Description du jeu non-coopératif	71
6.2.3	Notre solution optimale	75
6.3	Simulations et résultats	76
6.3.1	Paramètres des simulations	76
6.3.2	La répartition d'énergie entre les data centers	77

6.3.3	Les performances de notre approche en termes de taux de réduction de charge et d'émissions de dioxyde de carbone	80
6.4	Conclusion	80
7	Minimiser le coût énergétique des data centers en utilisant les Microgrids	83
7.1	Introduction et motivations	84
7.2	Modèle et architecture du système	84
7.2.1	Architecture du système	84
7.2.2	Niveau 1 : Description du jeu pour la répartition de l'électricité entre les data centers	87
7.2.3	Niveau 2 : Minimiser le coût énergétique des microgrids	90
7.3	Résultats des simulations	93
7.3.1	Paramètres des simulations	93
7.3.2	Répartition de l'électricité entre les data centers	94
7.3.3	Minimiser le coût d'énergie des microgrids	96
7.3.4	Les performances de notre approche en termes de taux de réduction de charge et des émissions de Co2	98
7.4	Conclusion	99
8	Classer les nœuds Fog pour la répartition des requêtes des utilisateurs : une approche basée sur la logique floue	101
8.1	Introduction et motivations	102
8.2	Une approche de classification basée sur des quantificateurs flous	104
8.2.1	Quantificateurs linguistiques	104
8.2.2	Propositions quantifiées floues	104
8.2.3	Notre approche de classification	104
8.3	Classer les nœuds Fog pour la répartition des requêtes des utilisateurs	105
8.3.1	Architecture et modélisation du problème	106
8.3.2	Application de la méthode de classification pour la répartition efficace des requêtes	107
8.4	Simulations et résultats	107
8.4.1	Paramètres des simulations	107
8.4.2	Métriques d'évaluation de notre méthode	110
8.4.3	Niveau de satisfaction des utilisateurs	110
8.4.4	Performances de notre approche en termes de délais et de consommation énergétique	111
8.5	Conclusion	112
9	Conclusion générale et perspectives	113
A	Démonstration du jeu non-coopératif	117
A.1	L'existence de l'équilibre de Nash	117
A.2	Unicité de l'équilibre de Nash	118

Table des figures

1.1	Croissance des utilisateurs Internet 2018-2023 [57]	16
2.1	Modèle conceptuel du smart grid vu par NIST [54]	28
2.2	Architecture réseau du Smart Grid [16]	31
3.1	Principaux composants d'un data center [90]	41
3.2	Vue globale d'un data center [90]	42
3.3	Interaction des data centers avec le smart grid et ses utilisateurs	44
5.1	Architecture d'un système smart grid - cloud	59
5.2	Variation des coûts des data centers	63
5.3	La quantité d'électricité allouée en utilisant notre approche basée sur un jeu non-coopératif	64
5.4	Comparaison des performances entre notre approche et l'approche traditionnelle : (A) Allocation d'énergie, (B) Allocation d'énergie Vs le coût de charge d'un data center , (C) Allocation d'énergie VS le coût des applications critiques, (D) Allocation d'énergie Vs le coût des applications critiques totale	65
6.1	Les 10 plus gros acheteurs d'énergie des États-Unis en 2019 (GW) [100]	68
6.2	L'énergie renouvelable dans une architecture smart grid - cloud	70
6.3	Production d'énergie renouvelable des data center	77
6.4	Variation des coûts des data centers	78
6.5	Performances de notre jeu non-coopératif par rapport à <i>REB</i>	79
6.6	Évaluation de notre proposition en termes de taux de charge électrique et des émissions de dioxyde de carbone	81
7.1	Architecture smart microGrid- cloud	87
7.2	Variations des coût des data centers du microgrid 01	95
7.3	Performances de notre approche de répartition d'énergie par rapport à l'approche traditionnelle et à celle basée sur le prix.	95
7.4	Variations de l'énergie achetée du smart grid et des microgrids	97
7.5	Comparer le coût énergétique des microgrids en utilisant notre proposition et BGS.	98

7.6	Performances de notre approche à deux niveaux en termes des émissions de dioxyde de carbone et de taux de charge électrique	99
8.1	Architecture fog-cloud (NIST 2018)	102
8.2	Architecture de notre système	106
8.3	Exemples de fonctions d'appartenance floue	108
8.4	Satisfaction des utilisateurs en utilisant notre approche par rapport aux approches basé sur la distance, le prix et à la latence dans chaque ensemble de nœuds fog et pour chaque classe d'utilisateurs	111
8.5	Temps d'exécution de notre approche par rapport aux autres approche tout en variant la classe d'utilisateurs et l'ensemble des nœuds fog	111
8.6	Performances de notre approche en termes de consommation d'énergie par rapport aux autres approches	112

Liste des tableaux

2.1	Comparaison entre le réseaux électrique traditionnel et le smart grid	27
4.1	Tableau comparatif des travaux qui se portent sur l'efficacité énergétique des data centers dans un système smart grid	55
5.1	L'efficacité de notre approche de répartition d'énergie basée sur un jeu non-coopératif	66
6.1	Nomenclature	72
6.2	Paramètres des simulations	77
6.3	Taux des émissions de dioxyde de carbone par l'énergie fossile et éolienne	80
7.1	Nomenclature	85
7.2	Paramètres des simulations	94
8.1	Paramètres des simulations	108
8.2	Les ensembles fog et les valeurs de leurs critères	109
8.3	Fonctions d'appartenances des classes des utilisateurs	109

Résumé

Dans cette thèse, nous étudions l'efficacité énergétique des infrastructures informatiques dans un système smart grid – cloud. Nous nous intéressons plus particulièrement aux réseaux de communication et aux data centers du cloud. Nous nous focalisons sur ces derniers à cause de leur grande consommation d'énergie et du rôle vital qu'ils jouent dans un monde connecté en pleine expansion, les positionnant, ainsi, comme des éléments importants dans un système smart grid - cloud. De ce fait, les travaux de cette thèse s'inscrivent dans le cadre d'un seul framework intégrant le smart grid, le microgrid, le cloud, les data centers et les utilisateurs. Nous avons, en effet, étudié l'interaction entre les data centers du cloud et le fournisseur d'énergie du smart grid et nous avons proposé des solutions d'allocation d'énergie et de minimisation du coût d'énergie en utilisant deux architectures : (1) une architecture smart grid-cloud et (2) une architecture microgrid-cloud. Par ailleurs, nous avons porté une attention particulière à l'exécution des requêtes des utilisateurs tout en leur garantissant un niveau de qualité de service satisfaisant dans une architecture fog - cloud. En comparaison avec les travaux de l'état de l'art, les résultats de nos contributions ont montré qu'ils répondent aux enjeux identifiés, notamment en réduisant les émissions de gaz à effet de serre et le coût d'énergie des data centers.

Mots-clés : Réseaux électriques intelligents ; Data centers ; Énergies renouvelables ; Cloud computing ; Fog computing ; Théorie des jeux.

Abstract

This thesis considers the energy efficiency of information and communication infrastructures in a smart grid - cloud system. It especially deals with communication networks and cloud data centers due to their high energy consumption, which confers them an important role in the network. The contributions of this thesis are implemented on the same framework integrating the smart grid, microgrid, cloud, data centers and users. Indeed, we have studied the interaction between the cloud data centers and the smart grid provider and we have proposed energy efficient power allocation solutions and an energy cost minimization scheme using two architectures: (1) a smart grid-cloud architecture and (2) a microgrid-cloud architecture. In addition, we paid close attention to execute user requests while ensuring a good quality of service in a fog-cloud architecture. In comparison with related works, the results of our contributions have shown that they respond to the identified challenges, particularly in terms of reducing carbon emissions and energy costs of cloud data centers.

Keywords: Smart grids; Data centers; Renewable energy sources; Cloud computing; Fog computing; Game Theory.

Chapitre **1**

Introduction

Sommaire

1.1	Transition numérique et augmentation du trafic	16
1.2	Numérisation et Cloud Computing	17
1.3	L'énergie et les data centers du Cloud	17
1.4	Le passage aux réseaux électriques intelligents	19
1.5	Les contributions de cette thèse	20
1.6	Plan du manuscrit	21

1.1 Transition numérique et augmentation du trafic

Au fil de ces dernières années, l'usage de la technologie numérique est devenu inévitable et passe d'un générateur de productivité marginal à un catalyseur fondamental d'innovation dans tous les secteurs de l'industrie moderne. D'ailleurs, la plupart des industries sont conscientes de cette évolution et commencent à aligner leurs processus commerciaux et expériences client existants aux exigences des entreprises et marchés actuels. Nous pouvons constater cette évolution dans les différents axes de la transformation numérique. En effet, les réseaux numériques connaissent un changement marquant avec le développement des technologies réseaux *NFV/SDN*, la *5G* et l'utilisation du cloud computing. La transformation de cet axe important permet d'avoir des échanges intelligentes et plus automatisées. De ce fait, plusieurs services numériques voient le jour comme : l'industrie 4.0, les maisons/villes intelligentes, les véhicules autonomes et la réalité virtuelle.

Ces nouveaux services sont basés sur des architectures offrant les ressources de calcul, stockage et réseaux nécessaires à la gestion du nombre colossal de données générées par les utilisateurs. En effet, le développement des services numériques engendre en parallèle une augmentation du nombre d'utilisateurs connectés à Internet. D'après le rapport annuel d'Internet 2018 – 2023 de Cisco [57], le nombre total d'utilisateurs d'Internet devrait passer de 3,9 milliards en 2018 à 5,3 milliards en 2023 avec un Taux de Croissance Annuel Composé (Compound Annual Growth Rate–CAGR) de 6% (figure 1.1). En termes de population, cela représente 51% de la population mondiale en 2018 et 66% de la population mondiale d'ici 2023.

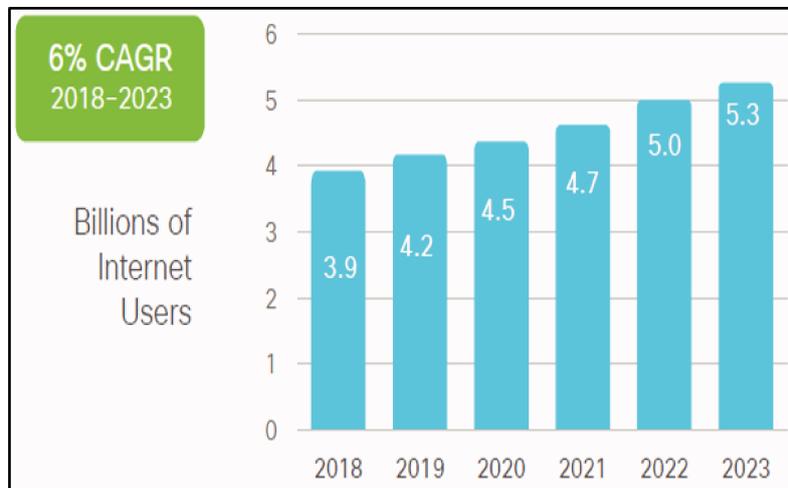


FIGURE 1.1 – Croissance des utilisateurs Internet 2018-2023 [57]

La croissance des utilisateurs Internet est accompagnée par l'apparition de nouvelles applications telles que les compteurs intelligents, la vidéosurveillance, la surveillance des soins de santé, le transport et le suivi des colis. Cela contribue de manière importante à la croissance des appareils et des connexions avec un nombre de 50 milliards d'objets connectés estimé en 2022 [2]. D'ici là, les connexions M2M représenteront la moitié du nombre total d'appareils et de connexions [57].

Ces appareils intelligents conçus avec différents formats sont généralement dotés de faibles capacités de calcul et de stockage. Ces caractéristiques poussent les industries à adopter de nouvelles architectures réseaux qui leur permettent une scalabilité plus flexible et économiser les efforts de configuration de leur infrastructure. Ainsi, ces appareils se voient généralement transférer leur trafic au cloud computing pour assurer un bon service aux utilisateurs.

1.2 Numérisation et Cloud Computing

Le processus de numérisation des entreprises doit, d'une part, s'accommoder à de nouveaux modèles commerciaux, et d'autre part s'adapter au volume colossal de données générées par leurs départements [55]. Par conséquent, une infrastructure flexible, automatisée et caractérisée par une adoption rapide est nécessaire pour mieux maîtriser le grand nombre de données générées des entreprises. Le cloud computing est une solution qui permet de répondre à ces besoins.

Le cloud computing a gagné en popularité durant ces dernières années. Il constitue un mélange de nombreux domaines informatiques ; Il fournit de l'infrastructure, des services et des applications sur Internet. De plus, il facilite la réduction du coût d'investissement, dissocie les services de la technologie sous-jacente et offre une flexibilité en termes de dimensionnement des ressources [87]. Le principal objectif du cloud computing est de permettre aux entreprises l'accès à des technologies via internet de manière rentable. En utilisant le cloud computing, les petites entreprises peuvent étendre leurs ressources informatiques en fonction des demandes de services et profiter des mêmes opportunités de croissance pour concurrencer les autres entreprises du marché [3].

De plus en plus d'entreprises migrent leur infrastructure vers les data centers du cloud. En effet, selon Gartner Inc, le marché mondial des services de cloud public devrait croître de 55,5% en 2022 pour atteindre 331,2 milliards de dollars, contre 214,3 milliards de dollars en 2019. De plus, le segment de marché qui connaît la croissance la plus rapide sera celui des services d'infrastructure de système cloud, ou Infrastructure en tant que service (IaaS), qui devrait croître de 97% en 2022 pour atteindre 76,6 milliards de dollars, contre 38,9 milliards de dollars en 2019 [61].

Par ailleurs, la transformation numérique n'est pas le seul facteur du développement du cloud. L'émergence de l'Internet des Objets contribue de manière importante à l'utilisation de cette technologie. En effet, les contraintes énergétiques, de calcul et de stockage des objets connectés les poussent à migrer leurs tâches aux data centers du cloud. En outre, en termes de contrainte réseau, Huawei montre que les futurs réseaux 5G auront 100 fois plus de bande passante que les réseaux 4G [29]. Par conséquent, le rôle majeur des data centers du cloud dans le stockage, le calcul et la gestion des données augmente de façon exponentielle au fil du temps.

1.3 L'énergie et les data centers du Cloud

La croissance de l'utilisation du cloud computing est principalement soutenue par des millions de serveurs et de périphériques réseaux disséminés dans des milliers de data centers géographiquement distribués. Les data centers sont l'infrastructure clé du cloud computing pour fournir les ressources nécessaires au traitement et au stockage des données. Parallèlement aux exigences croissantes des

utilisateurs finaux et à leur volonté d'utiliser des services basés sur le cloud, les applications actuelles sont devenues gourmandes en données, ce qui pose des exigences plus critiques pour la capacité de stockage des data centers [86]. Un nombre croissant de data centers ont été bâtis dans le monde pour stocker la quantité massive de données. Malgré les énormes progrès réalisés dans la conception de serveurs et d'équipements réseaux écoénergétiques, les coûts énergétiques encourus par les fournisseurs de services cloud (CSP) ont considérablement augmenté avec la croissance de la demande pour ces services. Il est estimé aujourd'hui que le secteur des data centers représente 1.4% de la consommation mondiale d'électricité [72] et que ce pourcentage a tendance à s'élever à 8% dans les cinq prochaines années [102]. De plus, la consommation massive d'énergie des data centers est responsable de 2% des émissions totales de gaz à effet de serre [103].

Les principaux fournisseurs de cloud computing tels que Microsoft, Google et Amazon ont construit les plus grands data centers au monde avec une énorme consommation d'énergie. Un data center typique comprend des centaines de milliers de serveurs informatiques, d'équipements de refroidissement et de transformateurs électriques. Le coût énergétique d'un data center provient principalement de l'énergie consommée par les serveurs physiques qui représente 45% de la consommation totale d'énergie [30]. Par exemple, le data center de Microsoft à Quincy, Washington, consomme 48MW, ce qui est suffisant pour alimenter 40000 foyers [112]. On estime que le coût total de l'électricité des serveurs et des data centers aux États-Unis est de 7,4 milliards de dollars par an [112] et domine tous les autres coûts reliés au cloud computing. Donc, l'importante consommation d'énergie des data centers présente deux défis majeurs. Tout d'abord, cette quantité de consommation d'énergie entraîne des coûts considérables pour les fournisseurs de cloud. Le coût de l'énergie étant une part importante du coût d'exploitation de chaque fournisseur de cloud, il est nécessaire et important de minimiser sa facture d'électricité. Deuxièmement, les data centers en tant que clients d'électricité ont un impact majeur sur le réseau électrique car ils augmentent considérablement la charge du réseau sur les sites où ils sont bâtis [25]. Plusieurs contributions ont été faites pour minimiser la consommation d'énergie des data centers en général [117] [13] [30]. Ces derniers proposent des solutions locales visant l'infrastructure des data centers notamment l'efficacité énergétique des serveurs et le système de refroidissement. Mais, les data centers du cloud ont les principales caractéristiques d'être distribués et flexibles. En fait, un fournisseur cloud peut déplacer la charge des data centers des heures de pointe aux heures creuses et migrer leur charge de travail d'un emplacement physique à un autre. Il peut aussi utiliser les générateurs sur site et les unités de stockage d'énergie pour gérer au mieux la consommation énergétique de ses data centers. De plus de l'effort mis pour minimiser la consommation d'énergie de l'infrastructure des data centers, ces caractéristiques du cloud peuvent être exploités par le fournisseur d'énergie pour réduire la consommation des data centers. Chose qui ne peut se faire avec les réseaux électriques traditionnels. Pour tirer profit de ces caractéristiques, le fournisseur d'énergie doit avoir une visibilité sur le comportement énergétique des data centers en temps réel ce qui lui permettrait d'offrir des services adaptés et précis [80].

1.4 Le passage aux réseaux électriques intelligents

Les réseaux électriques traditionnels ont été conçus et construits pour répondre aux exigences mises en place au siècle dernier. Cependant, plusieurs facteurs actuels les rendent moins efficaces et susceptibles d'être rénovés. En effet, une grande partie de l'infrastructure existante du système électrique remonte aux années 1950 ou même plus tôt et atteint la fin de sa durée de vie utile [52]. De plus, la demande d'électricité à travers le monde ne cesse d'augmenter, entraînant une charge élevée sur les équipements du système électrique. Selon une étude publiée par Bloomberg New Energy Finance (BNEF), la demande mondiale d'électricité va augmenter de 57% à l'horizon 2050 [59]. Compte tenu de la prédominance constante des énergies fossiles dans le paysage énergétique mondial, les émissions annuelles de CO₂ liées à l'énergie pourraient augmenter de 36 milliards de tonnes (Gt) en 2020 à 43 Gt en 2040, soit une hausse de 34% [58]. A ce propos, les services publics font pression sur les politiciens pour qu'ils réduisent leurs émissions de CO₂ en adoptant des sources d'énergie alternatives et mettent en place des réglementations pour accroître l'efficacité énergétique [58].

Pour faire face à ce nouveau paradigme, les réseaux électriques intelligents (smar grids) ont émergé du grand apport offert par les technologies de l'information et de la communication. La fusion de ces deux secteurs offre aux acteurs du système électrique des services supplémentaires en termes de réactivité, flexibilité et sécurité. Grâce à l'augmentation des données disponibles sur l'ensemble de la chaîne de valeur électrique, des outils productifs jusqu'à la consommation des usagers en passant par la conduite des réseaux de transmission et de distribution, les réseaux électriques intelligents ont comme objectif d'optimiser la gestion de l'équilibre offre/demande, impactant à la fois les flux d'énergie et financiers sur les marchés de l'électricité [42]. Aujourd'hui en France, des technologies modernes de production, de stockage sont étudiées, intégrées au smart grid pour fournir au consommateur de meilleures informations, favorisant sa participation dans l'optimisation du réseau électrique.

Grâce à l'infrastructure des réseaux de communications déployée dans les réseaux électriques intelligents, le fournisseur d'énergie peut avoir davantage d'informations sur le comportement énergétique de ses clients. Le smart grid est donc caractérisé d'une communication bidirectionnelle entre le client et son fournisseur, ce qui permet d'harmoniser la gestion d'énergie. Contrairement au système énergétique actuel qui a une architecture centralisée et permet aux flux d'énergie d'être unidirectionnels, du producteur vers les consommateurs, le smart grid permet d'intégrer une multitude de sources d'énergies renouvelables, de toutes tailles, et de permettre de nouveaux modes d'utilisation et de stockage de l'électricité, ainsi permettant l'échange bidirectionnelle de l'énergie entre le producteur et le consommateur [62].

Dans ce contexte, les réseaux électriques intelligents sont un atout important pour la gestion d'énergie des data centers du cloud. La flexibilité de ces derniers permet aux réseaux électriques intelligents d'apporter un plus à leur gestion d'énergie à travers des programmes de réponse à la demande. A cet égard, les fournisseurs de cloud computing ont récemment commencé à étudier de nouvelles approches pour développer des mécanismes de gestion de la charge de travail des data centers

afin d'augmenter leurs contributions aux programmes de réponse à la demande et de minimiser leurs coûts [71]. Cela permet au fournisseur de cloud de réduire ces factures d'électricité en utilisant des ressources plus économes en énergie, en accédant à des stratégies de gestion de l'énergie plus flexibles et en utilisant la variation des prix de l'électricité sur les sites des data centers [46].

1.5 Les contributions de cette thèse

Dans de cette thèse, nous étudions l'efficacité énergétique des infrastructures informatiques dans un système smart grid – cloud. Nous nous intéressons plus particulièrement aux réseaux de communication et aux data centers du cloud. Nous nous sommes focalisés sur ces derniers à cause de leur grande consommation d'énergie et du rôle vital qu'ils jouent dans un monde connecté en pleine expansion, les positionnant, ainsi, comme des éléments importants dans un système smart grid - cloud. De ce fait, nous avons étudié l'interaction des data centers du cloud et le fournisseur d'énergie du smart grid et nous avons proposé des solutions énergétiquement efficaces en utilisant deux architectures : une architecture smart grid-cloud et une architecture microgrid - cloud.

Notre première contribution vise à surmonter le problème de la surcharge énergétique que les data centers du cloud causent au smart grid. Ainsi, nous modélisons le comportement énergétique des data centers en tant que jeu non coopératif dans lequel chaque data center est représenté par un joueur. L'idée de base de notre proposition est d'attribuer la quantité optimale d'électricité à chaque data center afin d'obtenir une répartition équitable entre les data centers. Cette attribution prend en compte trois métriques : la charge des data centers, le nombre d'applications critiques en cours d'exécution au sein d'un data center et le nombre d'applications critiques en cours d'exécution au sein de tous les data centers. Cette contribution a fait l'objet d'un deuxième travail dans lequel nous avons considéré les énergies renouvelables dans une architecture smart grid - cloud et nous avons étudié le problème de répartition de l'énergie entre les data centers. Nous avons supposé que les data centers ont un caractère non coopératif pour demander de l'énergie aux centrales électriques et nous avons modélisé notre problème d'allocation d'énergie comme un jeu non coopératif. Nous avons modélisé la fonction de gain de notre jeu comme un problème d'optimisation non linéaire avant de le résoudre en utilisant les multiplicateurs de Lagrange et les conditions de KarushKuhn-Tucker (KKT). Ainsi, nous déterminons la quantité optimale attribuée à chaque data center en fonction de trois critères principaux : la consommation d'énergie renouvelable, le nombre d'applications critiques en cours d'exécution et la charge de travail. Nous avons effectué des simulations en comparant notre approche avec une approche existante d'optimisation d'énergie. Les résultats montrent que notre approche surpasse l'approche de comparaison avec un pourcentage de 31.2% en termes de taux de charge d'électricité et réduit considérablement l'émission de CO₂.

Nous avons étudié d'autres architectures notamment les architectures microgrid-cloud pour répartir l'énergie du smart grid vers les data centers du cloud. Dans un premier temps, nous modélisons la demande d'énergie entre les data centers et les microgrids comme un jeu non coopératif, en raison de leur comportement de demande d'énergie non coopératif. Ensuite, nous essayons d'allouer la quantité

optimale d'énergie à chaque data center en fonction de sa consommation d'énergie renouvelable, de l'utilisation de la bande passante du réseau et de la consommation d'énergie de ses équipements réseau. Deuxièmement, nous modélisons la fonction de gain du jeu comme un problème d'optimisation non linéaire et le résolvons en utilisant les multiplicateurs de Lagrange et les conditions de KarushKuhn-Tucker (KKT). Enfin, nous comparons les performances de notre approche avec deux algorithmes d'allocation d'énergie et les résultats ont montré que notre jeu est plus efficace et réduit la charge électrique du réseau du smart grid avec un taux de 40.5%.

Pour mettre plus en avant le rôle des microgrids dans une architecture Smart Microgrid-Cloud. Nous nous sommes penchés sur la problématique du coût énergétique élevé des data centers, et par conséquent des microgrids qui les alimentent. Contrairement à la précédente contribution, nous avons proposé dans cette contribution une architecture microgrid-cloud, où les contrôleurs des microgrids ont en plus d'une communication bidirectionnelle verticale avec le contrôleur du Smart Grid une communication bidirectionnelle horizontale entre eux-mêmes. Cela permettra d'acheter et de vendre de l'énergie entre les microgrids qui alimentent d'autres data centers. Notre objectif est de minimiser le coût total énergétique de notre système. Pour ce faire, nous nous basons en premier lieu sur les résultats de la répartition d'énergie via le jeu réalisé entre des data centers. Ensuite, nous utilisons ces résultats pour minimiser la quantité achetée du smart grid en prenant en considération les quantités achetées (et vendues) des (aux) autres microgrids.

Dans l'optique de combiner toutes les contributions sous un seul framework intégrant le microgrid, smart grid, cloud, data center et utilisateurs, nous nous sommes intéressés à l'architecture réseaux et son interaction avec le smart grid. Nous avons porté une attention particulière à l'exécution des requêtes des utilisateurs tout en leur garantissant un niveau de qualité de service satisfaisant. Pour cela nous avons considéré les requêtes reçus par les utilisateurs dans une architecture fog-cloud. Nous avons considéré la nécessité d'une méthode efficace de répartition des tâches dans une architecture fog-cloud. Notre proposition vise à envoyer les requêtes des utilisateurs au meilleur nœud fog, tout en satisfaisant leurs préférences et les contraintes des nœuds fog. Compte tenu du besoin de passage à l'échelle et du caractère dynamique des prix d'un système fog computing, les nœuds fog et les utilisateurs ont des exigences diverses. Pour résoudre ce problème, nous utilisons la logique floue pour sa capacité à couvrir un large éventail de conditions de fonctionnement et à exprimer les préférences des utilisateurs sous forme de quantificateurs flous. Dans ce travail, nous avons utilisé des quantificateurs linguistiques et des propositions quantifiées floues regroupant efficacement les préférences des utilisateurs et les contraintes des nœuds de fog afin de classer les nœuds fog du plus satisfaisant au moins satisfaisant.

1.6 Plan du manuscrit

Ce manuscrit est divisé en 8 chapitres, il apporte le positionnement des recherches menées par rapport à l'état de l'art et détaille les résultats obtenus durant cette thèse. Il est organisé comme suit :

Le chapitre 2 résume les principes fondamentaux et l'architecture du Smart Grid. Nous présentons

tout d'abord les limites des réseaux électriques traditionnels. Ensuite, nous détaillons l'architecture de communication du smart grid et expliquons ses principales caractéristiques.

Le chapitre 3 explique la gestion d'énergie des data centers du cloud. Nous donnons dans un premier temps les principes fondamentaux du cloud computing et mettons l'accent sur le rôle des data centers dans cette technologie. Ensuite, nous présentons les principaux composants des data centers et leurs comportements énergétiques. Enfin, ce chapitre explique l'interaction des data centers du cloud avec le smart grid.

Le chapitre 4 fait un tour d'horizon des besoins et des solutions existantes dans la gestion d'énergie des data centers. Nous présentons les différents travaux qui ont été menés dans ce contexte en termes d'architectures, algorithmes et outils d'aide à la décision. Nous concluons ce chapitre par une discussion sur les différents travaux recensés et notre positionnement par rapport à ces travaux.

Le chapitre 5 présente notre première contribution qui vise à répartir équitablement l'électricité entre les data centers dans un système smart grid - cloud. Nous détaillons l'architecture utilisée, la modélisation de notre approche de répartition d'énergie et nous discutons les résultats des simulations obtenus.

Le chapitre 6 présente une extension du travail décrit dans le chapitre précédent. Nous considérons dans ce chapitre les énergies renouvelables pour répartir équitablement l'électricité entre les data centers du cloud dans un système smart grid - cloud. La répartition se fait en fonction de trois critères : la consommation d'énergie renouvelable, le nombre d'applications critiques en cours d'exécution et la charge de travail des data centers. Nous comparons notre proposition avec un travail de la littérature et montrons les performances de notre approche en termes d'équité, de réduction de taux de charge d'électricité et de taux d'émission de CO₂.

Le chapitre 7 présente notre approche de minimisation à deux niveaux du coût d'énergie des microgrids dans une architecture microgrid - cloud. Nous présentons, dans ce chapitre, notre nouvelle architecture et détaillons la modélisation de notre problème d'optimisation du coût d'électricité. Enfin, nous montrons l'efficacité et les performances de notre approche à deux niveaux en la comparant avec l'approche présentée dans le chapitre précédent.

Le chapitre 8 traite le problème de répartition des tâches des utilisateurs dans les data centers (noeuds fog). Nous expliquons, dans ce chapitre, notre architecture fog - cloud et présentons les éléments de base de la théorie des ensembles flous et de la quantification floue. Nous détaillons aussi, dans ce chapitre, notre approche basée sur la décomposition qui sélectionne le noeud fog qui satisfait le mieux les préférences des utilisateurs et les contraintes des noeuds fog. Enfin, nous étudions les résultats des simulations et les performances de l'approche proposée en termes de satisfaction de l'utilisateur, le délai et la consommation d'énergie.

À la fin de ce manuscrit, nous tirons nos conclusions et nous présentons des perspectives que ces travaux laissent entrevoir.

Smart Grid : Concepts fondamentaux et architecture

Sommaire

2.1	Introduction	26
2.2	Vers un réseau électrique intelligent	26
2.3	Plus d'interaction pour plus d'efficacité	28
2.4	Architecture de communication du système smart grid	30
2.4.1	Home Area Network (HAN)	30
2.4.2	Neighborhood Area Network (NAN)	30
2.4.3	Wide Area Network (WAN)	31
2.5	Les principales caractéristiques d'un smart grid	32
2.5.1	La communication bidirectionnelle	32
2.5.2	Les Ressources d'Energie Distribuées (DERs)	32
2.5.3	La Gestion de l'équilibre du système électrique à travers les mécanismes de la réponse à la demande	33
2.5.4	La tarification intelligente	33
2.6	Conclusion	34

2.1 Introduction

Presque inchangée depuis près d'un siècle, l'infrastructure du réseau électrique telle que nous la connaissons aujourd'hui a réussi à répondre à nos besoins avec succès. Cependant, cette infrastructure vieillissant avec le temps, devient de moins en moins efficace, se heurte sans cesse à ses limites et s'efforce en permanence à faire face à nos exigences. De plus, ce réseau électrique produit l'électricité d'une manière centralisée et l'achemine à un grand nombre d'utilisateurs où la communication est unidirectionnelle : des producteurs vers les consommateurs. Les besoins en matière de fiabilité, de facilité de gestion de l'énergie et de production d'énergie renouvelable soulignent la nécessité d'un réseau modernisé et intelligent pour demain. Dans ce contexte, le réseau électrique intelligent ou Smart Grid ouvre la voie au futur réseau électrique qui vise à créer un système propre, sûr, sécurisé et fiable [77].

Afin de gérer le réseau électrique, il est nécessaire de procéder à des changements importants sans en perturber le fonctionnement. Cela se traduit, dans le cadre du réseau électrique, par la mise en place d'un réseau de capteurs et compteurs intelligents pour acheminer les données de production et de consommation vers le gestionnaire du Smart Grid. Ce dernier devient alors communicant et interactif.

Dans ce chapitre, nous faisons une petite introduction aux réseaux électriques intelligents et présentons leurs principales fonctionnalités et leurs standards. Nous détaillons aussi leur architecture de communication et résumons leurs principales caractéristiques.

2.2 Vers un réseau électrique intelligent

Le réseau électrique traditionnel est composé de quatre infrastructures principales, qui sont les couches de production, de transmission, de distribution et de consommation. Cette infrastructure a survécu par de grandes centrales de production qui sont installées à des niveaux de puissance de plusieurs MW, et elle est connectée à une sous-station de transmission où le système de transmission est mis en route [67]. La ligne de transmission est chargée de fournir l'énergie nécessaire aux sous-stations de distribution afin d'alimenter plusieurs réseaux de distribution avec l'énergie produite. Par conséquent, la connexion entre les couches de production et de consommation est gérée par les couches intermédiaires. Le réseau électrique traditionnel est unidirectionnel en termes de flux d'énergie puisque les sources de production distribuées n'ont pas été autorisées à participer à ce système. De plus, le prix de l'électricité est déterminé par les gestionnaires de réseaux de services publics, où les clients n'ont pas la possibilité de choisir leur tarif d'électricité [18].

L'architecture du réseau électrique traditionnel est principalement basée sur une structure verticale pour décrire un flux d'énergie unidirectionnel. Cependant, nous définissons le réseau électrique traditionnel comme étant horizontal, car le réseau intelligent est vertical en raison de son architecture multicouches composée d'une couche de technologies de l'information et de la communication (TIC) et d'une couche de contrôle et de gestion [67]. Les principales caractéristiques du réseau sont comparées pour les réseaux traditionnels et les réseaux intelligents dans le tableau 2.1 [41].

TABLE 2.1 – Comparaison entre le réseaux électrique traditionnel et le smart grid

Caractéristiques	Traditionnel	Smart grid
Contrôle	Électromécanique	Numérique
Communication	Unidirectionnel	Bidirectionnel
Génération	Centralisée	Distribuée
Détection	Capteurs limités	Utilisation généralisée des capteurs
Surveillance	Manuel	Autonome et distant
Restauration	Manuel	Autoréparation
Options offertes aux clients	Limitées	Diverses

Les caractéristiques limitées de contrôle et de surveillance du réseau traditionnel ont obligé les opérateurs de système indépendants (ISO) et les opérateurs de transmission régionaux (RTO) à améliorer les capacités de communication du réseau électrique existant pour obtenir un système plus flexible. Dans les années 80, les exigences en matière de comptage avancé ont été améliorées pour permettre de calculer la moyenne des prix de l'électricité et les limites de la sélection des tarifs ont été supprimées. À la fin des années 1990, les préoccupations environnementales se sont accrues pour empêcher la production basée sur les combustibles fossiles, qui a été l'un des jalons pour améliorer la production distribuée, la gestion de la demande et les opérations de contrôle et de surveillance décentralisées. Ainsi, des recherches sur un nouveau concept de réseau ont été lancées de manière intensive.

La production distribuée et le contrôle décentralisé ont été les principaux moteurs de l'utilisation des sources d'énergies renouvelables dans la production d'électricité. En outre, le terme de microgrid a surgi au début des années 2000, ce qui a contribué à accroître la capacité et la résilience du réseau électrique existant. Par conséquent, il a été possible de mentionner le flux d'énergie bidirectionnel et les communications bidirectionnelles dans une infrastructure améliorée du réseau électrique appelée "smart grid".

Ces améliorations ont également permis l'auto-génération et le contrôle généralisé de ce nouveau réseau électrique grâce à l'utilisation d'un grand nombre de capteurs presque à chaque nœud et ligne du réseau entier. Dès qu'une défaillance se produit dans une section du réseau, les capteurs détectent la panne et le système de protection gère le flux d'énergie en comprenant de nouvelles voies de relais pour l'énergie. Cette capacité de contrôle est fournie par trois grandes infrastructures techniques du réseau intelligent qui sont l'infrastructure intelligente, la gestion intelligente et les systèmes de contrôle intelligents. Le système de gestion intelligent est lié aux sous-systèmes de surveillance, de gestion et de prise de décision le long du réseau électrique. Le système de contrôle intelligent assure la détection et la protection contre les défaillances, la sécurité et la protection de la vie privée, les services et l'analyse du système pendant l'exploitation du réseau électrique [77] [54].

2.3 Plus d'interaction pour plus d'efficacité

Le "NIST Framework and Roadmap for future Smart Grids" [54] identifie sept domaines au sein du smart grid : Transport, Distribution, Opérations, Génération, Marchés, Consommateurs et Fournisseurs de services (figure 2.1). Selon la définition du NIST, le domaine du smart grid est un groupement de haut niveau d'organisations, de constructions, de personnes, de systèmes, d'appareils ou d'autres acteurs ayant des objectifs similaires et s'appuyant sur ou participant à des types d'applications similaires. Dans les sept domaines, de nombreux acteurs saisissent, transmettent, stockent, modifient et traitent les informations nécessaires aux applications du smart grid. Nous donnons dans ce qui suit une brève définition pour chaque domaine.

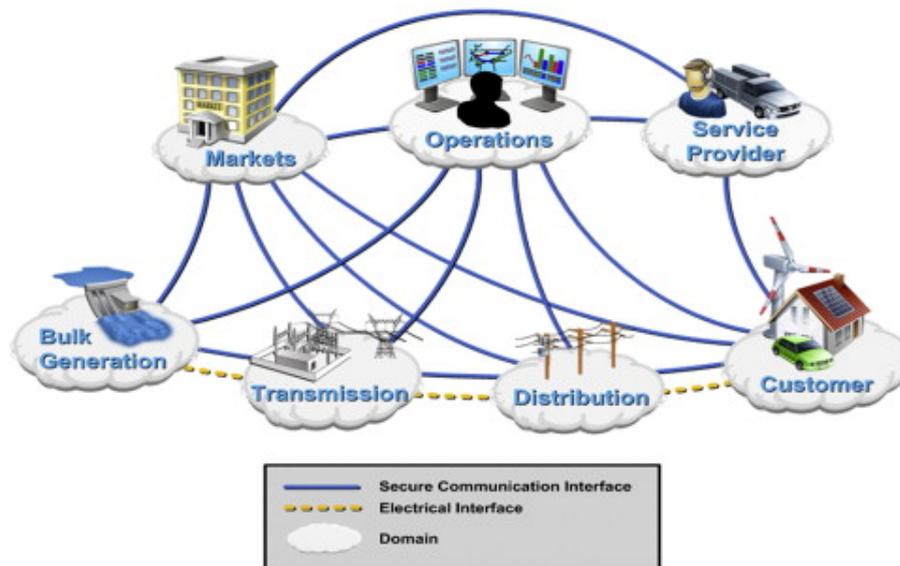


FIGURE 2.1 – Modèle conceptuel du smart grid vu par NIST [54]

1. *Génération* : l'électricité peut être produite à partir de différentes sources. Ces dernières peuvent être renouvelables ou non renouvelables. Pour les sources non renouvelables, il y a notamment le charbon, le pétrole, nucléaire, etc. Pour les sources renouvelables, il y a le solaire, l'éolien, l'hydraulique, etc. Ces sources peuvent également être centralisées ou distribuées.
2. *Transmission* : l'électricité produite est transmise au domaine de distribution par de multiples sous-stations et lignes de transmission. La transmission est généralement exploitée et gérée par un opérateur de transmission régional (RTO) ou un opérateur de système indépendant (ISO). Le RTO est chargé de maintenir la stabilité des lignes de transport régionales en équilibrant la demande et l'offre. Le domaine du transport peut également soutenir la production et le stockage d'énergie à petite échelle. Pour réaliser des fonctions d'auto-réparation et améliorer la prise de conscience de la situation dans une vaste zone et de contrôle, de nombreuses informations seront saisies à partir de la grille et envoyées au centre de contrôle supervisé par le Supervisory Control

And Data Acquisition (SCADA) (c'est-à-dire l'entité responsable de l'acquisition, du traitement, de la présentation et de la gestion des données reçues). Le centre de contrôle enverra également des réponses aux dispositifs des sous-stations distantes. Les communications bidirectionnelles entre le centre de contrôle et les sous-stations sont également traitées dans le domaine de la transmission.

3. *Distribution* : l'acheminement de l'électricité vers les utilisateurs finaux est mis en œuvre en utilisant les infrastructures électriques et de communication qui relient les domaines du transport et de la consommation. Ce domaine comprend les lignes d'alimentation de distribution et les transformateurs pour fournir l'électricité. Il interagit avec de nombreux équipements, tels que les ressources énergétiques distribuées (DER), les véhicules électriques, et les capteurs communicants. Le domaine de la distribution est chargé de fournir de l'électricité aux consommateurs en fonction des demandes des utilisateurs et de la disponibilité de l'énergie. Afin de fournir une électricité de qualité, la stabilité de ce domaine est surveillée et contrôlée.
4. *Consommation* : les clients consomment, produisent (en utilisant la DER) ou stockent de l'électricité. Ce domaine comprend les bâtiments résidentiels, commerciaux ou industriels. Il est électriquement connecté au domaine de la distribution et communique avec les domaines de la distribution, de l'exploitation, des fournisseurs de services et du marché. Le domaine des clients soutient également le processus de réponse à la demande. Pour permettre aux clients de participer activement au réseau, une interface de communication bidirectionnelle entre les locaux du client et le domaine de distribution est nécessaire. Cette interface est généralement appelée interface de services énergétiques (ESI) et est présente dans les locaux du client. Un réseau de communication dans les locaux du client est nécessaire pour permettre l'échange de données et de commandes de contrôle entre le service public et les appareils intelligents du client. Ce réseau s'appelle réseau domestique (Home Area Network - HAN). Il est censé prendre en charge des applications telles que le contrôle de la charge à distance, la surveillance et le contrôle du DER, la prise en charge de l'affichage à domicile (IHD) pour les utilisations des clients et l'intégration avec les systèmes de gestion des bâtiments.
5. *Opérations* : ce domaine maintient un fonctionnement efficace et optimal des domaines de la transmission et de la distribution en utilisant un système de gestion de l'énergie (EMS) dans le domaine de la transmission et un système de gestion de la distribution (DMS) dans le domaine de la distribution. Il utilise des réseaux de terrain et des réseaux étendus dans les domaines de la transmission et de la distribution pour obtenir des informations sur les activités du système électrique comme la surveillance, le contrôle, la gestion des pannes, la maintenance, l'analyse et le comptage. Les informations sont obtenues à l'aide des systèmes SCADA. Le domaine d'exploitation peut être subdivisé en sous-domaines pour le transport, la distribution et les opérations RTO/ISO. Ces sous-domaines peuvent être contrôlés par différentes organisations.

6. *Marchés* : l'équilibre entre l'offre et la demande d'électricité est maintenu par le domaine du marché. Ce domaine se compose des détaillants qui fournissent de l'électricité aux utilisateurs finaux, des fournisseurs d'électricité en vrac et des négociants qui achètent de l'électricité aux fournisseurs et la vendent aux détaillants. Une communication efficace entre les producteurs d'électricité en vrac, la DER et les marchés est essentielle pour faire correspondre la production d'électricité à la demande.
7. *Fournisseurs de services* : l'électricité est fournie aux clients et aux entreprises de par l'intermédiaire de fournisseurs de services. Ils gèrent des services tels que la facturation et la gestion des comptes clients pour les entreprises. Il communique avec le domaine d'exploitation pour obtenir les informations de comptage et pour la connaissance de la situation et le contrôle du système. Il doit également communiquer avec les clients via l'interface ESI pour fournir des services intelligents comme la gestion des utilisations de l'énergie et la production d'énergie à domicile.

2.4 Architecture de communication du système smart grid

Le smart grid peut être considéré comme un réseau de nombreux systèmes et sous-systèmes interconnectés intelligemment pour fournir de l'énergie d'une manière rentable et fiable. L'intelligence apportée au smart grid est réalisée par l'ajout des Technologies de l'Information et de la Communication (TIC) au réseau électrique [50]. Cette infrastructure de communication permet la collecte de données sur la production et le transport ainsi que sur les réseaux de distribution. En termes d'infrastructure de communication, le smart grid peut être divisé en trois principales entités ; le réseau domestique de communication (HAN pour Home Area Network), le réseau voisin de communication (NAN pour Neighborhood Area Network) et le réseau de communication à grande distance (WAN pour Wide Area Network) [104].

2.4.1 Home Area Network (HAN)

Le HAN est le plus petit sous système de la chaîne hiérarchique du Smart Grid. Il représente une unité résidentielle unique dotée d'appareils intelligents, d'outils de contrôle de la consommation d'énergie, de solutions de stockage, de panneaux solaires, de petites éoliennes, de véhicules électriques et de compteurs intelligents. Ces composants communiquent entre eux en utilisant des technologies de communication filaires comme les courants porteurs en ligne (CPL) et des technologies sans fil comme le Wi-Fi et Zigbee. Les technologies sans fil sont en train de devenir un choix populaire contrairement aux technologies filaires en raison de leur faible coût d'installation et de leur flexibilité [119].

2.4.2 Neighborhood Area Network (NAN)

Le réseau NAN est un ensemble de réseaux HAN, qui correspond à un groupe de maisons éventuellement alimentées par le même transformateur. Le NAN met en œuvre la connexion qui permet d'offrir plusieurs applications Smart Grid tel que le comptage intelligent (Smart Metering), la gestion

de la charge, l'automatisation de la distribution d'énergie et la gestion des pannes. L'infrastructure avancée de comptage (AMI pour Advanced Metering Infrastructure) collecte les données des équipements intelligents d'un réseau NAN et les regroupe avant qu'elles ne soient envoyées au WAN. Ce dernier, connecte les réseaux NANs à l'opérateur du réseau électrique.

2.4.3 Wide Area Network (WAN)

Le WAN relie plusieurs systèmes de distribution et sert de pont entre les réseaux NAN et HAN et le réseau d'opérateur. Comme le montre la figure 2.2, le réseau WAN fournit l'infrastructure de communication nécessaire pour connecter les équipements des clients du Smart Grid à l'opérateur du réseau électrique. Cela peut être réalisé en adoptant plusieurs technologies de communication (Ethernet, réseaux cellulaires, etc.) pour transférer les données issues des réseaux NANs à l'opérateur réseau. Parmi les principales applications qu'offre le réseau WAN on peut citer la surveillance, le contrôle et la protection à grande distance. Ces applications représentent des solutions efficaces pour améliorer la planification, l'exploitation et la protection du réseau électrique dans le Smart Grid [98]. Les applications de surveillance, de contrôle et de protection dans un réseau WAN offrent une résolution de données supérieure et un temps de réponse plus court que les systèmes classiques de contrôle, de supervision et d'acquisition de données (SCADA pour Supervisory Control And Data Acquisition) et de gestion de l'énergie (EMS pour Energy Management System). Ces derniers fournissent un intervalle de mise à jour des mesures de plusieurs secondes, voire de quelques minutes, alors que les applications de surveillance, de contrôle et de protection sur le réseau WAN fournissent des données de haute résolution, à savoir 60 échantillons par seconde [68].

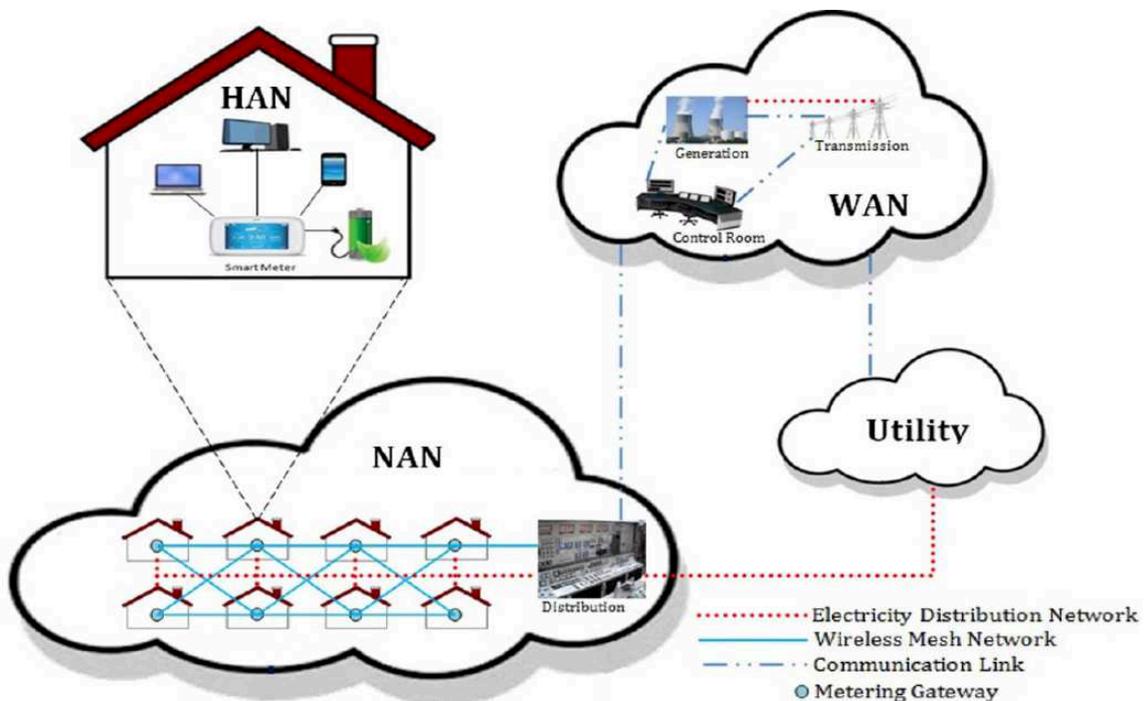


FIGURE 2.2 – Architecture réseau du Smart Grid [16]

2.5 Les principales caractéristiques d'un smart grid

L'un des majeurs objectifs du Smart Grid est le contrôle de la consommation d'électricité chez les clients en mettant en place différentes méthodes d'optimisation. Pour atteindre cet objectif, le smart grid utilise principalement : le comptage intelligent (communication bidirectionnelle) et les microgrids (Ressources d'Energie Distribuées), les mécanismes de la réponse à la demande (Demand Response) et la tarification intelligente (Smart Pricing) [16].

2.5.1 La communication bidirectionnelle

Afin de créer un réseau de distribution automatisé et largement distribué, le smart grid ajoute au réseau électrique traditionnel les avantages des technologies de la communication, notamment, des capteurs, des objets connectés ainsi que des compteurs intelligents (Smart Meter –SM) pour fournir des informations en temps réel et permettre un équilibre quasi instantané de la gestion de l'offre et de la demande. À cet égard, il est important de mettre en place une infrastructure de communication fiable permettant d'établir un transport de données en temps réel robuste via des réseaux étendus (WAN) jusqu'au niveau du client [16].

2.5.2 Les Ressources d'Energie Distribuées (DERs)

Le Smart Grid permet d'intégrer des ressources d'énergie distribuées qui comprennent plusieurs technologies : les piles à combustible, les panneaux photovoltaïques, les éoliennes, etc. Le contrôle coordonné du DER ainsi que les charges contrôlables des dispositifs de stockage, comme les volants d'inertie et les batteries, sont au cœur du concept de microgrid (MG).

Le microgrid est un réseau électrique composé : de sources de consommation énergétiques, de systèmes de production d'électricité (renouvelables ou thermiques), de systèmes de stockage de l'énergie et de systèmes de contrôle et de gestion des données [49]. Il constitue une solution efficace pour surmonter les problèmes liés aux infrastructures traditionnelles. La capacité du MG à intégrer des générateurs locaux, comme ressources de production d'énergie distribuées et des systèmes de stockage offre de nombreux avantages par rapport aux systèmes existants. Les microgrids peuvent opérer tout en étant connectés au réseau principal de distribution, où leurs systèmes de stockage permettent une optimisation locale de l'offre et la demande. De plus, ils peuvent fonctionner en mode îlot d'une manière isolée en étant déconnectées du réseau pendant une période de temps [1]. Les deux modes de fonctionnement sont gérés par l'opérateur en fonction de la situation technique ou économique du système électrique.

L'un des avantages de l'intégration des microgrids dans l'architecture du smart grid consiste à accroître la fiabilité. Cela se fait en assurant une alimentation de secours en cas de panne du réseau ou lorsque le prix de l'énergie augmente. En outre, les microgrids permettent de réduire les émissions de CO₂ en alimentant les clients par des sources d'énergie renouvelables [7].

2.5.3 La Gestion de l'équilibre du système électrique à travers les mécanismes de la réponse à la demande

La réponse à la demande (ou DR pour Demand Response), est une approche utilisée pour réduire la charge sur le réseau électrique et améliorer la fiabilité du système. La DR est une approche dans laquelle les utilisateurs finaux modifient leurs modèles de consommation d'électricité en fonction des variations des prix. Du point de vue smart grid, la réponse à la demande est un moyen efficace de réordonner la consommation d'énergie des utilisateurs afin de réduire les dépenses d'exploitation. De plus, cette technologie rend le système électrique plus fiable, améliore la transparence et l'efficacité du marché de l'électricité et engendre des avantages financiers mutuels pour l'opérateur smart grid et ses clients.

Le mécanisme de la réponse à la demande peut être réalisé en utilisant trois méthodes [32] :

- Réduire la consommation d'énergie de pointe, afin d'empêcher la charge électrique de dépasser la capacité d'alimentation des sous-stations de distribution. La satisfaction des utilisateurs serait réduite, car cette méthode atténue une partie de leur demande.
- Promouvoir la consommation d'énergie en période creuse par le biais de dispositifs de stockage d'énergie, tels que les batteries rechargeables et les véhicules électriques (EVs).
- Équilibrer la consommation d'énergie sur un horizon temporel. Par exemple, transférer une partie de la charge électrique des heures de pointe aux heures creuses.

Le concept de la réponse à la demande a été utilisé dans le secteur commercial et industriel pendant un certain temps pour améliorer la stabilité du réseau. Avec l'émergence de réseaux électriques intelligents, la DR a maintenant le potentiel d'être étendue à grande échelle aux marchés de l'électricité résidentielle [119].

2.5.4 La tarification intelligente

La tarification intelligente se présente comme une alternative au prix fixe de l'électricité. Les programmes basés sur les prix sont principalement intégrés aux programmes de la réponse à la demande dans le smart grid. Ils offrent aux utilisateurs différents prix de l'électricité à des moments différents. Sur la base de ces informations, les utilisateurs consommeront naturellement moins d'électricité lorsque les prix de l'électricité seront élevés et réduiront donc la demande aux heures de pointe. En d'autres termes, ces programmes incitent indirectement les utilisateurs à modifier de manière dynamique leurs schémas d'utilisation de l'énergie en fonction de la variation des prix de l'électricité, au lieu de contrôler directement leurs charges. On peut trouver différents programmes de tarifications basés sur l'échelle de temps [32] :

1. **Time of Use (ToU)** : représente la forme de tarification dynamique la plus utilisée. Ce mode

de tarification décompose la journée en blocs de temps auxquels est associé un prix spécifique. Lorsque les utilisateurs consomment de l'énergie à différents moments de la journée ou à différentes saisons de l'année, ils sont facturés à des prix différents. Afin d'inciter les utilisateurs à déplacer leurs charges durant la journée, le prix de l'électricité annoncé aux périodes de pointes est plus élevé à celui fixé aux périodes creuses. Dans ce type de tarification, les prix sont souvent fixés et annoncés à l'avance et restent inchangés pendant une longue période [105].

2. **Critical Peak Pricing (CPP)** : Principalement basé sur la tarification ToU sauf dans certains jours. La tarification CPP permet d'envoyer un jour à l'avance un signal d'urgence pour prévenir de la période d'extrême pointe [17]. Durant cette période de pointe, le prix augmente de manière significative par rapport au tarif habituel, de manière à inciter fortement les utilisateurs à ne pas consommer en ce moment. CPP n'est employé que pendant un nombre limité d'heures ou de jours par an, pour assurer la fiabilité du système et garantir l'équilibre entre l'offre et la demande.
3. **Real Time Pricing (RTP)** : se réfère à la tarification dynamique, où le prix de l'électricité varie généralement à des intervalles de temps différents dans la journée (toutes les 15 minutes ou toutes les heures) [5]. RTP est la tarification la plus efficace, car elle permet de transmettre heure par heure au consommateur les variations du prix du kWh sur le marché de gros. Elle donne le coût réel de l'électricité au moment où elle est consommée.
4. **Inclining Block Rate (IBR)** : est une tarification progressive qui vise à réduire la consommation globale de l'électricité. Cette tarification est conçue avec des structures tarifaires à deux niveaux (blocs inférieur et supérieur), de sorte que le prix unitaire de l'électricité, que le consommateur devra payer, augmente avec la quantité d'énergie consommée. En d'autres termes, le prix de l'électricité par consommation d'énergie atteindra une valeur supérieure si la consommation d'énergie horaire, journalière ou mensuelle de l'utilisateur dépasse un certain seuil. IBR incite les utilisateurs à répartir leurs charges entre différents moments de la journée afin d'éviter des taux plus élevés, ce qui contribue à réduire le ratio pic / moyenne du réseau [43].

2.6 Conclusion

Nous avons résumé dans ce chapitre les principes fondamentaux et l'architecture du smart grid. Ensuite, nous avons détaillé l'architecture de communication du smart grid et expliqué ses principales caractéristiques.

Nous expliquons par la suite comment tirer profit des fonctionnalités de cette technologie pour améliorer l'efficacité énergétique des data centers du cloud.

Chapitre 3

Gestion d'énergie des data centers du Cloud

Sommaire

3.1	Introduction	36
3.2	Cloud computing	36
3.2.1	Principes fondamentaux	36
3.2.2	Le rôle des data centers dans le cloud computing	39
3.3	Data centers et énergie	40
3.3.1	Composants d'un data center	40
3.3.2	Principales métriques d'efficacité énergétique des data centers	42
3.4	Interaction des data centers du cloud avec le smart grid	43
3.5	Conclusion	45

3.1 Introduction

Techniquement, le concept du Cloud Computing est loin d'être nouveau, il est même présent depuis des décennies. On en trouve les premières traces dans les années 1960, quand John McCarty affirmait que cette puissance de traitement informatique serait accessible au public dans le futur. Le terme en lui-même est apparu plus couramment aux alentours de la fin du XXe siècle et il semblerait qu'Amazon.com soit l'un des premiers à avoir assemblé des data centers et fournit des accès à des clients. Les entreprises comme IBM et Google ainsi que plusieurs universités ont seulement commencé à s'y intéresser sérieusement aux alentours de 2008, quand le Cloud Computing est devenu un concept à la mode.

3.2 Cloud computing

Nous donnons dans cette section une petite introduction aux concepts du cloud computing. Nous mettons aussi en avant les différences entre les data centers du cloud et les data centers traditionnels.

3.2.1 Principes fondamentaux

Le "cloud computing" est un terme qui désigne les services hébergés en ligne. Ces services sont accessibles via l'internet, qui est métaphoriquement appelé "cloud". Le cloud computing a reçu une attention croissante ces dernières années en raison de son modèle qui requiert une quantité limitée de ressources du côté du client. Le vaste champ d'application du cloud computing dépend de l'endroit où se trouvent les ressources informatiques, de qui les gère, et de la manière dont on y accède et dont on les paie. Le cloud computing peut recourir à la virtualisation pour séparer les logiciels des caractéristiques des serveurs physiques. En utilisant cette approche, il est souvent possible d'optimiser les serveurs et de réduire certaines caractéristiques coûteuses en énergie.

Un bon nombre d'utilisateurs interagissent déjà avec divers types de services du cloud computing. Par exemple, en utilisant un compte de courrier électronique avec le service de courrier électronique en ligne Gmail, au lieu d'exécuter un programme de courrier électronique sur l'ordinateur, l'utilisateur se connecte à distance à son compte de courrier électronique. Le logiciel et le stockage du compte n'existent pas sur l'ordinateur - ils se trouvent dans le cloud du fournisseur de services de courrier électronique.

Divers facteurs ont contribué à une utilisation de plus en plus répandue du cloud computing, notamment la disponibilité accrue de l'internet, les technologies modernes telles que la virtualisation et les nouveaux modèles de fourniture de services sur le web. Le modèle commercial du fournisseur du cloud computing est basé sur la fourniture d'une infrastructure composée de grands pools de ressources informatiques à haute performance et de dispositifs de stockage à haute capacité qui sont partagés entre les services offerts par le fournisseur [11]. Les clients qui s'abonnent aux services voient leurs données hébergées par le fournisseur et se voient allouer des ressources informatiques à la demande à partir du pool, selon les besoins. L'offre du fournisseur du cloud computing peut également s'étendre

aux applications logicielles requises par les clients. Pour réussir, le modèle du cloud computing nécessite également un réseau à haut débit pour assurer la connexion entre les clients et l'infrastructure du fournisseur du cloud computing.

Le cloud computing offre des avantages économiques aux entreprises, car elles peuvent réduire leurs dépenses informatiques et bénéficier d'une plus grande flexibilité et d'une meilleure scalabilité. Un avantage financier global est également potentiellement accru par des économies importantes dans le budget énergétique d'une entreprise. Traditionnellement, chaque entreprise disposait de ses propres data centers dédiés avec leurs propres serveurs, applications logicielles et plateformes de développement. Ceux-ci doivent être payés, gérés et entretenus en interne. L'achat de matériel et de logiciels entraîne des dépenses d'investissement élevées et la mise à jour et la maintenance de ces systèmes peuvent prendre beaucoup de temps.

Les avantages du cloud computing ont beaucoup à voir avec les importantes économies d'échelle que les fournisseurs de cloud computing peuvent réaliser. Au lieu de disposer de ressources informatiques internes pour chaque entreprise, un fournisseur de cloud computing fournit des services à plusieurs entreprises en partageant ses ressources. Le concept de partage des ressources permet d'exploiter les ressources informatiques de manière beaucoup plus efficace que les ressources informatiques internes. Des utilisateurs distants plus diversifiés et situés à différents endroits permettent de répartir les charges informatiques sur la journée, ce qui permet d'accroître l'utilisation des équipements. Les installations dans le nuage présentent également un avantage substantiel en permettant aux départements de se consacrer à l'efficacité. La technologie évolue si rapidement qu'il est difficile pour les personnes qui ne se consacrent pas à l'efficacité de suivre aussi bien que celles qui le font.

Certaines grandes entreprises ont adopté une approche de cloud "privé", ce qui signifie qu'elles ont consolidé tous leurs serveurs et applications en un seul système, qui est maintenu derrière leur pare-feu et accessible via leur intranet, c'est-à-dire qu'elles partagent leurs ressources entre les différents bureaux et unités commerciales. Cela signifie qu'ils peuvent faire un meilleur usage de leurs ressources, plus efficace, car elles sont partagées par un plus grand nombre de personnes. Cela signifie également que leurs systèmes sont uniformes dans toute l'entreprise.

Cependant, le terme familier du cloud computing fait généralement référence aux clouds publics, par lesquels les entreprises achètent les services nécessaires à des fournisseurs de cloud computing spécialisés. Au lieu de posséder et de gérer leurs propres data centers, les entreprises accèdent aux services qu'elles ont achetés de n'importe où via Internet. Les systèmes de gestion de la relation client d'une entreprise, par exemple, peuvent être gérés par un fournisseur de cloud computing et les informations sont stockées dans les centres de données du fournisseur, tandis que le système de messagerie électronique de l'entreprise peut être géré par un autre fournisseur de cloud computing. Ces services peuvent être fournis à distance, rapidement et à la demande, de manière autonome, et l'entreprise ne paie que ce dont elle a besoin et ce qu'elle utilise.

Le cloud computing n'est pas une solution universelle. Il a des avantages et des inconvénients, et il est essentiel de les comprendre pour décider s'il répond à un besoin particulier.

Les principaux avantages du cloud computing sont les suivants :

- Économie : partage des ressources informatiques entre différents clients.
- Paiement à l'utilisation : les clients paient pour le service au lieu d'acheter des licences de logiciels et du matériel.
- Utilisation/flexibilité à la demande : les services cloud peuvent être utilisés presque instantanément et peuvent facilement être étendus
- Stockage externe des données : les données des clients sont stockées à l'extérieur, chez le fournisseur cloud computing
- Des services hautement fiables : les clouds se gèrent d'eux-mêmes en cas de défaillance ou de dégradation des performances.

D'autre part, les inconvénients les plus cités du cloud computing sont

- Dépendance à l'égard de la connexion Internet : une connexion constante est nécessaire
- Perte de contrôle : quelqu'un d'autre héberge du matériel, des logiciels et des données, ce qui peut poser des problèmes de sécurité
- Coût imprévisible : le paiement à l'utilisation signifie que le coût de l'informatique sera différent chaque mois

Le modèle de service du cloud computing est un modèle commercial solide, tant du point de vue du fournisseur que du client. Les fournisseurs investissent dans l'infrastructure et la gestion nécessaires et reçoivent en retour un flux de revenus régulier de l'investissement par le biais d'abonnements aux services [20]. Étant donné qu'à tout moment, un nombre important de clients sont inactifs, le fournisseur profite des bénéfices des économies d'échelle et peut facturer des frais d'abonnement moins élevés. Les clients voient à leur tour les avantages de la disponibilité des données et des services en tout lieu, de la gestion centralisée des sauvegardes de données, de la disponibilité d'une capacité accrue en cas de besoin et de la facturation en fonction de l'utilisation [115], [53].

Le modèle de cloud computing peut être décomposé en trois segments : "logiciel", "plate-forme" et "infrastructure". Chaque segment sert un objectif différent et offre des produits différents aux entreprises et aux particuliers dans le monde entier. Les services fondamentaux et pratiques du cloud computing sont connus sous le nom de :

- Logiciel en tant que service (SaaS)
- Plate-forme en tant que service (PaaS - parfois appelé middleware)
- Infrastructure as a service (IaaS - parfois appelé Hardware-as-a-Service) [28].

L'IaaS donne au consommateur le contrôle d'un type particulier de système d'exploitation, de stockage, de base de données, de réseau et d'applications déployées sur le cloud. Le PaaS donne au consommateur la possibilité de déployer des applications, mais lui interdit de gérer ou de contrôler l'infrastructure du cloud sous-jacente. Le PaaS simplifie le processus de développement de logiciels d'un ordre de grandeur, mais il n'est pas aussi flexible que l'IaaS car certains détails sont cachés à l'utilisateur final. Enfin, le SaaS permet au consommateur d'utiliser les applications du fournisseur fonctionnant sur une infrastructure de cloud, généralement via une interface de client léger, comme

un navigateur web. Dans les environnements SaaS, le fournisseur de cloud computing a un contrôle total sur l'application, y compris les capacités, les mises à jour et la maintenance.

3.2.2 Le rôle des data centers dans le cloud computing

Le secteur des data centers évolue rapidement pour offrir un mélange complexe de services, de modèles commerciaux et de déploiements d'infrastructures, car les consommateurs et les entreprises exigent une connectivité 24 heures sur 24, 7 jours sur 7, une plus grande flexibilité des ressources et une réduction des coûts des services informatiques. Bien que les salles de serveurs des petites entreprises représentent toujours une grande partie du marché, diverses options de data centers ont vu le jour, des data centers privés et hybrides aux fournisseurs d'hébergement de data centers en colocation ou à plusieurs locataires. Cette évolution est également marquée par la croissance des services de cloud computing qui sont hébergés et exploités en grande partie par des entreprises telles qu'Amazon, Apple, eBay, Facebook, Google, Microsoft et Salesforce.com [83].

Toute organisation a besoin d'un data center, quelle que soit sa taille ou son secteur d'activité. Un data center est traditionnellement une installation physique que les entreprises utilisent pour stocker leurs informations ainsi que d'autres applications qui font partie intégrante de leur fonctionnement. Un data center est souvent composé d'équipements techniques il peut s'agir de routeurs, de dispositifs de sécurité, de systèmes de stockage. Pour maintenir tout le matériel et les logiciels à jour et en fonctionnement, un data center nécessite une quantité importante d'infrastructures. Ces installations peuvent comprendre des systèmes de ventilation et de refroidissement, des alimentations électriques sans coupure, des générateurs de secours, etc.

Un data center du cloud est très différent d'un data center traditionnel, il n'y a rien de similaire entre ces deux systèmes, si ce n'est le fait qu'ils stockent tous deux des données. Un data center du cloud n'est pas physiquement situé en interne d'une organisation particulière et tout est en ligne. Lorsque les données sont stockées sur des serveurs du cloud, elles sont automatiquement fragmentées et dupliquées à différents endroits pour un stockage sécurisé. En cas de défaillance, votre fournisseur de services cloud s'assurera qu'il y a également une sauvegarde de vos données.

Avec un data center traditionnel, vous devrez faire divers achats, notamment les serveurs et les équipements réseau. Non seulement c'est un inconvénient en soi, mais ce matériel doit être remplacé à mesure qu'il vieillit et devient obsolète. En plus du coût d'achat du matériel, il est également nécessaire d'engager du personnel pour superviser son fonctionnement. Lorsque les clients hébergent leurs données sur des serveurs du cloud, ils utilisent essentiellement le matériel et l'infrastructure d'un autre organisme, ce qui permet d'économiser beaucoup d'argent. De plus, il prend en charge divers éléments liés à la maintenance, ce qui aide les clients à mieux optimiser leurs ressources. Nous présentons dans ce qui suit la différence entre les data centers du cloud et les data centers traditionnels en termes de : Accessibilité, Sécurité et Scalabilité.

1. **Accessibilité** : un data center traditionnel offre une certaine flexibilité en termes d'équipement, ce qui permet de savoir exactement quels logiciels et matériels utiliser. Cela facilite les personnalisations ultérieures et permet d'apporter des modifications selon les besoins.
Avec l'hébergement en cloud, l'accessibilité peut devenir un problème. Dans l'absence de connexion Internet, par exemple, les données à distance deviendront inaccessibles, ce qui pourrait poser un problème pour certains. Toutefois, d'un point de vue réaliste, ces cas de non connexion à Internet peuvent être très rares et cet aspect ne devrait donc pas poser trop de problèmes.
2. **Sécurité** : les data centers traditionnels doivent être protégés de la manière traditionnelle : il faudra engager du personnel de sécurité pour garantir la sécurité des données. L'avantage est qu'il y aura un contrôle total sur les données et les équipements, ce qui les rend plus sûrs dans une certaine mesure.
L'hébergement dans le cloud peut, du moins en théorie, être plus risqué, car toute personne disposant d'une connexion Internet peut illégitimement avoir accès aux données des clients. En réalité, la plupart des fournisseurs de services dans le cloud ne négligent aucun détail pour garantir la sécurité des données. Ils emploient du personnel expérimenté pour s'assurer que toutes les mesures de sécurité requises sont en place afin que les données soient toujours entre de bonnes mains.
3. **Scalabilité** : la construction de sa propre infrastructure à partir de zéro exige beaucoup d'efforts, tant sur le plan financier qu'humain. Il faudra, entre autres, superviser sa propre maintenance ainsi que son administration. La mise en place d'un data center traditionnel est une opération qui coûte cher. Avec l'hébergement dans le cloud, en revanche, il n'y a pas de coût initial en termes d'achat d'équipement, ce qui permet de réaliser des économies qui peuvent être utilisées ultérieurement pour faire évoluer les ressources. Les fournisseurs de services cloud proposent de nombreux plans flexibles pour répondre aux besoins du client, en lui permettant d'acheter plus d'espace de stockage au moment où il le souhaite.

3.3 Data centers et énergie

Nous présentons dans cette section les principaux composants des data centers, et nous détaillons leurs métriques d'efficacité énergétique les plus utilisées.

3.3.1 Composants d'un data center

Les data centers sont un élément essentiel pour le bon fonctionnement des entreprises, des communications, des systèmes universitaires et gouvernementaux. Les petites entreprises possèdent des data centers internes qui sont généralement installés dans de grands bâtiments commerciaux. Les plus grandes entreprises construisent généralement leurs data centers sur des infrastructures de bâtiments indépendants (distants) pouvant atteindre plusieurs centaines de milliers de mètres carrés et dédiés uniquement à l'hébergement de périphériques informatiques. De plus, les grandes entreprises peuvent contenir des dizaines, voire des centaines de data centers. Un petit centre de données peut consommer

jusqu'à des dizaines de kilowatts (KWh) alors que la consommation électrique de pointe d'un grand centre de données peut atteindre des dizaines de mégawatts (MWh) [88]

Un data center comporte trois éléments principaux : (a) les équipements informatiques, (b) le système d'alimentation électrique et (c) le système de refroidissement. La figure 3.1 montre les composants typiques d'un data center. Nous décrivons ci-dessous plus en détails chacun de ces trois principaux composants [90] :

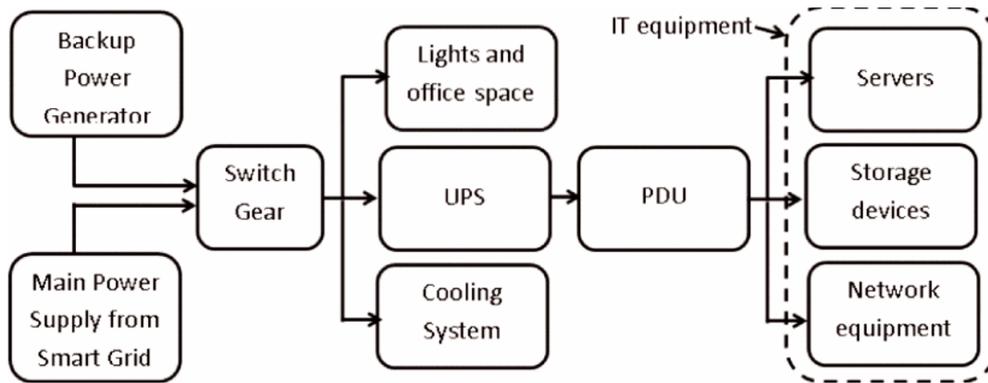


FIGURE 3.1 – Principaux composants d'un data center [90]

1. Les équipements IT :

Il existe principalement trois types d'équipements électroniques hébergés dans un data center typique : (i) les serveurs pour le traitement des données, (ii) les équipements de stockage pour le stockage des données, et (iii) les équipements de réseau pour les communications de données. Ces trois types d'équipements sont connus collectivement sous le nom d'équipements "IT".

2. Système d'alimentation électrique :

Contient principalement des unités de conversion d'énergie (AC en DC et DC en AC), des régulateurs de tension et des équipements de sauvegarde. Le stockage de l'énergie est souvent assuré par une unité d'alimentation sans coupure (UPS) qui empêche l'équipement informatique de subir des interruptions de courant et d'éventuelles perturbations de l'activité ou pertes de données.

3. Système de refroidissement :

Le système de refroidissement dans les data centers est souvent assuré par les unités de climatisation des salles informatiques (CRAC). En général, l'air entre dans la partie supérieure d'une unité CRAC où il est encore conditionné en passant sur des serpentins contenant de l'eau froide pompée d'un refroidisseur situé à l'extérieur du bâtiment. L'air refroidi est ensuite acheminé vers l'équipement informatique (principalement les serveurs) par un plénum de plancher sur-élevé. L'air froid, en passant à travers des dalles de sol perforées, est aspiré par les ventilateurs situés à l'intérieur des serveurs (figure 3.2).

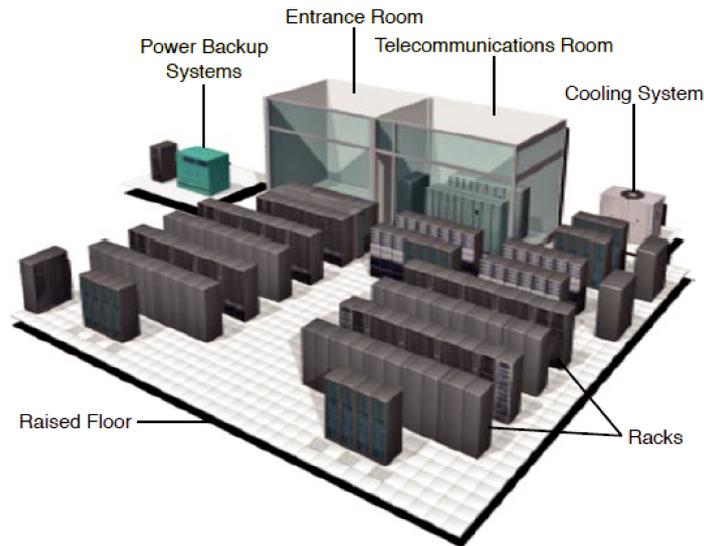


FIGURE 3.2 – Vue globale d’un data center [90]

3.3.2 Principales métriques d’efficacité énergétique des data centers

Dans cette section, nous présentons les principales métriques qui permettent d’avoir un data center écologique. Les principaux objectifs du développement des métriques écologiques sont 1) de décrire quantitativement le degré écologique d’un data center, 2) de surveiller et de rapporter la consommation d’énergie et les émissions de carbone d’un data center pour augmenter son efficacité écologique, et 3) de fournir une visualisation claire de l’efficacité d’un data center pour aider ses opérateurs à réduire leurs factures d’électricité et leurs taxes carbone. [63]

Diverses métriques axées sur l’efficacité énergétique des data centers sont apparues ces dernières années. Nous résumons ci-dessous les métriques écologiques les plus populaires :

1. Power Usage Effectiveness (PUE) :

C’est la mesure la plus populaire en matière d’efficacité énergétique des data centers. Il s’agit du rapport entre l’énergie totale consommée par le data center et l’énergie consommée par ses équipements informatiques [12].

$$PUE = \frac{\text{Énergie totale consommée}}{\text{Énergie consommée par les équipements IT}} \quad (3.1)$$

2. Efficacité de l’infrastructure des data centers (Data Center infrastructure Efficiency - DCiE) :

Cette mesure est simplement l’inverse du PUE, et les deux ont été développés par The Green Grid pour mesurer l’efficacité énergétique dans les data centers. Barroso et al. [12] ont suggéré que la valeur moyenne du PUE est de 1.83, c’est-à-dire que le DCiE est de 0.54.

$$DCiE = \frac{1}{PUE} \quad (3.2)$$

3. Efficacité énergétique de calcul (Compute Power Efficiency - CPE) :

Une autre approche pour évaluer les performances d'un data center consiste à calculer sa surcharge électrique, ensuite à prendre en compte l'utilisation moyenne des processeurs des serveurs qu'il héberge. Connue sous le nom de "Compute Power Efficiency" (CPE), cette mesure a été introduite pour la première fois en 2007 [63] :

$$CPE = \frac{\textit{Utilisation des équipements IT} \times \textit{Consommation d'énergie des équipements IT}}{\textit{Consommation d'énergie totale}} \quad (3.3)$$

4. Coefficient d'énergie verte (Green Energy Coefficient - GEC) :

Proposé en 2009, il fournit le rapport entre l'énergie renouvelable produite sur place et l'énergie totale consommée [101]. Il incite les data centers à utiliser des énergies renouvelables.

$$GEC = \frac{\textit{Énergie renouvelable produite (kwh)}}{\textit{Énergie totale consommée (kwh)}} \quad (3.4)$$

Dans l'ensemble, les paramètres des data centers sont encore en évolution. Même ceux qui sont les plus utilisés aujourd'hui sont relativement récents, introduits il y a seulement quelques années. De nouvelles formules continuent d'être développées et proposées tous les deux ou trois ans, le plus souvent par des groupes industriels ou des sociétés de conseil en technologie, tandis que les mesures existantes sont affinées et clarifiées en permanence par leurs concepteurs.

3.4 Interaction des data centers du cloud avec le smart grid

Victimes de leur développement, le cloud computing et le Smart grid sont confrontés à plusieurs défis. En particulier, les grands fournisseurs de services de cloud computing voient leur facture annuelle d'électricité augmenter et qui pourraient atteindre les 67 millions de dollars [33], chiffre qui continue d'augmenter avec l'essor des services du cloud computing et la hausse du prix de l'électricité. En parallèle, le smart grid intègre un grand nombre de ressources d'énergie distribuées, telles que les panneaux solaires et les éoliennes et doit également garantir une stabilité opérationnelle élevée. Cela peut générer des problèmes économiques, vu la nature intermittente de la production d'énergie décentralisée.

Les préoccupations susmentionnées du cloud computing et du smart grid peuvent être atténuées par une coopération appropriée entre les deux parties. La réponse à la demande des data centers peut être un grand atout pour le smart grid, car elle est motivée par les demandes des utilisateurs pouvant être réparties sur des data centers géographiquement distribués et desservies par plusieurs sources d'énergie. Ainsi, l'électricité consommée par un data center à partir du smart grid peut être ajustée de manière flexible en équilibrant la charge de travail, ou en modulant la génération de l'électricité sur le site. Dans ce contexte, l'adoption de moyens de stockage par les data centers dans le smart grid permet de résoudre plusieurs problèmes et présente de multiples intérêts aux différents acteurs concernés : producteurs, consommateurs et gestionnaires de réseaux. Les systèmes de stockage d'énergie pourront permettre de garantir l'équilibre et la stabilité du réseau tout en améliorant l'efficacité des installations

de transmission et de distribution. Ils représentent des solutions ayant un rôle primordial à jouer aussi bien au niveau économique qu'environnemental en garantissant une fourniture d'énergie électrique efficace et durable [42].

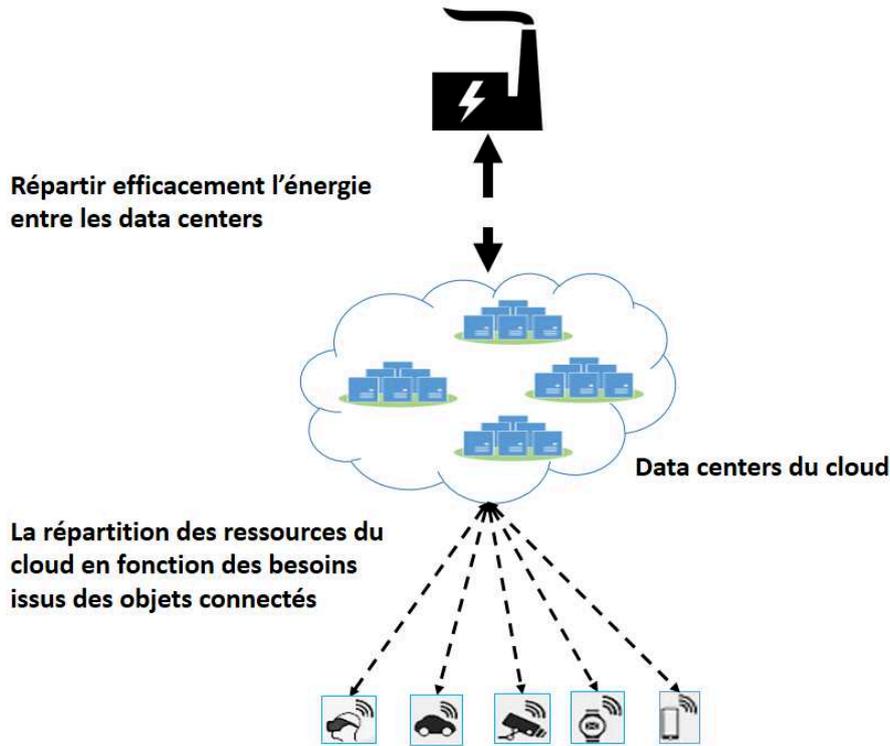


FIGURE 3.3 – Interaction des data centers avec le smart grid et ses utilisateurs

L'interaction des data centers du cloud avec le gestionnaire du smart grid se fait d'une manière directe, en échangeant les données via l'infrastructure de communication du smart grid. Mais aussi, d'une manière indirecte via les requêtes issues des utilisateurs vers les data centers du cloud. La figure 3.3 illustre l'interaction des data centers du cloud avec le smart grid et ses utilisateurs.

Dans un système smart grid – cloud, les requêtes des utilisateurs sont regroupées par un opérateur cloud. Ce dernier se charge de répartir les requêtes vers les data centers d'un cloud suivant une politique prédéfinie. Exécuter ces requêtes tout en garantissant un niveau de qualité de service satisfaisant reste parmi les plus grands challenges à relever. Dans ce contexte, les différentes requêtes issues des objets connectés (IoTs) exigent un traitement intensif et une faible latence, en particulier les applications temps réel. Il est très difficile pour le cloud computing de faire face à cette communication de plus en plus contraignante et de satisfaire toutes les demandes.

3.5 Conclusion

Nous avons présenté dans ce chapitre, les principes fondamentaux du cloud computing et avons mis l'accent sur le rôle des data centers dans cette technologie. Puis, nous avons présenté les principaux composants d'un data center et les métriques d'efficacité d'énergie les plus utilisées dans les data centers. Enfin, nous avons expliqué l'interaction des data centers du cloud avec le smart grid. Nous présentons dans ce qui suit un état de l'art sur les besoins et les solutions existantes dans la gestion d'énergie des data centers du cloud.

Gestion d'énergie des data centers : Solutions existantes et besoins

Sommaire

4.1	Introduction	48
4.2	Techniques d'efficacité énergétique des équipements non-IT d'un data center	48
4.3	Techniques d'efficacité énergétique des serveurs d'un data center	48
4.4	Etat de l'art sur la gestion d'énergie des data centers dans un système Smart Grid – Cloud	49
4.4.1	Architecture Smart Grid – Cloud	50
4.4.2	Architecture Smart Microgrid - Cloud	50
4.5	Les outils d'aide à la décision dans un système Smart Grid-Cloud	51
4.5.1	Système multi-agents et la théorie des jeux	52
4.5.2	Optimisation convexe	53
4.5.3	Processus de décision Markovien	53
4.5.4	La logique floue	54
4.6	Discussion et conclusion	54

4.1 Introduction

L'amélioration de l'efficacité énergétique des data centers en termes de coût, de consommation ou bien d'impact sur l'environnement devient un sujet qui attire beaucoup de chercheurs. De nombreux travaux de recherche traitent le problème de la consommation élevée d'énergie des data centers et proposent des solutions mettant en œuvre différentes techniques. Dans cette section, nous regroupons les efforts qui ont été menés à ce sujet en trois sous-sections : les techniques d'efficacité énergétique des installations hors équipements informatiques, les techniques d'efficacité énergétique des serveurs d'un data center et enfin les techniques d'efficacité énergétique d'un ensemble de data centers.

4.2 Techniques d'efficacité énergétique des équipements non-IT d'un data center

Un data center regroupe plusieurs installations : des serveurs, des onduleurs, des équipements de refroidissement, une salle d'opération, etc. La majeure partie de l'énergie consommée par un data center est utilisée pour refroidir les serveurs et les onduleurs. La règle de base pour la gestion des installations des data centers est d'organiser de manière appropriée des îlots d'air chaud et des îlots d'air froid. Un îlot d'air froid, se forme de la prise d'air froid des serveurs se faisant face, tandis que l'îlot d'air chaud se forme de la sortie d'air chaud des rangées de serveurs. L'élévation des sols ainsi que l'utilisation d'équipements de refroidissement à des températures plus élevées font partie des techniques permettant d'économiser de l'énergie. Afin de réduire la consommation d'énergie des conteneurs des data centers, Endo et al. [38] ont combiné le refroidissement direct à l'air frais au refroidissement par évaporation pour atténuer la chaleur produite par les équipements IT du data center. Somani et al. [95] ont présenté un algorithme appelé gestion de la charge basée sur l'intelligence ambiante, qui améliore la capacité de dissipation de la chaleur du data center. L'algorithme utilise la température d'entrée des racks de serveurs et répartit les charges de travail en fonction de l'environnement thermique du data center. Pour analyser l'efficacité des techniques mises en œuvre, la dynamique des fluides numériques (CFD pour Computational fluid dynamics) est utilisée pour modéliser le flux d'air et le transfert de chaleur dans les data centers. Joshi et al. [64] ont classifié les modèles CFD des data centers en deux catégories : modélisation du flux d'air permettant de prévoir les débits des tuiles perforées et modélisation des effets thermiques de la disposition du rack et de la distribution de l'énergie. Ils ont aussi présenté des schémas alternatifs d'approvisionnement et de retour et quelques métriques de performances thermiques. Généralement, l'efficacité énergétique des installations d'un data center est indépendante des concepts du Smart Grid.

4.3 Techniques d'efficacité énergétique des serveurs d'un data center

Considérant la disponibilité d'un ensemble de data centers, cette catégorie améliore la consommation énergétique de chaque data center individuellement, par le biais de plusieurs méthodes comme, par exemple, minimiser l'utilisation de leurs équipements informatiques tels que les processeurs. L'efficacité énergétique des équipements informatiques au sein d'un data center est mise en œuvre via un

matériel économe en énergie, ainsi que par la virtualisation et la consolidation dynamique des charges de travail [22]. Bien que les data centers contiennent des commutateurs réseaux intra-data centers, la consommation totale des commutateurs centraux, d'agrégation et d'accès représente environ la moitié de la consommation électrique des équipements IT [70]. Par conséquent, la plupart des efforts réalisés à ce sujet dans la littérature portent sur l'efficacité énergétique des serveurs d'un data center.

Gandhi et al. [44] proposent un algorithme qui permet de trouver une allocation optimale de l'électricité aux serveurs d'un data center, afin de minimiser leur temps moyen de réponse. Pour ce faire, ils exploitent deux mécanismes de redimensionnement de la tension et de la fréquence du processeur : le dimensionnement dynamique de la fréquence (DFS pour Dynamic Frequency Scaling) et le dimensionnement dynamique de la tension et de la fréquence (DVFS pour Dynamic Voltage and Frequency Scaling). Ces derniers, comptent parmi les techniques les plus exploitées pour réduire la consommation d'énergie des serveurs. Di et al. [108] proposent une application qui permet de virtualiser la distribution d'électricité dans un data center. Leur proposition permet aux serveurs de définir leurs besoins en énergie et par conséquent, l'application répartit l'électricité d'une manière équitable. Liu et al. [74] considèrent diverses sources d'énergie et proposent une architecture de gestion d'énergie nommée DiPSN. Cette dernière, connecte efficacement différentes sources d'énergie à des serveurs appropriés.

L'allocation de ressources et la migration de machines virtuelles (VM) considérant l'optimisation d'énergie ont été largement prises en compte. Dai et al. [31] ont modélisé le placement des machines virtuelles dans les serveurs d'un data center en tant que problème d'optimisation linéaire. Les auteurs ont développé deux algorithmes d'approximation appelés MinES et MinCS afin d'obtenir la solution optimale qui minimise la consommation d'énergie et garantit le niveau de service des utilisateurs. Sharma et al. [93] proposent une approche hybride qui combine Algorithmes Génétiques (AG) et Optimisation de l'essaimage de particules (PSO pour Particale Swarm Optimization) appelée HGAPSO. Cet algorithme permet la migration des machines virtuelles tout en réduisant la consommation d'énergie et en évitant la violation du niveau de service prédéfini. Dans un autre travail, Thuan et al. [36] traitent le problème du placement et migration optimale des machines virtuelles pour minimiser l'utilisation des ressources et la consommation d'énergie des data centers du cloud. Ils ont modélisé le problème en tant que fonction multi-objectif et ont résolu le problème en proposant deux algorithmes multi-niveaux qui combinent placement et migration des machines virtuelles.

4.4 Etat de l'art sur la gestion d'énergie des data centers dans un système Smart Grid – Cloud

Nous considérons dans cette sous-section les techniques conçues pour un ensemble de data centers géographiquement distribués tels que ceux détenus et exploités par Google, Yahoo, Microsoft, Amazon, eBay, etc. L'idée générale est de diriger les requêtes des utilisateurs vers un ou plusieurs data centers d'une manière à minimiser leur coût, leur consommation ou leurs émissions en CO₂.

Cette catégorie tire parti des fonctionnalités offertes par le smart grid pour améliorer l'efficacité énergétique des data centers dans le cloud. En outre, nous examinerons l'adoption des data centers pour

l'hébergement de services de smart grid afin de broser un tableau complet des interactions entre le Smart grid et les data centers du cloud. Nous nous concentrerons sur les techniques d'efficacité énergétique pour plusieurs data centers et leurs réseaux de transport correspondants. Ces techniques sont appliquées sous deux architectures : une architecture smart grid – cloud et une architecture Smart microgrid – cloud.

4.4.1 Architecture Smart Grid – Cloud

Plusieurs travaux de recherches ont considéré l'interaction des data centers du cloud avec le smart grid pour les fonctionnalités qu'offre ce dernier comme la tarification dynamique et la réponse à la demande (DR). Gu et al. [47] considèrent un système Smart Grid – cloud où les data centers du cloud sont alimentés par le Smart Grid et par l'énergie renouvelable qu'il autoproduit. Ils proposent un algorithme de répartition de requêtes des utilisateurs qui minimise les coûts énergétiques des data centers du Cloud. Ce travail a été étendu dans [48] en proposant un ordonnanceur vert qui minimise à la fois le coût énergétique et les émissions de carbone. Dans un autre travail, Huang et al. [110] ont modélisé l'interaction entre les réseaux électriques et les data centers du cloud sous forme de problème en deux étapes. Ils utilisent d'abord la tarification dynamique pour réaliser l'équilibre de charge, puis, dans un second temps, les data centers réagissent aux prix de l'énergie et gèrent la répartition des requêtes des utilisateurs pour minimiser leur coût total en énergie. Wang et al. [109] proposent un algorithme de réponse à la demande qui considère les énergies renouvelables afin de minimiser les coûts énergétiques. Leur proposition optimise conjointement l'approvisionnement en énergie et l'allocation de la charge de travail issue des utilisateurs dans les data centers du cloud. Kiani et al. [69] visent à maximiser les profits associés à l'exploitation des énergies renouvelables dans les data centers répartis géographiquement. Ils considèrent dans leur système à la fois la production d'énergie renouvelable et la tarification dynamique des marchés de l'électricité. Deng et al. [34] proposent un algorithme stochastique de répartition des ressources afin d'optimiser les coûts énergétiques et l'émission du CO₂ des data centers. Kawsar et al. [51] visent à minimiser le coût énergétique des data centers et proposent une approche qui planifie la charge de travail des data centers, en tenant compte des énergies renouvelables disponibles, du sous-système de refroidissement et de la structure tarifaire de l'électricité. Lu et al. [75] considèrent le transfert de données en masse entre les data centers pour réduire les coûts opérationnels des data centers géographiquement distribués. Ils modélisent une approche efficace d'optimisation en deux étapes pour réduire le coût énergétique des data centers dans un marché à plusieurs sources d'électricité. Yuan et al. [129] utilisent la variation spatiale des prix de l'électricité et la production d'énergie renouvelable pour optimiser conjointement les revenus et le coût énergétique du fournisseur cloud. Ils modélisent le problème comme un problème d'optimisation multi-objectifs et le résolvent par un algorithme évolutionnaire.

4.4.2 Architecture Smart Microgrid - Cloud

Considérer l'interaction des data centers du cloud avec un opérateur smart grid a conduit à des résultats significatifs en termes d'efficacité énergétique. Cependant, certaines études estiment que faire tourner des data centers dans des microgrids peut être plus efficace en termes d'efficacité énergétique,

étant donné que les microgrids peuvent améliorer la durabilité et la fiabilité des services électriques. Yu et al. [127] étudient le problème de la gestion de l'énergie des data centers du Cloud alimentés par des Microgrids. Ils minimisent le coût énergétique en considérant les paramètres d'efficacité de chargement et de déchargement des batteries. Ils étudient également dans [128] le problème de minimisation des coûts de l'énergie et des émissions de carbone des data centers alimentés par des Microgrids. Pour ce faire, ils modélisent un problème d'optimisation stochastique et le résolvent en utilisant la technique d'optimisation de Lyapunov. Jiang et al. [126] ont modélisé de la même manière le problème de minimisation du coût énergétique des data centers en prenant en considération les pannes de courant. Ding et al. [35] proposent une approche innovante de la gestion de l'énergie pour réduire les frais opérationnels d'un data center alimenté par un microgrid. Ils ont modélisé la chaleur résiduelle recueillie du data center avec d'autres ressources afin de minimiser le coût d'exploitation du data center. Baris et al. [4] étudient les effets de l'augmentation de la consommation énergétique des data centers dans un système smart microgrid - cloud et proposent une solution innovante qui prend en compte toutes les limitations des demandes et minimise les facteurs d'instabilité sur le réseau électrique.

4.5 Les outils d'aide à la décision dans un système Smart Grid-Cloud

Les problèmes d'efficacité énergétiques dans un système Smart Grid – Cloud sont souvent modélisés en problèmes d'optimisation. Ces derniers peuvent être résolus par plusieurs approches. Dans cette section, nous présentons les principales approches de modélisation utilisées dans les travaux étudiés dans l'état de l'art. Nous nous sommes basés sur ces outils d'aide à la décision pour apporter nos contributions.

Les techniques d'efficacité énergétique dans un système Smart Grid – Cloud nécessitent l'interaction des data centers avec l'opérateur du Smart Grid. Chacune de ces deux entités a un comportement vis-à-vis de l'autre. Ces comportements peuvent être modélisés en utilisant les fonctions d'utilité et les fonctions de coûts.

- Fonction d'utilité : les comportements des différents data centers sont modélisés par différents choix de fonctions d'utilité. Plus formellement, l'utilité représente le niveau de confort / satisfaction obtenue par le data center en fonction de sa consommation d'énergie, non décroissante et concave. Généralement, les fonctions d'utilité les plus considérées sont les fonctions d'utilité quadratiques [92].
- Fonction de coût : le coût de la production et de la fourniture d'électricité par l'opérateur Smart Grid est modélisé par la fonction de coût, qui est croissante et strictement convexe. Deux alternatives peuvent être adoptées : la fonction de coût linéaire par morceaux (correspondant à IBR pour Inclining Block Rate) et la fonction de coût quadratique [81].

4.5.1 Système multi-agents et la théorie des jeux

Un système multi-agents (SMA) est un groupe d'agents autonomes qui agit dans un environnement pour atteindre un objectif commun. Cela se fait par la coopération ou la compétition, le partage ou non des connaissances entre eux [10].

Les systèmes multi-agents ont été largement utilisés dans de nombreux domaines en raison des avantages qu'ils offrent. Certains des avantages offerts par l'utilisation de cette technologie sont les suivants :

- Augmentent la vitesse et l'efficacité des opérations grâce au calcul parallèle et au fonctionnement asynchrone.
- Permettent une dégradation graduelle du système en cas de panne d'un ou plusieurs agents. Ils augmentent ainsi la fiabilité et la robustesse du système.
- Scalabilité et flexibilité : des agents peuvent être ajoutés au fur et à mesure des besoins.
- Réduisent les coûts : les agents individuels coûtent en effet beaucoup moins cher qu'une architecture centralisée.
- Les agents ont une structure modulaire et peuvent être facilement remplacés dans d'autres systèmes ou mis à niveau plus facilement qu'un système monolithique

L'un des outils d'aide à la décision qui peut être appliqué aux systèmes multi-agents est la théorie des jeux.

La théorie des jeux est un modèle de processus décisionnel interactif qui étudie le comportement égoïste ou rationnel des individus. On appelle ces derniers les joueurs et ils représentent l'élément le plus important dans un jeu. Un joueur peut représenter un seul individu ou un groupe d'individus prenant une décision. Après avoir défini l'ensemble des joueurs, on peut distinguer deux types de modèles : les jeux dont les éléments de base sont des actions des joueurs individuels ou ceux basés sur les actions jointes d'un groupe de joueurs. Les modèles du premier type sont appelés les jeux non coopératifs et ceux de second type, les jeux coopératifs [123]. Un jeu G se compose de trois éléments fondamentaux : les joueurs N ; les stratégies $\{S_i\}_{i \in N}$, et la fonction de gain $\{P_i\}_{i \in N}$. Chaque joueur choisit une stratégie $s_i \in S_i$ pour maximiser sa fonction de gain $P_i(s_{-i}, s_i)$, qui dépend non seulement de sa stratégie s_i , mais aussi des stratégies des autres joueurs s_{-i} .

L'un des plus importants concepts dans la théorie des jeux est : l'Équilibre de Nash. Il représente la stratégie stable et statique dans laquelle aucun joueur n'a envie de dévier unilatéralement, étant données les stratégies jouées par les autres joueurs. Les approches basées sur la théorie des jeux pour l'optimisation d'énergie des data centers convergent généralement vers un équilibre de Nash, cela nous conduit à avoir une solution optimale à la fois pour l'offre et la demande. Ainsi, la théorie des jeux représente une approche efficace pour faciliter la prise de décision d'une manière intelligente dans un

système Smart Grid – Cloud.

Dans ce contexte, Yanzhi et al. [113] proposent un jeu Stackelberg à deux niveaux entre le contrôleur du Smart Grid et le contrôleur du Cloud Computing. L'objectif du contrôleur du Smart Grid est d'effectuer l'équilibre de charge dans les bus électriques en se basant sur une politique de tarification définie. Tandis que l'objectif du contrôleur du Cloud est de maximiser le prix total obtenu en répondant aux demandes des utilisateurs. Une extension de ce travail est présentée dans [114], où les auteurs considèrent une structure de réseau électrique existante (la structure IEEE 24 bus) et l'intégration des sources d'énergies renouvelables dans le système. Par conséquent, la charge sur chaque bus d'alimentation dépend à la fois de la consommation d'énergie et de la production d'énergie renouvelable.

4.5.2 Optimisation convexe

Dans l'optimisation convexe, on considère des fonctions objectifs qui sont convexes. Généralement, les méthodes d'efficacités énergétiques dans les data centers visent à maximiser la fonction d'utilité ou minimiser la fonction de coût. Il faut noter que la fonction d'utilité est concave, alors que la fonction de coût est convexe. En outre, le problème de maximiser une fonction concave f peut être reformulé de la même manière lorsqu'on veut minimiser la fonction $-f$ qui est convexe [32].

Hao et al. [109] utilisent l'optimisation convexe pour modéliser le problème de réduction du coût énergétique des data centers dans le Cloud. Ils proposent une approche qui permet à un opérateur Cloud d'optimiser conjointement l'approvisionnement en énergie et l'allocation de la charge de travail des utilisateurs dans des data centers distribués. Chen et al. [24] considèrent un data center doté de dispositifs de stockage d'énergie. Leurs travaux visent à minimiser la combinaison linéaire entre le coût de l'électricité et le temps moyen de réponse des requêtes des utilisateurs. Des solutions basées sur des techniques d'optimisation convexes sont proposées pour résoudre le problème.

4.5.3 Processus de décision Markovien

Le processus de décision Markovien fait référence à la prise de décision séquentielle basée sur l'observation périodique ou continue des systèmes dynamiques de Markov. Les actions de ces derniers sont aléatoires, mais les probabilités de transition d'état présentent des propriétés de Markov.

Rao et al. [89] modélisent l'incertitude du prix d'énergie et la charge de travail traitée par les data centers du cloud en un processus de décision Markovien. Ils visent à réduire le coût de l'électricité tout en garantissant la qualité de service des utilisateurs. Zhang et al. [121] utilisent un processus de décision Markovien pour modéliser le problème d'approvisionnement de ressources des serveurs d'un data center. Ils proposent un algorithme d'autogestion basée sur l'apprentissage automatique. Leur proposition fournit les ressources aux serveurs d'une manière autonome et dynamique pour réduire le coût de l'électricité des data centers.

4.5.4 La logique floue

La logique floue est un sous-ensemble de la logique conventionnelle booléenne qui a été étendue pour prendre en charge le concept de vérité partielle : les valeurs de vérité qui sont entre « complètement vrai » et « complètement faux ». La logique floue [130] est un outil qui traite des informations incertaines, imprécises ou qualitatives dans des systèmes qui ne sont pas définis avec des modèles mathématiques formels. En théorie, l'ensemble flou indique l'appartenance progressive d'un élément à un ensemble. Soit A un ensemble flou, défini sur un univers X par une fonction d'appartenance $\mu_A : X \rightarrow [0, 1]$. μ_A indique le degré d'appartenance de x à A . Particulièrement, $\mu_A = 1$, indique l'appartenance complète, $\mu_A = 0$ indique la non-appartenance absolue et $0 < \mu_A < 1$ indique l'appartenance partielle (plus μ_A s'approche de 1, plus x appartient à A).

La logique floue couvre un large éventail de conditions de fonctionnement. En outre, les algorithmes à base de logique floue peuvent traiter les incertitudes et sont robustes aux environnements à forte mutation [130]. De plus, le processus d'inférence est simple, comparé à d'autres systèmes [78], ce qui permettra d'économiser de l'énergie. Dans notre cas, cette fonctionnalité est très intéressante, car nous considérons un système temps réel où les requêtes des utilisateurs doivent être exécutées dans des brefs délais tout en minimisant la consommation d'énergie.

Chopra et al. [84] ont mis en œuvre un système de surveillance et de contrôle de l'électricité en utilisant la logique floue et le Cloud Computing. Ce système est responsable de l'ajustement automatique du temps de travail des appareils électriques. Le système prend l'humidité et la température comme entrées et calcule le temps de travail des appareils en sortie. Les résultats ont montré que le système proposé permet d'économiser de l'énergie et aide à avoir une utilisation énergétiquement efficace.

4.6 Discussion et conclusion

Dans cette thèse, nous portons notre intérêt sur les techniques prenant en considération un ensemble de data centers interagissant avec le smart grid. Le tableau 4.1 présente une étude comparative entre les travaux susmentionnés dans cette catégorie. La comparaison est effectuée selon quatre critères : réduction des coûts énergétiques, réduction des émissions de gaz à effet de serre, réduction de la charge du réseau électrique et intégration des énergies renouvelables. Nous constatons que la plupart des solutions proposées dans un système smart grid – cloud est appliquée du côté des fournisseurs de services cloud et nous notons que la prise en considération du prix de l'énergie et l'utilisation des énergies renouvelables demeurent deux facteurs principaux de l'efficacité énergétique dans les data centers. Cependant, peu de travaux ont examiné le rôle du smart grid dans la gestion de l'allocation d'énergie aux data centers du cloud. En outre, les approches proposées ne traitent pas l'impact que l'énorme consommation d'énergie des data centers peut avoir sur le smart grid.

C'est sur la base de ces limitations que se portent les premières contributions de cette thèse. Nous avons proposé, dans un premier lieu, une approche basée sur un jeu non-coopératif qui vise à répartir

TABLE 4.1 – Tableau comparatif des travaux qui se portent sur l’efficacité énergétique des data centers dans un système smart grid

Travaux de recherche	Métriques considérées				
	Coût d’énergie	Émissions de CO2	Taux de charge électrique	Énergies renouvelables	Architecture Microgrid - Cloud
Gu et al [47] [48]	X	X		X	
Huang et al [110]	X		X		
Wang et al [109]	X		X		
Kiani et al [69]	X			X	
Deng et al [34]	X	X		X	
Kawsar et al [51]	X	X		X	
Lu et al [75]	X		X		
Yuan et al [129]	X	X		X	
Yu et al [127] [128]	X	X		X	X
Jiang et al [126]	X	X		X	X
Ding et al [35]	X	X		X	X
Baris et al [4]			X		X
Benblidia et al [14] (chapitre 5)			X		
Benblidia et al (chapitre 6)		X	X	X	
Benblidia et al [15]		X	X	X	X
Benblidia et al (chapitre 7)	X	X	X	X	X

équitablement l’énergie entre les data centers dans un système smart grid - cloud [14] (chapitre 5). Ensuite, afin de réduire les émissions de CO2 que causent les data centers, nous avons considéré les énergies renouvelables et utiliser notre jeu non-coopératif pour éviter un déséquilibre d’attribution d’énergie entre les data centers et essayer de répartir équitablement l’énergie entre eux (chapitre 6).

Généralement, les travaux de recherche utilisant des architectures à base de microgrids ont pour objectif de réduire le coût énergétique et l’émission de carbone des data centers du cloud. Cependant, ils ne présentent pas une analyse détaillée de la manière dont leurs approches peuvent affecter le réseau électrique, car ils ne considèrent pas le fait qu’un smart grid ait une quantité d’énergie limitée à fournir.

Compte tenu de ces manquements, nous avons proposé, dans notre travail [15], une architecture smart microgrid cloud et étudié le problème de la répartition d’énergie entre les data centers du cloud. Mais nous n’avons pas pris en compte le coût d’énergie des microgrids lors de l’allocation d’énergie aux data centers. Ce qui nous a amené, dans le travail qui suit (chapitre 7), à proposer une architecture microgrid - cloud, où les microgrids peuvent échanger des informations et de l’énergie entre eux. Nous avons donc proposé une approche à deux étapes qui, dans un premier temps répartit l’énergie équitablement entre les data centers, et dans un second temps, minimise le coût énergétique des microgrids.

Grâce à cette étude de l’état de l’art, on peut constater qu’il existe peu de travaux de recherche qui ont étudié l’efficacité énergétique des équipements IT dans un data center du point de vue du smart grid. La majorité des travaux réalisés à ce sujet se concentre sur l’efficacité énergétique d’un seul élément IT sans tirer profit des fonctionnalités qu’apporte le smart grid.

Un jeu non-coopératif pour une répartition équitale de l'énergie entre les data centers

Sommaire

5.1	Introduction et motivations	58
5.2	Architecture et modélisation du problème	58
5.2.1	Architecture du système	58
5.2.2	Modélisation de notre jeu non-coopératif	60
5.2.3	Solution optimale du jeu	62
5.3	Résultats des simulations	62
5.3.1	Paramètres des simulations	63
5.3.2	Évaluation des performances de notre approche	63
5.4	Conclusion	66

5.1 Introduction et motivations

Étant la source d'énergie de multiples systèmes tels que les véhicules électriques, les maisons intelligentes, etc. [76], le Smart Grid prend également en charge l'énorme demande en électricité des data centers du cloud. Compte tenu des milliers de serveurs, d'équipements de stockage, d'installations de refroidissement et de transformateurs de puissance que possède un data center, le secteur des data centers représenterait 1.4% de la consommation mondiale d'électricité [120].

De plus, les data centers à l'heure actuelle procurent de grandes quantités d'énergie du smart grid, et cela sans tenir compte de la quantité d'énergie limitée, désignée à alimenter les data centers. D'une part, cela peut priver un data center, dont le besoin en énergie est plus indispensable que d'autres data centers, de disposer de la quantité d'énergie suffisante pour remplir ses fonctions. D'autre part, cela peut surcharger les centrales électriques et causer de graves dommages au réseau électrique.

Notre première contribution vise à surmonter ce problème. Ainsi, nous modélisons le comportement égoïste des data centers en tant que jeu non coopératif dans lequel chaque data center est représenté par un joueur. L'idée de base de notre proposition est d'attribuer la quantité optimale d'électricité à chaque data center afin d'obtenir une répartition équitable entre les data centers. Cette attribution prend en compte trois métriques : la charge des data centers, le nombre d'applications critiques en cours d'exécution au sein d'un data center et le nombre d'applications critiques en cours d'exécution au sein de tous les data centers.

5.2 Architecture et modélisation du problème

Nous présentons dans cette section l'architecture de notre système, la modélisation de la problématique. Nous décrivons aussi notre jeu non-coopératif ainsi que la solution optimale du jeu.

5.2.1 Architecture du système

Dans notre système, le fournisseur du réseau électrique est chargé de répondre aux demandes des consommateurs en leur fournissant de l'électricité. L'alimentation en électricité est fournie par plusieurs centrales qui communiquent avec les data centers du cloud (communication bidirectionnelle). Par conséquent, le fournisseur d'énergie aura une visibilité sur le comportement énergétique de chaque data center. L'architecture de notre système smart grid - cloud est illustrée dans la figure 5.1.

Nous considérons un ensemble de data centers géographiquement répartis et gérés par un seul fournisseur de services cloud. Notons d_i un data center i ; $\forall i \in I$ avec $I = \{1, 2, \dots, i, \dots, n\}$. Chaque data center d_i est alimenté par une centrale électrique dédiée et toutes les centrales électriques sont gérées par le même contrôleur smart grid et nous supposons que :

1. x_i^{Max} est la quantité maximale d'électricité qu'un data center d_i peut acheter.

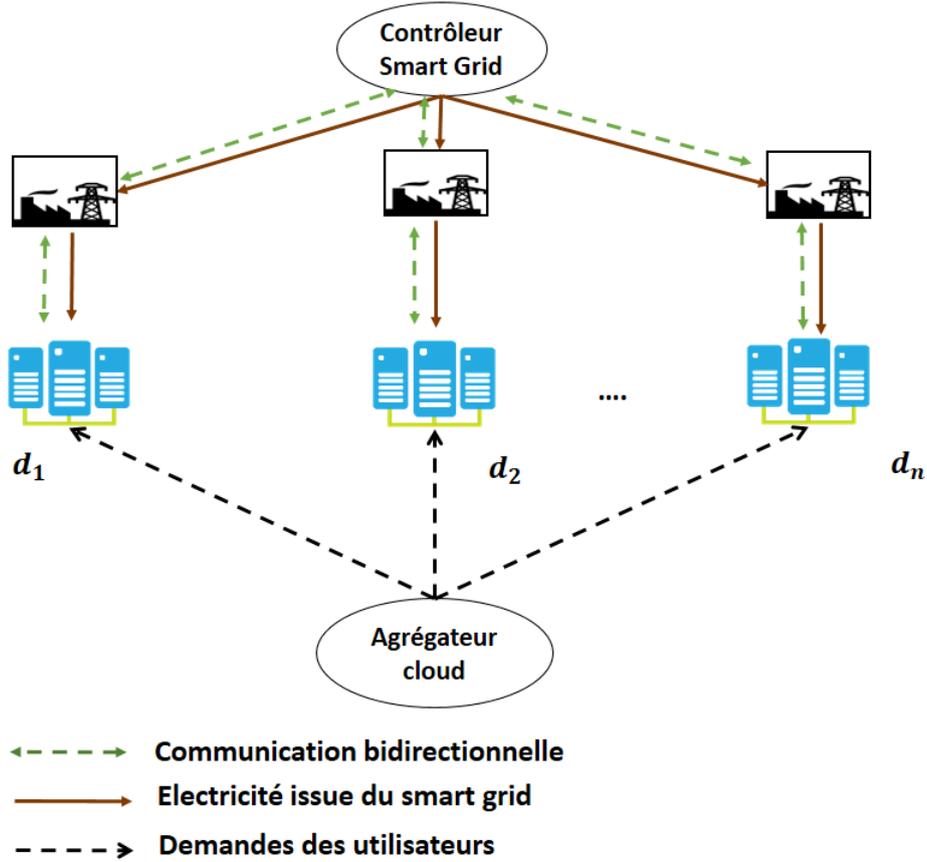


FIGURE 5.1 – Architecture d'un système smart grid - cloud

2. Chaque data center a une énergie stockée, nous notons par Soc_i le niveau d'énergie stockée d'un data center d_i .

Chaque data center d_i a trois priorités $Dp_i, Ap_i, Tp_i \in [0, 1] \forall i \in I$, définies comme suit :

- Cp_i : La priorité de la charge d'un data center : reflète la priorité d'un data center en termes de son nombre de serveurs actifs et la valeur du niveau de charge de son énergie stockée.
- Ap_i : La priorité des applications critiques : indique la priorité d'un data center en termes de nombre d'applications critiques en cours d'exécution, par rapport au nombre total d'applications en cours d'exécution.
- Tp_i : La priorité des applications critiques totale : reflète la priorité d'un data center en termes de nombre d'applications critiques en cours d'exécution, par rapport au nombre total d'applications critiques en cours d'exécution dans tous les data centers.

En se basant sur les suppositions de [48] et [109], nous considérons, dans ce travail, que les data centers visent à obtenir le maximum d'énergie du smart grid, afin de stocker leur énergie et d'exécuter le maximum d'applications des utilisateurs. Cela permettra d'améliorer les bénéfices du fournisseur de service cloud. De plus, chaque data center essaie d'augmenter son état de charge d'énergie et demande de l'énergie de manière égoïste sans tenir compte de l'énergie limitée prévue pour servir les autres data

centers. Cela peut surcharger les centrales électriques et causer de sérieux dégâts au réseau électrique.

5.2.2 Modélisation de notre jeu non-coopératif

Nous avons utilisé la théorie des jeux non coopératifs pour modéliser la répartition de l'électricité entre les data center. Ainsi, nous considérons le jeu : $G = (I, S_i, gain_i)_{i \in I}$ où :

1. I : représente un ensemble de joueurs (data centers), $t_1, \dots, t_j, \dots, t_n$ où n est le nombre de data centers participants au jeu. .
2. S_i ; $\forall i \in I$: Stratégies : reflète les actions que le joueur d_i peut entreprendre à n'importe quel stade du jeu. Un joueur peut demander une quantité d'électricité comprise entre un minimum de zéro et un maximum de x^{Max} . Par conséquent, $S_i = [0, x_i^{Max}]$ et le profil stratégique de tous les joueurs est $S = \prod_{i=1}^n S_i = [0, x_1^{Max}] \times \dots \times [0, x_i^{Max}] \times \dots \times [0, x_n^{Max}]$.
3. $gain_i : S_i \rightarrow \mathbb{R}$ représente la fonction de gain du joueur $d_i \forall i \in I$. Chaque joueur vise à maximiser sa fonction de gain $gain$, afin d'augmenter son profit x_i ;
 $0 \leq x_i \leq x_i^{Max}, \forall i \in I$.

Nous modélisons la fonction de gain de notre jeu de manière à refléter quatre fonctions principales :

- (i) La première fonction calcule la demande du data center pour avoir une grande quantité d'énergie (fonction d'utilité). Les trois autres fonctions sont des fonctions de coût des trois priorités qui calculent respectivement (ii) le coût de la charge du data center (iii) le coût des applications critiques, et (iiii) le coût total des applications critiques.

La fonction de gain comprend donc les quatre fonctions suivantes :

1. Fonction d'utilité : dans notre approche, la fonction d'utilité a été modélisée de sorte que chaque joueur obtienne plus de profits lorsqu'il augmente la valeur de sa demande d'énergie. Parmi les fonctions d'utilité couramment utilisées dans les travaux de recherches sur les réseaux [111], nous avons choisi de modéliser notre fonction d'utilité à l'aide de la fonction logarithmique. Donc, nous définissons la fonction d'utilité d'un joueur d_i comme suit :

$$\gamma_i(x_i) = \log(x_i + 1), \forall i \in I \quad (5.1)$$

où x_i est la quantité d'électricité qu'un joueur d_i demande.

2. La fonction de coût de la charge du data center : cette fonction définit la pénalité que le joueur doit payer en fonction de sa demande de quantité d'énergie x_i et de sa priorité de charge Dp_i . Dans notre approche, la priorité Dp_i est calculée en fonction des serveurs actifs As_i et de l'état de charge Soc_i du joueur d_i comme suit :

$$Dp_i = 1 - \frac{1}{As_i + Soc_i} \quad (5.2)$$

Nous avons donc la fonction de coût suivante :

$$\rho_i(x_i, Dp_i) = x_i \times Dp_i \quad (5.3)$$

3. La fonction du coût des applications critiques : dans cette fonction de coût, nous prenons en considération les applications critiques en cours d'exécution d'un data center d_i . Nous désignons, dans notre approche, l'application critique comme une application en temps réel. La fonction est calculée en fonction de la quantité d'énergie demandée par un data center x_i et sa priorité d'applications critiques Ap_i . Nous calculons la priorité Ap_i comme étant le rapport entre le nombre d'applications critiques en cours d'exécution notée $Critical_apps_i$ et le nombre total des applications notées $Total_apps_i$. Ainsi, nous définissons la fonction comme suit :

$$\sigma_i(x_i, Ap_i) = x_i \times Ap_i \quad (5.4)$$

Où :

$$Ap_i = 1 - \frac{Critical_apps_i}{Total_apps_i} \quad (5.5)$$

4. La fonction du coût total des applications critiques : cette fonction de coût est en fonction de la quantité d'électricité demandée par le data center x_i et sa priorité des applications critiques totale Tp_i . La priorité Tp_i est le rapport entre le nombre des applications critiques en cours d'exécution dans un data center et la somme de toutes les applications critiques en cours d'exécution dans tous les data centers participants au jeu. Nous avons donc la fonction de coût suivante :

$$\tau_i(x_i, Tp_i) = x_i \times Tp_i \quad (5.6)$$

Où :

$$Tp_i = 1 - \frac{Critical_apps_i}{\sum_{i=0}^n Critical_apps_i} \quad (5.7)$$

Après avoir défini les quatre fonctions d'un joueur d_i , nous donnons la fonction de gain ci-après :

$$gain_i(x_i, x_{-i}) = \alpha_i \gamma_i(x_i) + \beta_i \rho_i(x_i, Dp_i) + \omega_i \sigma_i(x_i, Ap_i) + \psi_i \tau_i(x_i, Tp_i) \quad (5.8)$$

Où :

- $x_{-i} = \{x_k\} (k \in I)$ et $i \neq k$ est le vecteur de la quantité d'électricité demandée (stratégies) de tous les joueurs sauf le joueur d_i .
- $s = (x_i, x_{-i}) \in S$ est le profil de stratégie.
- $\alpha_i, \beta_i, \omega_i$ et ψ_i sont les coefficients de préférence des joueurs des fonctions $\gamma_i, \rho_i, \sigma_i$ et τ_i , respectivement. Ces paramètres sont définis pour satisfaire l'objectif et les exigences du système, par exemple, dans le cas où la valeur ω_i est plus grande, un data center qui a un ratio d'applications critiques élevé obtient plus de profit (électricité) qu'un data center qui a un ratio d'applications critiques faible et vice versa.

5.2.3 Solution optimale du jeu

Nous calculons dans cette sous-section la solution optimale de notre jeu x_i^* maximisant son profit (fonction de gain). Pour ce faire, nous modélisons un problème d'optimisation non linéaire (F) comme suit :

$$\begin{aligned} & \underset{i}{\text{Maximiser}} && \text{gain}_i(x_i, x_{-i}) \\ \text{S.C :} & && x_i \geq 0, \\ & && x_i - x_i^{\max} \leq 0, \forall i \in I. \end{aligned}$$

Pour résoudre le problème (\mathbf{F}), nous avons utilisé la méthode des multiplicateurs de Lagrange pour maximiser notre fonction qui est soumise à deux contraintes d'inégalité. Nous introduisons deux multiplicateurs de Lagrange, λ_i et ℓ_i , et étudions la fonction de Lagrange définie par

$$L_i(x_i, \lambda_i, \ell_i) = \text{gain}_i(x_i, x_{-i}) + \lambda_i x_i + \ell_i (x_i^{\max} - x_i) \quad (5.9)$$

Ainsi, notre fonction de Lagrange satisfait les conditions KKT suivantes pour chaque joueur d_i

$$\begin{aligned} \lambda_i, \ell_i &\geq 0 \\ x_i &\geq 0 \\ x_i^{\max} - x_i &\geq 0 \\ \nabla_{x_i} \text{gain}_i(x_i, x_{-i}) + \lambda_i \nabla_{x_i} x_i + \ell_i \nabla_{x_j} (x_i^{\max} - x_i) &= 0 \\ \lambda_i x_i + \ell_i (x_i^{\max} - x_i) &= 0 \end{aligned}$$

Après avoir résolu le problème (\mathbf{F}), nous déterminons la quantité d'électricité optimale à assigner (x_i^*) au joueur d_i , $\forall i \in I$, comme suit :

$$x_i^* = \begin{cases} 0 & \text{si condition 1} \\ x_i^{\max} & \text{si condition 2} \\ \frac{\alpha_i}{\beta_i \rho_i(x_i, Dp_i) + \omega_i \sigma_i(x_i, Ap_i) + \psi_i \tau_i(x_i, Tp_i)} - 1 & \text{sinon} \end{cases} \quad (5.10)$$

Où les conditions 1 et 2, respectivement, sont :

$$\beta_i \rho_i(x_i, Dp_i) + \omega_i \sigma_i(x_i, Ap_i) + \psi_i \tau_i(x_i, Tp_i) \geq \alpha_i \quad (5.11)$$

$$\beta_i \rho_i(x_i, Dp_i) + \omega_i \sigma_i(x_i, Ap_i) + \psi_i \tau_i(x_i, Tp_i) \leq \frac{\alpha_i}{x_i^{\max} + 1} \quad (5.12)$$

5.3 Résultats des simulations

Dans cette section, nous présentons les résultats de nos simulations et démontrons l'efficacité de notre approche de répartition d'énergie basée sur un jeu non-coopératif.

5.3.1 Paramètres des simulations

Pour valider notre proposition, nous la comparons avec la méthode traditionnelle dans laquelle les data centers demandent périodiquement de l'énergie au contrôleur du smart grid, sans tenir compte de sa quantité limitée d'énergie destinée à servir les data centers.

Nous considérons quatre data centers géographiquement répartis et qui sont connectés au même contrôleur smart grid. Nous considérons que chaque data center a au moins un serveur $5MW$ comme minimum d'énergie stockée (Soc). Nous mesurons la demande d'énergie de chaque data center à partir du contrôleur smart grid toutes les 2 heures. Nous effectuons nos simulations sur deux périodes différentes : les heures creuses et les heures de pointe, avec un horizon temporel de 24 heures. Les heures creuses sont la période où les demandes d'électricité et de service cloud sont faibles (de $22h$ à $7h$) et les heures de pointe sont la période où nous avons une demande d'électricité énorme et où les data centers exécutent un grand nombre d'applications critiques (de $7h$ à $22h$). Dans notre schéma, nous avons défini les paramètres $\alpha_i = 6, \beta_i = 2, 7, \omega_i = 2, \psi_i = 2$ et $x_i^{max} = 5MW, i \in I$.

5.3.2 Évaluation des performances de notre approche

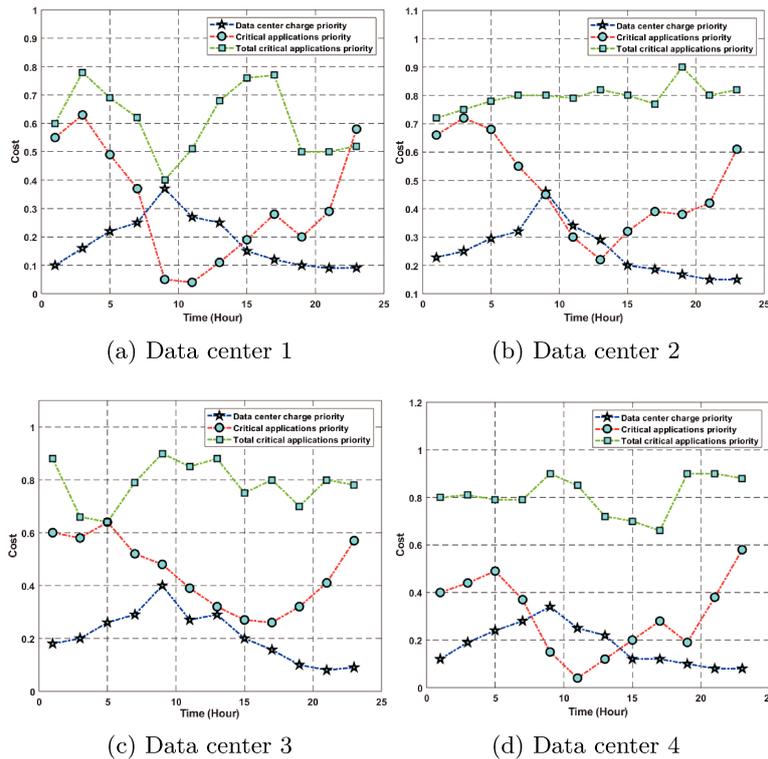


FIGURE 5.2 – Variation des coûts des data centers

La figure 5.2 montre les variations des coûts de chaque data center sur une journée (le coût de la priorité de charge du data center, la priorité des applications critiques et la priorité totale des applications critiques) en utilisant les équations (5.2), (5.5) et (5.7), respectivement. Nous pouvons observer que dans chaque data center, la priorité de charge du data center est faible aux heures creuses et élevée

aux heures de pointe. Cela peut s'expliquer par le fait que le data center n'a pas beaucoup de serveurs actifs et d'énergie stockée aux heures creuses, contrairement aux heures de pointe où le data center a chargé son énergie et a un nombre élevé de serveurs actifs. D'autre part, les priorités des applications critiques et des applications critiques totales sont élevées aux heures creuses et faibles aux heures de pointe. Cela signifie que les applications critiques par rapport à l'ensemble des applications en cours d'exécution dans un data center sont beaucoup plus nombreuses aux heures de pointe qu'aux heures creuses.

La figure 5.3 illustre la répartition d'énergie des quatre data centers en une heure particulière (12h) en utilisant notre approche de répartition d'énergie basée sur un jeu non-coopératif. Nous pouvons voir que notre approche prévoit une répartition équitable en fonction des coûts des priorités de chaque data center.

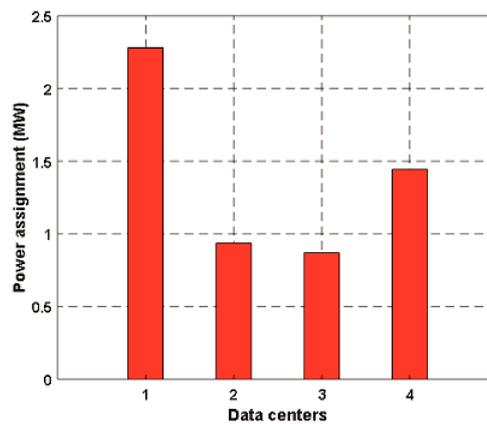


FIGURE 5.3 – La quantité d'électricité allouée en utilisant notre approche basée sur un jeu non-coopératif

La figure 5.4 compare la variation de l'énergie allouée à un data center choisi aléatoirement, en utilisant notre jeu non-coopératif et l'approche traditionnelle. Nous observons que l'approche traditionnelle est toujours fournie avec la quantité totale d'énergie demandée, quelles que soient les applications critiques ou la charge énergétique qu'un data center peut avoir. Cela peut surcharger le réseau et la centrale électrique. D'autre part, la quantité d'énergie assignée en utilisant notre approche varie dans le temps et peut diminuer jusqu'à 1,27MW. Cela se produit généralement pendant les heures creuses (par exemple à 3 heures du matin dans la figure 5.4-a) lorsque le data center est en train de charger son énergie et n'a pas beaucoup d'applications en cours d'exécution. De plus, il peut atteindre la valeur maximale de 5MW comme à 11h et 19h. Les résultats présentés sont dus à la théorie des jeux utilisée dans notre approche qui attribue la quantité suffisante d'énergie en tenant compte (i) des serveurs actifs et de l'état de charge de chaque data center, (ii) des applications critiques en cours d'exécution dans le data center et (iii) des applications critiques en cours d'exécution dans tous les data centers, comme le montrent respectivement la figure 5.4-b, la figure 5.4-c et la figure 5.4-d.

La figure 5.4-b montre la quantité d'énergie assignée en fonction de la priorité de charge du data center

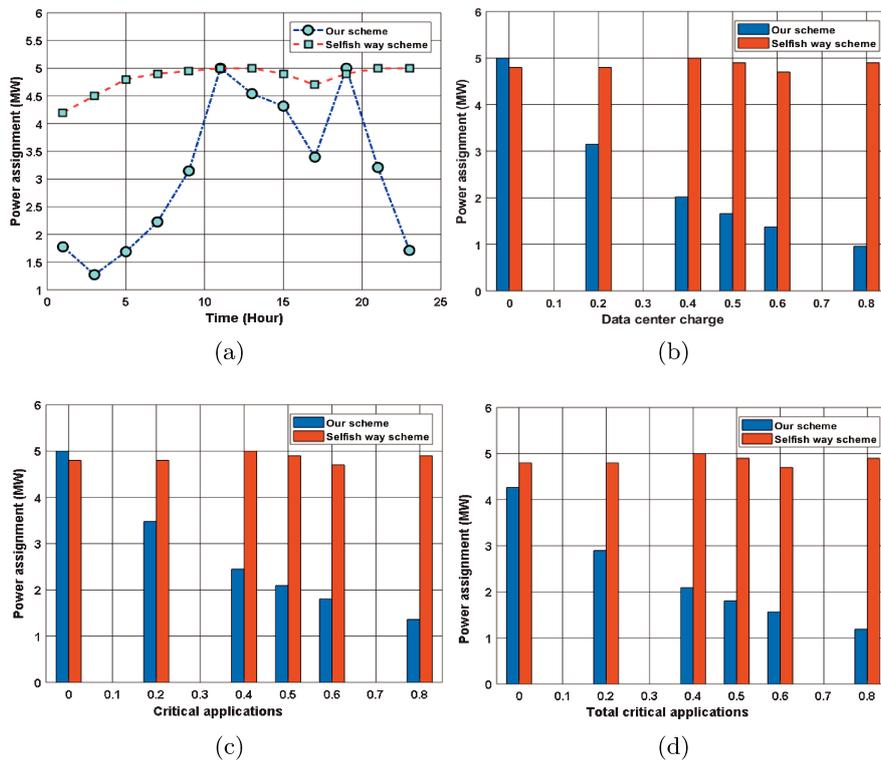


FIGURE 5.4 – Comparaison des performances entre notre approche et l’approche traditionnelle : (A) Allocation d’énergie, (B) Allocation d’énergie Vs le coût de charge d’un data center , (C) Allocation d’énergie VS le coût des applications critiques, (D) Allocation d’énergie Vs le coût des applications critiques totale

TABLE 5.1 – L’efficacité de notre approche de répartition d’énergie basée sur un jeu non-coopératif

Data center \ Approche	Traditionnelle (MW)	Notre approche (MW)	Efficacité
DC 01	5	2.28	54.39 %
DC 02	4.7	0.93	80.07 %
DC 03	4.9	0.86	62.32 %
DC 04	4.75	1.44	69.6 %

en termes de serveurs actifs et d’état de charge. Nous pouvons observer que la quantité d’énergie assignée dans notre approche diminue au fur et à mesure que la priorité diminue. Cependant, l’approche traditionnelle est toujours desservit par le smart grid d’une manière stable sans tenir compte d’une quelconque priorité. En outre, les figures 5.4-c et 5.4-d montrent également que la priorité des applications critiques et la priorité totale des applications critiques, respectivement, ont une relation inverse avec l’allocation d’énergie. Un joueur a plus d’énergie lorsque ses applications critiques augmentent et vice versa.

Nous mesurons l’efficacité dans notre jeu par le pourcentage de charge électrique qui a été réduit sur le réseau électrique. Le tableau 5.1 montre l’efficacité de notre approche, où notre jeu peut réduire la charge sur le réseau électrique avec un pourcentage de 81%.

5.4 Conclusion

Nous avons étudié, dans ce chapitre, l’interaction des data centers du cloud avec le smart grid pour avoir une répartition équitable de l’énergie. Nous avons modélisé l’attribution d’énergie aux data centers comme un jeu non coopératif. Ensuite, nous avons calculé la quantité optimale d’énergie à attribuer à chaque data center. Les résultats de la simulation ont montré que notre approche peut réduire de manière significative la charge électrique sur le réseau électrique, par rapport à l’approche traditionnelle.

Plusieurs paramètres n’ont pas été pris en compte dans notre architecture smart grid - cloud proposée dans ce travail. Une extension de ce travail considérant les énergies renouvelables dans notre système sera le sujet du prochain chapitre.

Chapitre 6

Répartition efficace de l'énergie entre les data centers alimentés par des énergies renouvelables

Sommaire

6.1	Introduction et motivations	68
6.2	Architecture et modélisation du problème	69
6.2.1	Architecture du système	69
6.2.2	Description du jeu non-coopératif	71
6.2.3	Notre solution optimale	75
6.3	Simulations et résultats	76
6.3.1	Paramètres des simulations	76
6.3.2	La répartition d'énergie entre les data centers	77
6.3.3	Les performances de notre approche en termes de taux de réduction de charge et d'émissions de dioxyde de carbone	80
6.4	Conclusion	80

6.1 Introduction et motivations

L'impact des data centers sur l'environnement est un sujet dont on parle assez souvent. Et à mesure que l'intelligence artificielle et les systèmes connectés se développent, la demande de stockage de données augmente également. Les entreprises comprennent que les data centers ont une empreinte carbone extrêmement importante, c'est pourquoi certaines des plus grandes entreprises cherchent à utiliser des énergies renouvelables propres.

L'engagement des entreprises technologiques en faveur des énergies renouvelables se manifeste par la quantité d'énergie renouvelable qu'elles achètent. La Renewable Energy Buyer's Alliance a classé les dix premiers acheteurs d'énergie renouvelable aux États-Unis en 2019, avec de nombreuses entreprises technologiques dans la liste (figure 6.1).



FIGURE 6.1 – Les 10 plus gros acheteurs d'énergie des États-Unis en 2019 (GW) [100]

AWS, Google, Facebook et Microsoft ont tous récemment annoncé des achats d'énergie renouvelable à grande échelle. Google a annoncé en 2017 qu'il a atteint 100% d'énergie renouvelable dans toutes ses opérations, y compris les data centers. Il est important de noter que Google et les autres acteurs majeurs dans ce domaine comptent sur les crédits d'énergie renouvelable (Renewable Energy Credit - REC) pour compenser leur utilisation des énergies fossiles et revendiquer une utilisation à 100% des énergies renouvelables tout en étant connectés à des réseaux qui dépendent des énergies fossiles pour la production d'électricité [6].

AWS, de loin le plus grand fournisseur de cloud computing, avec plus d'un tiers du marché, s'est engagé, à long terme, à utiliser 100% d'énergies renouvelables et a atteint 50% dans toute l'entreprise d'Amazon en 2018 [8].

Microsoft est responsable de 16 millions de tonnes d'émissions de gaz à effet de serre par an, ce qui inclut ses data centers, la fabrication de composants électroniques et la consommation d'énergie des utilisateurs de consoles de jeux Xbox à la maison [79].

Les gaz à effet de serre comme le dioxyde de carbone peuvent provoquer des changements climatiques et créer une pollution de l'air contribuant aux maladies respiratoires. Ils peuvent également provoquer

des conditions climatiques extrêmes, une augmentation des incendies de forêt et des perturbations de l’approvisionnement alimentaire [99].

Dans l’optique de surmonter les problématiques évoquées précédemment, nous nous concentrons dans ce chapitre sur la gestion de l’énergie des data centers du cloud par le smart grid. Contrairement à notre travail présenté dans le chapitre précédent, qui considère le smart grid comme la seule source d’énergie pour les data centers, dans ce travail, nous considérons un ensemble de data centers géographiquement dispersés, alimentés à la fois par un contrôleur smart grid et par leur propre énergie renouvelable. Notre approche répartit de manière optimale l’énergie entre les data centers en fonction de leurs priorités. Notre objectif est d’éviter un déséquilibre d’attribution d’énergie entre les data centers et d’essayer de répartir équitablement l’énergie entre eux. Pour ce faire, nous définissons cinq priorités de data centers en fonction de leur charge de travail, du nombre d’applications critiques en cours d’exécution et de leur utilisation d’énergie renouvelable. Nous modélisons la répartition d’énergie du smart grid vers les data centers du cloud comme un jeu non coopératif, ensuite nous calculons la quantité optimale d’énergie à fournir à chaque data center.

Afin d’évaluer notre proposition, nous l’avons comparée à une approche existante de minimisation de l’énergie et nous avons prouvé l’équité de notre approche de répartition d’énergie en utilisant l’indice d’équité de Jain.

6.2 Architecture et modélisation du problème

Nous présentons dans cette section notre architecture qui considère les énergies renouvelables dans un système smart grid - cloud. Nous détaillons la modélisation des cinq priorités des data centers, ainsi que la fonction de gain du jeu non-coopératif. Enfin, nous déterminons la solution optimale de notre problème.

6.2.1 Architecture du système

Comme l’illustre la figure 6.2, nous considérons, dans notre système, un contrôleur de smart grid qui approvisionne en énergie ses différents consommateurs. Le contrôleur de smart grid est connecté à des centrales électriques qui fournissent de l’énergie à des data centers géographiquement répartis. Les data centers sont gérés par différents fournisseurs de cloud computing qui collectent les demandes des clients et les assignent à leurs data centers respectifs. Afin de remplir leurs fonctions et d’exécuter les demandes de leurs clients, les data centers consomment de l’électricité provenant de deux sources : le réseau principal et leur propre énergie renouvelable autoproduite par des turbines éoliennes. Nous utilisons les turbines éoliennes comme seules sources d’énergie renouvelable en raison de leur efficacité en termes de production d’électricité et d’émission de carbone par rapport aux panneaux solaires [116] [45]. Pour faire face à la nature intermittente des énergies renouvelables, les data centers sont équipés de batteries pour stocker leur énergie.

Le smart grid fournit de l’électricité à un ensemble de data centers géographiquement répartis.

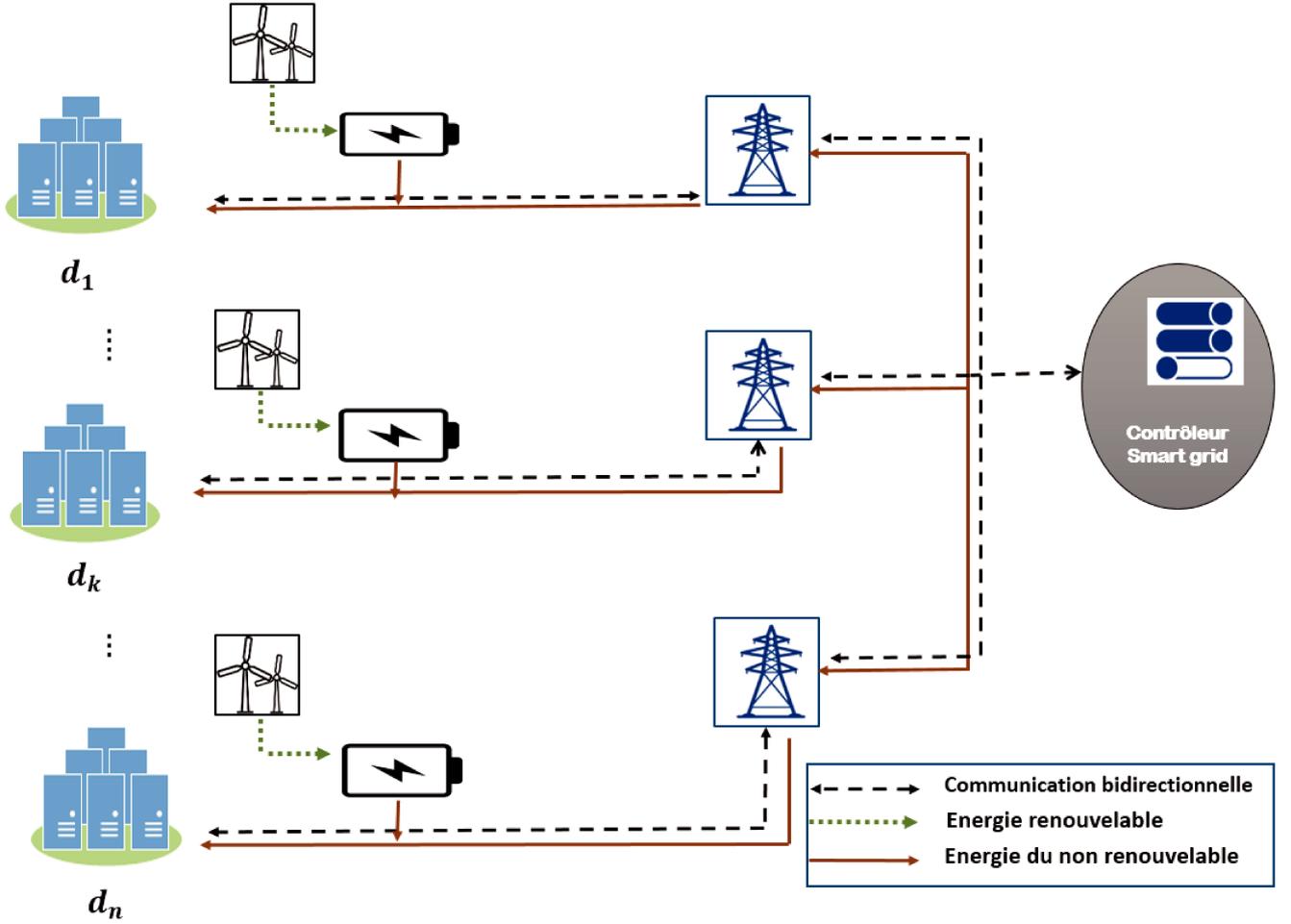


FIGURE 6.2 – L'énergie renouvelable dans une architecture smart grid - cloud

Nous indiquons par d_k un data center k ; $\forall k \in K$ avec $K = \{1, 2, \dots, k, \dots, n\}$. Nous considérons dans notre système une période de répartition d'énergie de $T = \{1, \dots, t, \dots, \tau\}$ et nous divisons le temps quotidien en $\tau = 24$. Nous supposons que chaque data center d_k , dans chaque tranche de temps $t \in T$, a un nombre de serveurs actifs s_k^t et un état de charge énergétique SEC_k^t stocké sur sa batterie. Nous supposons que le smart grid peut alimenter chaque data center d_k avec une quantité maximale d'électricité q_k^{max} . De plus, chaque data center d_k a cinq priorités $DCp_k^t, CAp_k^t, TAp_k^t, GECp_k^t$ et $REPP_k^t \in [0, 1] \forall k \in K, \forall t \in T$ défini comme suit [73] :

1. La priorité de charge du data center DCp_k^t : reflète dans chaque instant $t \in T$ la priorité d'un data center en termes de son nombre de serveurs actifs s_k^t et la valeur du niveau de charge de son énergie stockée SEC_k^t .
2. La priorité des applications critiques CAp_k^t : représente dans chaque instant $t \in T$ le rapport entre le nombre d'applications critiques en cours d'exécution dans un data center, et le nombre total d'applications en cours d'exécution. Nous considérons dans notre système les applications en temps réel comme les applications critiques.
3. La priorité des applications critiques totale TAp_k^t : représente dans chaque instant $t \in T$ la priorité du data center, calculée par le rapport entre son nombre d'applications critiques en

cours d'exécution et le total des applications critiques en cours d'exécution dans tous les data centers présents dans le système.

4. Priorité du coefficient d'énergie verte $GECp_k^t$: reflète dans chaque intervalle de temps $t \in T$ la consommation d'énergie verte d'un data center par rapport à sa consommation d'énergie totale.
5. Priorité de production d'énergie renouvelable $REPP_k^t$: correspond à la priorité d'un data center en termes de production d'énergie renouvelable par rapport à la production d'énergie renouvelable de tous les data centers présents dans le système.

Afin d'améliorer leurs profits, les fournisseurs de services cloud, s'efforcent, via leurs data centers respectives, à exécuter le maximum d'applications issues des clients. Pour ce faire, les data centers sont en concurrence pour obtenir le maximum d'énergie du smart grid sans tenir compte de sa capacité de production d'énergie. Ce comportement irrationnel de demande d'énergie peut endommager les lignes de transmission ou les systèmes de distribution d'énergie.

6.2.2 Description du jeu non-coopératif

Nous modélisons l'allocation d'électricité aux data centers comme un jeu non coopératif. Nous considérons le jeu : $G = (K, St_k, Gn_k)_{k \in K}$ où :

- K : est un ensemble de joueurs de data centers, $d_1, \dots, d_k, \dots, d_n$ où n est le nombre de joueurs du jeu.
- $St_k; \forall k \in K$: représente les stratégies ou les actions que le joueur d_k peut avoir à n'importe quel stade du jeu. Dans notre cas, l'action est la quantité d'électricité qu'un data center demande, nous indiquons cette quantité par q_k et sa valeur maximale par q_k^{max} . Ainsi, la quantité d'électricité demandée peut varier entre zéro et q_k^{max} . Ainsi, $St_k = [0, q_k^{max}]$ et le profil stratégique de tous les joueurs est $St = \prod_{k=1}^n St_k = [0, q_1^{max}] \times \dots \times [0, q_k^{max}] \times \dots \times [0, q_n^{max}]$.
- $Gn_k : St_k \rightarrow \mathbb{R}$ représente la fonction de gain du joueur d_k pour tout $k \in K$. L'objectif de chaque joueur est de maximiser sa fonction de gain, $Gn, 0 \leq Gn_k \leq q_k^{max}$ pour tout $k \in K$, afin d'augmenter ses profits.

Afin d'assurer une répartition équitable de l'énergie aux data centers, le contrôleur de smart grid répartit l'énergie en fonction des cinq priorités des data centers préalablement définies. Par conséquent, nous avons conçu la fonction de gain comprenant six fonctions : une fonction d'utilité qui calcule la demande d'énergie des data centers pour avoir une quantité élevée d'énergie, et une fonction de coût pour chaque priorité de data center. Nous définissons les six fonctions impliquées dans la fonction de gain de notre jeu comme suit :

1. Fonction d'utilité $\gamma_k(q_k^t)$: nous modélisons la fonction d'utilité de manière à ce que plus le data center augmente sa valeur de quantité d'énergie, plus il a de profit. Nous choisissons de concevoir cette fonction avec la fonction logarithmique, qui est strictement concave. Ainsi, nous modélisons la fonction d'utilité d'un joueur d_k comme suit :

$$\gamma_k(q_k^t) = \log(q_k^t + 1), \forall k \in K, \forall t \in T \quad (6.1)$$

TABLE 6.1 – Nomenclature

K	Ensemble de data centers
T	Période de l'horizon temporel
data center d_k	k
q_k^{max}	Quantité maximale d'énergie qu'un data center peut demander
q_k^t	Quantité optimale d'énergie attribuée à un data center d_k
DCp_k^t	Priorité de charge d'un data center
CAP_k^t	Priorité des applications critiques d'un data center d_k
TAP_k^t	Priorité des applications critiques totale d'un data center d_k
$GECp_k^t$	Priorité du coefficient d'énergie verte d'un data center d_k
$REPP_k^t$	Priorité de production d'énergie renouvelable d'un data center d_k
Gn_k	Fonction de gain
$\gamma_k(q_k^t)$	Fonction d'utilité
$\theta_k(q_k^t, DCp_k^t)$	La fonction de coût de charge d'un data center
SEC_k^t	Niveau de charge énergétique d'un data center d_k
s_k^t	Nombre de serveurs actifs d'un data center d_k
$\phi_k(q_k^t, CAP_k^t)$	Fonction de coût des applications critiques
rt_k^t	Nombre d'applications en temps réel dans un data center d_k
$apps_k^t$	Nombre total des applications en cours d'exécution dans un data center d_k
$\rho_k(q_k^t, TAP_k^t)$	Fonction de coût total des applications en temps réel
$\varphi_k(q_k^t, GECp_k^t)$	La fonction de coût de l'utilisation de l'énergie verte
GC_k^t	La consommation d'énergie verte d'un data center d_k (renouvelable)
PC_k^t	La consommation totale d'énergie d'un data center d_k (renouvelable + non renouvelable)
PUE_k^t	Efficacité de l'utilisation de l'énergie de d_k
C_{peak}	La consommation de pointe d'un serveur dans un data center
C_{idle}	La consommation d'un serveur au repos dans un data center
ut_k^t	L'utilisation moyenne des serveurs d'un data center d_k
$\varrho_k(q_k^t, REPP_k^t)$	Fonction de coût de production des énergies renouvelables
GP_k	Quantité d'énergie renouvelable produite par un data center d_k
$\alpha, \beta, \psi, \omega, \varpi, \vartheta$	Coefficients de préférences
DEC_k	émission de dioxyde de carbone d'un data center d_k
G_{ce}	Taux d'émission de carbone de l'énergie du réseau électrique (non renouvelable)
W_{ce}	Taux d'émission de carbone de l'énergie éolienne

Où q_k^t est la quantité d'énergie qu'un data center d_k demande dans un instant $t \in T$.

2. Fonction de coût de charge : le coût qu'un data center doit payer en fonction de la quantité d'énergie qu'il demande q_k^t et de sa priorité DCp_k^t . Un data center a une priorité élevée lorsqu'il a des serveurs moins actifs et un niveau de charge énergétique faibles. Nous définissons donc cette fonction de coût comme suit :

$$\theta_k(q_k^t, DCp_k^t) = q_k^t \times DCp_k^t, \forall k \in K, \forall t \in T \quad (6.2)$$

Où

$$DCp_k^t = 1 - \frac{1}{(C_{idle} * s_k^t) + SEC_k^t} \quad (6.3)$$

et C_{idle} est la consommation d'énergie au repos d'un serveur dans le data center.

3. Fonction de coût des applications critiques : représente le coût qu'un data center doit payer en fonction de sa demande de quantité d'énergie q_k^t et de sa priorité d'application critique CAP_k^t . Cette priorité est calculée par le rapport entre les applications en temps réel rt_k^t exécutées dans un data center d_k et le nombre total d'applications en cours d'exécution $apps_k^t$. Nous définissons donc cette fonction comme suit :

$$\phi_k(q_k^t, CAP_k^t) = q_k^t \times CAP_k^t, \forall k \in K, \forall t \in T \quad (6.4)$$

Où :

$$CAP_k^t = 1 - \frac{rt_k^t}{apps_k^t} \quad (6.5)$$

4. La fonction de coût de l'ensemble des applications en temps réel : Elle est basée sur la demande de quantité d'énergie du data center q_k^t et sa priorité TAP_k^t . Un data center a un coût inférieur lorsque le nombre de ses applications en temps réel par rapport au nombre total des applications en temps réel de tous les data centers est élevé. Nous modélisons cette fonction de coût comme suit :

$$\rho_k(q_k^t, TAP_k^t) = q_k^t \times TAP_k^t, \forall k \in K, \forall t \in T \quad (6.6)$$

Où :

$$TAP_k^t = 1 - \frac{rt_k^t}{\sum_{k=1}^n rt_k^t} \quad (6.7)$$

5. La fonction de coût de l'utilisation de l'énergie verte : pénalise le joueur en fonction de sa demande de quantité d'énergie q_k^t et de sa priorité $GECp_k^t$. Cette fonction de coût favorise les data centers à forte consommation d'énergie verte. $GECp_k^t$ calcule le coefficient d'énergie verte GEC [101] d'un data center. Ainsi, nous définissons la fonction comme suit :

$$\varphi_k(q_k^t, GECp_k^t) = q_k^t \times GECp_k^t \quad (6.8)$$

Nous modélisons le coefficient d'énergie verte GEC de manière à impliquer deux paramètres

du data center. Le premier est la consommation d'énergie verte d'un data center d_k notée par GC_k^t et le second est la consommation totale d'énergie d'un data center d_k notée par PC_k^t . Nous donnons GEC_k^t comme suit :

$$GECp_k^t = 1 - \frac{GC_k^t[kwh]}{PC_k^t[kwh]} \forall k \in K, \forall t \in T \quad (6.9)$$

Où

$$0 < GC_k^t \leq PC_k^t, \forall k \in K, \forall t \in T \quad (6.10)$$

En se basant sur le modèle présenté dans [47], nous modélisons la consommation électrique totale du data center PC_k^t comme suit :

$$PC_k^t = s_k^t [C_{idle} + (C_{peak} - C_{idle}) ut_k^t] PUE_k^t \quad (6.11)$$

Où C_{peak} et C_{idle} sont respectivement la consommation de pointe et la consommation au repos d'un serveur dans le data center. ut_k^t représente l'utilisation moyenne des serveurs d'un data center d_k dans un instant t . PUE_k^t représente l'efficacité de la consommation électrique d'un data center d_k dans un instant t . (PUE) est la mesure la plus populaire en matière d'efficacité énergétique des data centers. Il s'agit du rapport entre l'énergie totale consommée par le data center et l'énergie consommée par ses équipements informatiques [40].

6. La fonction de coût de la production d'énergie renouvelable : cette fonction de coût est basée sur la demande de quantité d'énergie du data center q_k^t et sa priorité de production d'énergie renouvelable $REPP_k^t$. La priorité $REPP_k^t$ est calculée comme étant la quantité d'énergie renouvelable produite par un data center divisée par l'énergie renouvelable produite par tous les data centers. Ainsi, nous avons :

$$\varrho_k(q_k^t, REPP_k^t) = q_k^t \times REPP_k^t \quad (6.12)$$

Où :

$$REPP_k^t = \frac{GP_k}{\sum_{k=1}^n GP_k} \quad (6.13)$$

Dans notre système, les data centers produisent leur énergie renouvelable en utilisant des éoliennes, nous avons choisi d'estimer la production d'énergie éolienne en utilisant la méthode de distribution de Weibull. Cette méthode est l'une des méthodes largement acceptées pour estimer l'énergie éolienne [106]. Ainsi, nous modélisons la production d'énergie renouvelable d'un data center comme une variable aléatoire de Weibull GP_k avec une fonction de densité de probabilité [9] :

$$f_{GP_k}(GP_k; \lambda_1, \lambda_2) = \begin{cases} \frac{\lambda_2}{\lambda_1} \left(\frac{GP_k}{\lambda_1}\right)^{\lambda_2-1} e^{-(GP_k/\lambda_1)^{\lambda_2}} & ; GP_k \geq 0 \\ 0 & ; GP_k < 0 \end{cases} \quad (6.14)$$

Où $\lambda_1 > 0$ est le paramètre d'échelle et $\lambda_2 > 0$ est le paramètre de forme de la distribution.

Compte tenu des fonctions détaillées ci-dessus, nous donnons la fonction de gain de notre jeu comme suit :

$$Gn_k^t(q_k^t, q_{-k}^t) = \alpha\gamma_k(q_k^t) - \beta\theta_k(q_k^t, DCP_k^t) - \psi\phi_k(q_k^t, CAP_k^t) - \omega\rho_k(q_k^t, TAP_k^t) - \varpi\varphi_k(q_k^t, GECp_k^t) - \vartheta\rho_k(q_k^t, REPP_k^t) \quad (6.15)$$

Où :

- $q_{-k}^t = \{q_o^t\}(o \in K)$ and $k \neq o$ représente les stratégies (quantité d'énergie) de tous les data centers sauf le data center d_k .
- $st = (q_k^t, q_{-k}^t) \in St$ le profile de stratégies.
- Les coefficients de préférences $\alpha, \beta, \psi, \omega, \varpi$ et ϑ doivent être supérieure à 0.

6.2.3 Notre solution optimale

Avant de calculer la solution optimale de notre jeu, il est nécessaire de démontrer que notre jeu non coopératif a une solution. Par conséquent, nous détaillons en annexe de ce manuscrit la démonstration de l'existence de l'équilibre de Nash (NE), et son unicité.

Pour calculer notre solution optimale, nous modélisons un problème d'optimisation non linéaire sous contraintes, et le résolvons à l'aide de la méthode des multiplicateurs de Lagrange.

Dans notre jeu, chaque data center vise à maximiser son profit en maximisant ses demandes d'énergie. Nous modélisons donc notre problème (**F'**) comme suit :

$$\begin{aligned} \text{Maximiser}_{q_k^t} \quad & Gn_k^t(q_k^t, q_{-k}^t) \\ \text{S.C :} \quad & q_k^t \geq 0, \\ & q_k^t - q_k^{max} \leq 0, \forall k \in K. \end{aligned}$$

Pour résoudre le problème (**F'**), nous utilisons deux multiplicateurs de Lagrange : ℓ_k et ℓ'_k . Nous donnons la fonction de Lagrange comme suit :

$$M_k(q_k^t, \ell_k, \ell'_k) = Gn_k^t(q_k^t, q_{-k}^t) + \ell_k q_k^t + \ell'_k (q_k^{max} - q_k^t) \quad (6.16)$$

Par conséquent, la fonction de Lagrange définie satisfait aux conditions KKT suivantes pour chaque joueur d_k :

$$\begin{aligned} \ell_k, \ell'_k &\geq 0 \\ q_k^t &\geq 0 \\ q_k^{max} - q_k^t &\geq 0 \\ \nabla_{q_k^t} Gn_k^t(q_k^t, q_{-k}^t) + \ell_k \nabla_{q_k^t} q_k^t + \\ &\ell'_k \nabla_{q_k^t} (q_k^{max} - q_k^t) = 0 \\ \ell_k q_k^t + \ell'_k (q_k^{max} - q_k^t) &= 0 \end{aligned} \quad (6.17)$$

Après résolution, nous déterminons la quantité d'énergie optimale demandée (q_k^{t*}) pour le data center $d_k, \forall k \in K$, comme suit :

$$q_k^{t*} = \begin{cases} 0 & \text{si cdt 1} \\ q_k^{max} & \text{si cdt 2} \\ \frac{\alpha}{\beta\theta_k(q_k^t, DCP_k^t) + \psi\phi_k(q_k^t, CAP_k^t) + \omega\rho_k(q_k^t, TAP_k^t) + \varpi\varphi_k(q_k^t, GECp_k^t) + \vartheta\varrho_k(q_k^t, REPP_k^t)} - 1 & \text{sinon} \end{cases} \quad (6.18)$$

Où cdt 1 et 2 sont, respectivement :

$$\beta\theta_k(q_k^t, DCP_k^t) + \psi\phi_k(q_k^t, CAP_k^t) + \omega\rho_k(q_k^t, TAP_k^t) + \varpi\varphi_k(q_k^t, GECp_k^t) + \vartheta\varrho_k(q_k^t, REPP_k^t) \geq \alpha \quad (6.19)$$

$$\beta\theta_k(q_k^t, DCP_k^t) + \psi\phi_k(q_k^t, CAP_k^t) + \omega\rho_k(q_k^t, TAP_k^t) + \varpi\varphi_k(q_k^t, GECp_k^t) + \vartheta\varrho_k(q_k^t, REPP_k^t) \leq \frac{\alpha}{q_k^{max} + 1} \quad (6.20)$$

Après avoir déterminé la solution optimale, nous exploitons ce résultat pour montrer l'efficacité de notre jeu non-coopératif dans la section suivante.

6.3 Simulations et résultats

Cette section présente les résultats de nos simulations et étudie les performances de notre approche par rapport à une autre approche existante.

6.3.1 Paramètres des simulations

Nous considérons un ensemble de trois data centers géographiquement dispersés qui demandent un maximum d'énergie de $q_k^{max} = 10MWh$ au smart grid. De plus, les trois data centers disposent d'une deuxième source d'énergie et produisent leur propre énergie renouvelable à l'aide d'éoliennes. Chaque data center est équipé de 6000 serveurs homogènes. Nous effectuons notre simulation sur une période de 24 heures et mesurons la demande d'énergie du data center chaque 2 heures à deux périodes différentes : en heures de pointe (de 6 heures à 10 heures) et en heures creuses (de 10 heures à 6 heures). On notera que les valeurs des coefficients de la fonction de gain sont fixées par le contrôleur du smart grid en fonction de ses besoins. Nous avons fixé $\alpha = 0,5$ afin d'inciter les data centers à acheter fortement l'énergie du smart grid, également parce que la somme des pénalités ($\beta, \psi, \omega, \varpi$ et ϑ) est égale à 0,5.

Les simulations ont été réalisées sur un Dell OptiPlex 7050 (Intel Core i7 CPU 2.90GHz avec 32G RAM). La détermination de la solution optimale nous a permis d'exécuter notre programme en 0,94 seconde avec 100 itérations. Les paramètres de notre simulation sont détaillés dans le tableau 6.2.

TABLE 6.2 – Paramètres des simulations

Paramètres	Valeurs
q_k^{max}	10 MWh
Nombre de data centers (n)	3
s_k^t	[1 – 6000]
Taux d'utilisation des serveurs (ut_k^t)	[2.2 – 3]
C_{idle}	[70 – 100] MW
C_{peak}	[200 – 300] MW
PUE_k^t	[1.1 – 2]
$\alpha, \beta, \psi, \omega, \varpi, \vartheta$	(0.5, 0.1, 0.1, 0.1, 0.1, 0.1)
Heures pleines	[6am – 10pm])
Heures creuses	[10pm – 6am])

Pour évaluer notre approche, nous la comparons avec une approche existante présentée dans [109]. Les auteurs de ce travail considèrent les énergies renouvelables afin de minimiser le coût énergétique des data centers du cloud et optimiser leur consommation électrique. Ils abordent le problème de la gestion de l'énergie des data centers du point de vue du fournisseur de cloud computing et tentent d'assigner les demandes des clients au data center qui a suffisamment d'énergie renouvelable. Nous avons appelé cette approche "Renewable Energy Based (*REB*) ". Dans nos simulations, nous définissons globalement les mêmes paramètres que ceux utilisés dans [109] et considérons la même infrastructure de production d'énergie renouvelable dans chaque data center. La disponibilité de l'énergie éolienne dans chaque data center sur une journée (24 heures) est représentée dans la figure 6.3.

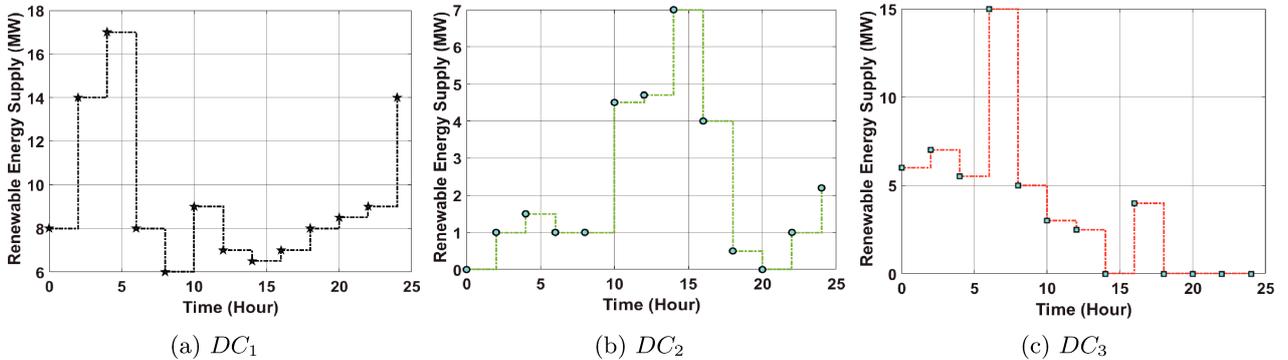


FIGURE 6.3 – Production d'énergie renouvelable des data center

6.3.2 La répartition d'énergie entre les data centers

La figure 6.4 indique la variation des coûts des priorités des data centers pendant une journée. Nous observons dans la figure 6.4-a que le coût de DC_p des trois data centers est faible aux heures creuses lorsque les data centers n'ont pas beaucoup de serveurs actifs et d'énergie stockée, et élevé aux heures de pointe lorsqu'ils ont chargé leur énergie et ont un nombre élevé de serveurs actifs. Dans notre simulation, nous avons défini beaucoup plus d'applications en temps réel aux heures de pointe

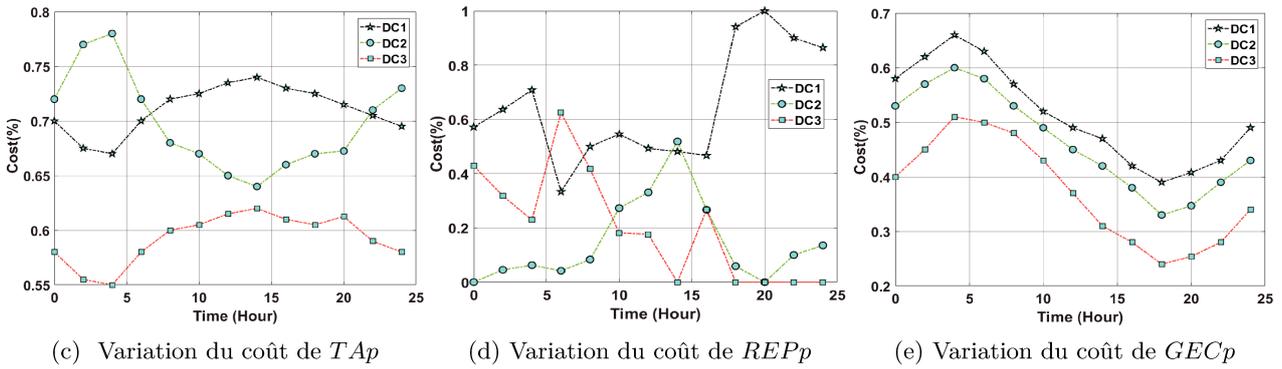
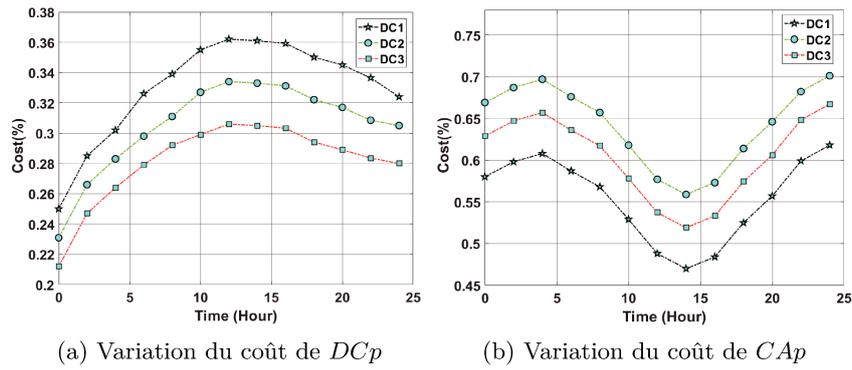


FIGURE 6.4 – Variation des coûts des data centers

par rapport au nombre d'applications fonctionnant aux heures creuses. Par conséquent, les variations du coût de CAp illustrés dans la figure 6.4-b sont plus élevés aux heures creuses qu'aux heures de pointe. En outre, nous pouvons observer que les coûts de TAp dans la figure 6.4-c varient correctement en fonction du nombre d'applications en temps réel de chaque data center par rapport au nombre total d'applications en temps réel en cours d'exécution dans tous les data centers. En outre, nous voyons dans la figure 6.4-d que les coûts de la priorité REP des trois data centers varient en fonction de leur production d'énergie éolienne représentée dans la figure 6.3. De plus, nous avons modélisé notre approche de façon à ce que le data center encourage l'utilisation de l'énergie du réseau à un prix bas en dehors des heures de pointe et stocke l'énergie des éoliennes pour une utilisation ultérieure aux heures de pointe lorsque le prix de l'énergie du réseau est élevé. Nous pouvons le vérifier dans la figure 6.4-e où les coûts de $GECp$ sont élevés aux heures creuses et faibles aux heures de pointe dans chaque data center.

La figure 6.5 montre la quantité d'énergie qui a été attribuée à chaque data center pendant une journée en utilisant notre approche de jeu non coopératif et l'approche de comparaison REB [109]. Nous remarquons que l'énergie assignée aux data centers en utilisant la méthode REB est faible aux heures creuses et élevée aux heures de pointe. Cela correspond à leur approvisionnement en énergie renouvelable, comme le montre la figure 6.3, où ils disposent d'un approvisionnement élevé en énergie renouvelable aux heures creuses et d'un faible approvisionnement en énergie renouvelable aux heures de pointe. Comme pour l'approche REB , l'énergie attribuée par notre approche est faible aux heures

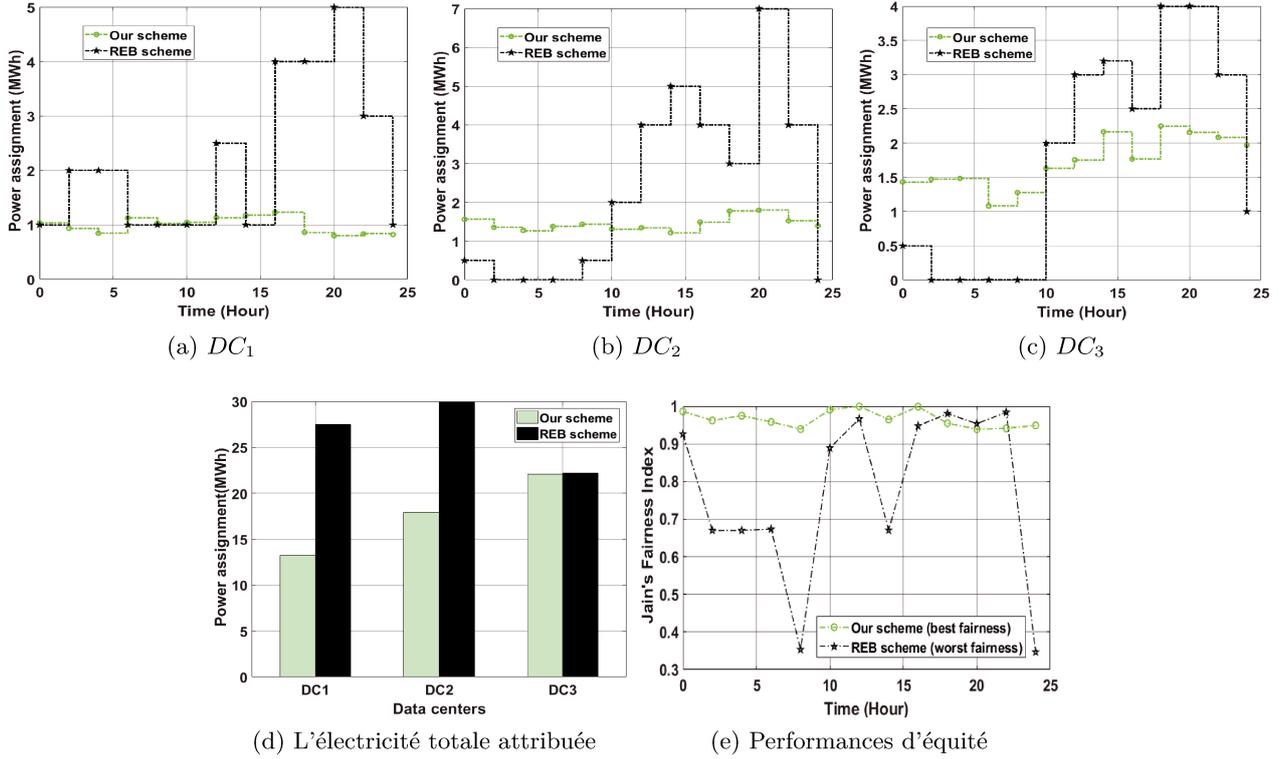


FIGURE 6.5 – Performances de notre jeu non-coopératif par rapport à *REB*

creuses et élevée aux heures de pointe, mais notre approche permet d'avoir un équilibre entre l'utilisation de l'énergie du smart grid et l'énergie renouvelable. En fait, notre approche pénalise les data centers de ne pas avoir plus d'énergie lorsqu'ils ont une priorité *REP* élevée et une faible priorité *GEC*. Cela signifie que notre approche incite les data centers à être plus efficaces en termes d'utilisation de l'énergie verte et à consommer plus d'énergie renouvelable en parallèle avec l'énergie du smart grid. Pour mieux montrer les performances de notre approche, nous présentons dans la figure 6.5-d l'énergie totale attribuée à chaque data center pendant une journée. Nous pouvons remarquer que la quantité d'énergie assignée en utilisant notre approche est inférieure à la quantité d'énergie assignée en utilisant l'approche *REB*. Cela est dû à notre jeu non coopératif où le contrôleur smart grid répartit équitablement l'énergie entre les data centers en les pénalisant en fonction de leurs différentes priorités, contrairement à l'approche *REB* qui ne prend en compte que l'approvisionnement en énergie renouvelable des data centers.

Pour mieux montrer les performances de notre proposition en termes d'équité, nous comparons dans la figure 6.5-e l'indice d'équité de Jain de notre approche par rapport à l'approche *REB*. L'indice d'équité de Jain est la mesure d'équité entre les flux actifs (qui sont les data centers dans notre cas) dans un réseau. L'indice se situe entre $\frac{1}{n}$ et 1, représentant respectivement la *pire* et la *meilleure* équité [131]. Nous calculons l'indice d'équité de Jain à l'aide de la fonction suivante [94] :

$$Jain(Y) = \frac{[\sum_{k=1}^n Y_k]^2}{\sum_{i=1}^m Y_i^2} \quad (6.21)$$

TABLE 6.3 – Taux des émissions de dioxyde de carbone par l'énergie fossile et éolienne

Sources d'énergie	Fossile	Éolienne
CER(gCO_2e/kwh)	968	22.5

Où

$$Y_k = \frac{q_k^{max}}{q_k} \quad (6.22)$$

Nous devons noter que la répartition tend à être plus équitable lorsque l'indice de Jain est plus proche de 1, et nous pouvons voir sur la figure 6.5-e que notre approche a une répartition plus équitable par rapport à celle de *REB*.

6.3.3 Les performances de notre approche en termes de taux de réduction de charge et d'émissions de dioxyde de carbone

Dans cette sous-section, nous considérons deux métriques pour évaluer notre approche : les émissions de dioxyde de carbone et le taux de charge électrique. Le taux de charge électrique représente le rapport entre l'énergie totale attribuée par le réseau principal et la quantité maximale d'énergie qui peut être attribuée au cours d'une journée. La métrique des émissions de dioxyde de carbone mesure la quantité de dioxyde de carbone qui a été émise par chaque data center au cours d'une journée. Nous calculons le taux des émissions de dioxyde de carbone DEC_k de chaque data center en utilisant la fonction suivante :

$$DEC_k = \sum_{t=1}^T (q_k^t * G_{ce}) + ((q_k^{max} - x_k^t) * W_{ce}) \quad (6.23)$$

Où G_{ce} et W_{ce} sont respectivement le taux des émissions de carbone *CER* [116] de l'énergie du grid (non renouvelable) et de l'énergie renouvelable des éoliennes présentés dans le tableau 6.3.

Nous donnons dans la figure 6.6-a le taux de charge électrique de chaque data center en utilisant notre jeu non-coopératif et l'approche de comparaison *REB*. Nous voyons que dans chaque data center notre approche a un taux de charge électrique inférieur à celui de *REB*. De plus, notre approche basée sur un jeu non-coopératif surpasse l'approche *REB* avec un pourcentage de 31,2%.

La figure 6.6-b montre les performances de notre approche en termes d'émissions de Co2. Nous pouvons voir que notre proposition permet de réduire considérablement les émissions de dioxyde de carbone par rapport à l'approche de comparaison *REB*. Cela peut s'expliquer par le fait que notre jeu favorise l'utilisation des énergies renouvelables tout en répartissant équitablement l'énergie du réseau entre les data centers.

6.4 Conclusion

Nous avons présenté, dans ce chapitre, notre deuxième contribution qui consiste à considérer les énergies renouvelables dans la répartition d'énergie entre les data centers dans un système smart grid - cloud. Nous avons modélisé la répartition de l'énergie du smart grid vers les data centers comme

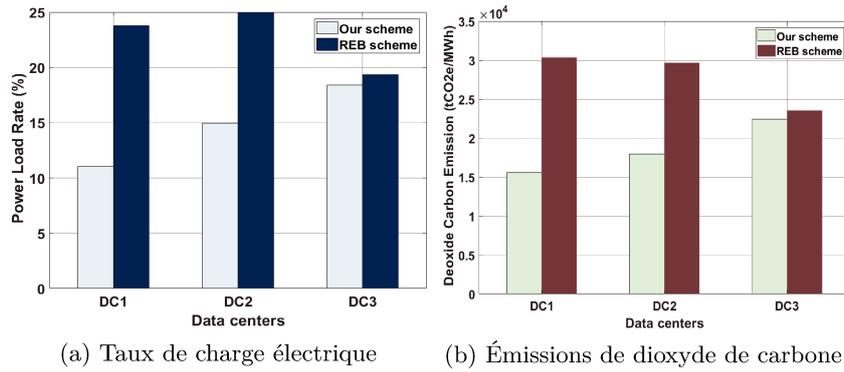


FIGURE 6.6 – Évaluation de notre proposition en termes de taux de charge électrique et des émissions de dioxyde de carbone

un jeu non coopératif, ensuite nous avons calculé la quantité optimale d'énergie à fournir à chaque data center. Les résultats des simulations ont montré que notre approche attribue l'énergie aux data centers de manière équitable en fonction de leurs différentes priorités.

Afin de valider notre proposition, nous l'avons comparée avec une approche existante de minimisation d'énergie, et avons montré que notre approche est plus efficace en termes de taux de charge d'énergie avec un pourcentage de 31,2% et qu'elle réduit de manière significative l'émission de dioxyde de carbone. De plus, nous avons prouvé le caractère équitable de notre approche d'allocation d'énergie en utilisant l'indice d'équité de Jain.

Notre approche présentée dans ce chapitre permet d'avoir une équité dans la répartition d'énergie entre les data centers, mais ne montre en aucun cas son efficacité en termes de réduction du coût lié à l'énergie. Dans ce contexte, nous présentons dans le chapitre suivant notre contribution qui vise à minimiser le coût d'énergie des data centers dans une architecture microgrid - cloud.

Minimiser le coût énergétique des data centers en utilisant les Microgrids

Sommaire

7.1	Introduction et motivations	84
7.2	Modèle et architecture du système	84
7.2.1	Architecture du système	84
7.2.2	Niveau 1 : Description du jeu pour la répartition de l'électricité entre les data centers	87
7.2.3	Niveau 2 : Minimiser le coût énergétique des microgrids	90
7.3	Résultats des simulations	93
7.3.1	Paramètres des simulations	93
7.3.2	Répartition de l'électricité entre les data centers	94
7.3.3	Minimiser le coût d'énergie des microgrids	96
7.3.4	Les performances de notre approche en termes de taux de réduction de charge et des émissions de Co2	98
7.4	Conclusion	99

7.1 Introduction et motivations

Les microgrids sont conçus pour fonctionner à échelle locale. Ils sont de petite taille, et ont pour objectif de fournir un approvisionnement électrique fiable à un petit nombre de consommateurs. Les microgrids permettent une meilleure résistance face aux imprévus ou aux dommages subis sur le réseau d’approvisionnement principal [60].

Dans ce contexte, faire fonctionner des data centers du cloud dans des microgrids peut être plus fiable et plus efficace que de les faire gérer directement par l’opérateur du smart grid. Cependant, les microgrids doivent faire face au coût élevé de l’énergie que leurs data centers peuvent engendrer. Dans ce contexte, la gestion de l’énergie allouée aux data centers et la minimisation du coût énergétique des microgrids constituent un véritable défi.

Dans ce chapitre, nous modélisons une approche d’optimisation en deux étapes dans une architecture microgrids - cloud. Dans un premier temps, nous considérons que les data centers sont gérés par différents fournisseurs de cloud et tentent d’avoir le maximum d’énergie de leur microgrid respectif, afin de stocker le maximum d’énergie et de pouvoir exécuter le maximum d’applications d’utilisateurs. Un tel comportement énergétique requiert une répartition optimale de l’énergie de la part du microgrid afin d’atténuer le risque de pannes et de coupures de courant. Ainsi, nous modélisons l’énorme demande d’énergie des data centers comme un jeu non coopératif. Le contrôleur du microgrid calcule la quantité optimale d’énergie à attribuer à chaque data center en fonction de (i) son Power Usage Effectiveness (PUE), (ii) du nombre d’applications en temps réel et (iii) de son utilisation de la bande passante du réseau. Dans un deuxième temps, les microgrids tentent de minimiser leur coût énergétique en achetant de façon optimale l’énergie du réseau principal et des autres microgrids, et aussi en revendant leur énergie stockée.

Des simulations sont effectuées en comparant notre approche avec trois autres approches de réduction de la consommation d’énergie. Les résultats montrent que notre approche est plus performante que les autres approches de comparaison en termes de coût énergétique, de taux de charge électrique et d’émission de dioxyde de carbone.

7.2 Modèle et architecture du système

Nous présentons dans cette section notre nouvelle architecture microgrid - cloud et détaillons notre approche de minimisation à deux niveaux. Nous résumons les principaux paramètres utilisés dans notre système dans le tableau 7.1.

7.2.1 Architecture du système

Comme l’illustre la figure 7.1, nous considérons dans notre système un contrôleur smart grid qui gère l’approvisionnement en énergie de ses différents consommateurs. Le contrôleur smart grid est lié aux centrales électriques qui fournissent de l’énergie à des microgrids géographiquement répartis.

TABLE 7.1 – Nomenclature

J	Ensemble des microgrids
N	Ensemble des data centers
T	Horizon temporel
m_j	Microgrid j
dc_{ij}	Data center i du microgrid j
B_{ij}^t	Débit du data center dc_{ij}
R_{ij}^t	Nombre d'applications en temps réel d'un data center dc_{ij}
q_{ij}^{max}	Quantité maximale d'énergie qu'un data center peut demander
q_{ij}^t	Quantité d'énergie attribuée à un data center dc_{ij}
PUE_{ij}^t	Efficacité de la consommation d'énergie de dc_{ij}
CAR_{ij}^t	Taux d'application critique de dc_{ij}
NBU_{ij}^t	Utilisation de la bande passante du réseau de dc_{ij}
Gn_{ij}	Fonction de gain du jeu
$\alpha, \beta, \gamma, \delta$	Coefficients de préférences
C_{ij}^1	Fonction du coût de PUE
C_{ij}^2	Fonction du coût de CAR
C_{ij}^3	Fonction du coût de NBU
y_j^{max}	Quantité maximale d'énergie qu'un microgrid m_j peut demander du smart grid
SEC_j^t	Niveau de l'énergie stockée du microgrid
S_j^{min}	La quantité minimale d'énergie stockée dans chaque microgrid
S_j^{max}	La quantité maximale d'énergie qu'un microgrid peut stocker
Sl_j	Le seuil à partir duquel un microgrid peut vendre son énergie
Pg_j	Prix de vente de l'électricité du smart grid
Pm_j	Prix de vente d'un microgrid $m_k \forall k \neq j$
Ps_j	Prix de vente d'un microgrid m_j
yg_j^t	Quantité d'électricité achetée par un microgrid m_j du smart grid
ym_{jk}^t	Quantité d'électricité achetée d'un microgrid par un autre microgrid m_j
ys_j^t	La quantité d'électricité vendue par un microgrid m_j
π_j^t	Prix de base
sp_j	Paramètre de sensibilité
$Q_Avg_j^t$	Quantité moyenne d'énergie allouée par un microgrid m_j à ses data centers à $(t - 1)$
$C_Sell_j^t$	Les gains de la quantité d'énergie vendue par un microgrid m_j
$C_Buy_j^t$	Coût de l'énergie achetée d'un microgrid m_j
DE_j	Émissions de dioxyde de carbone par un microgrid m_j

Chaque microgrid j désigné par $m_j, \forall j \in J$ avec $J = \{1, 2, \dots, j, \dots, J\}$ a la responsabilité d'alimenter n_j data center avec de l'électricité achetée du réseau principal et son énergie renouvelable auto-générée par des éoliennes. Pour faire face à l'intermittence des énergies renouvelables, chaque microgrid est équipé de batteries pour stocker son énergie. En outre, nous considérons dans notre système un échange d'énergie à deux sens où chaque microgrid peut acheter de l'énergie du réseau principal et vendre de l'énergie vers les autres microgrids du système.

Dans un système smart grid, une communication bidirectionnelle peut être établie entre les consommateurs et les fournisseurs de services énergétiques. Ainsi, dans notre proposition, le contrôleur du smart grid, les contrôleurs microgrids et les data centers échangent périodiquement des informations sur l'énergie par le biais de la communication par courant porteur en ligne (CPL). Ainsi, nous avons, d'une part, une communication verticale où les contrôleurs du smart grid et des microgrids ont une visibilité totale sur le comportement énergétique de chaque data center. D'autre part, nous avons une communication horizontale entre les contrôleurs qui échangent périodiquement l'état de leur charge énergétique et le prix de vente correspondant [66].

Les data centers homogènes qui fonctionnent dans chaque microgrid sont gérés par différents fournisseurs de cloud. Chaque data center i alimenté par un microgrid $m_j, \forall j \in J$ with $J = \{1, 2, \dots, j, \dots, J\}$ est désigné par $dc_{ij}, \forall i \in N$ with $N = \{1, 2, \dots, i, \dots, n\}$. Nous considérons dans notre système une période de répartition d'énergie de $T = \{1, \dots, t, \dots, \tau\}$ et nous divisons le temps quotidien en $\tau = 24$. Nous supposons que chaque data center dc_{ij} , dans chaque tranche de temps $t \in T$, exécute un certain nombre Rt_{ij}^t d'applications en temps réel, utilise B_{ij}^t de bande passante et consomme ec_{ij}^t d'énergie. Nous considérons également que le smart grid peut alimenter chaque microgrid m_j avec une quantité maximale d'énergie de y_j^{max} et que les data centers peuvent de la même manière acheter une quantité maximale d'énergie de q_{ij}^{max} du microgrid. Par conséquent, afin d'allouer l'énergie de manière optimale aux data centers, les microgrids prennent en compte trois paramètres principaux des data centers : PUE_{ij}^t, CAR_{ij}^t et NBU_{ij}^t qui sont respectivement l'utilisation de l'efficacité énergétique, le taux d'applications critiques et l'utilisation de la bande passante du réseau [101] [40]. Nous définissons ces métriques comme suit :

1. Le Power Usage Effectiveness PUE_{ij}^t : Cette métrique représente dans chaque tranche de temps $t \in T$ le rapport entre l'énergie totale consommée par le data center et l'énergie consommée par ses équipements informatiques [101].
2. Taux d'applications critiques CAR_{ij}^t : représente dans chaque instant $t \in T$ le ratio du nombre d'applications en temps réel en cours d'exécution dans un data center, par rapport au total des applications en temps réel en cours d'exécution dans tous les data centers alimentés par le même microgrid.
3. Utilisation de la bande passante NBU_{ij}^t : reflète dans instant $t \in T$ le rapport entre la bande passante utilisée par le data center et sa capacité totale en bande passante.

Pour répondre à leurs demandes et améliorer les bénéfices de leurs fournisseurs de cloud respectifs, les data centers se font concurrence pour obtenir le maximum d'énergie du microgrid. Un tel comportement nécessite une répartition optimale de l'énergie de la part du microgrid afin d'éviter le risque de pannes et de coupures de courant. C'est pourquoi, dans un premier temps, nous modélisons

le comportement énergétique des data centers comme un jeu non coopératif, dans lequel le contrôleur du microgrid calcule la quantité d'énergie optimale à attribuer à chaque data center en fonction de ses paramètres : $PU E_{ij}^t, CAR_{ij}^t$ et NBU_{ij}^t . Ensuite, en se fondant sur le résultat du jeu de répartition d'énergie et sur son niveau d'énergie stockée, le microgrid essaie, dans un deuxième temps, de minimiser son coût énergétique en achetant de façon optimale l'énergie du réseau principal et des autres microgrids.

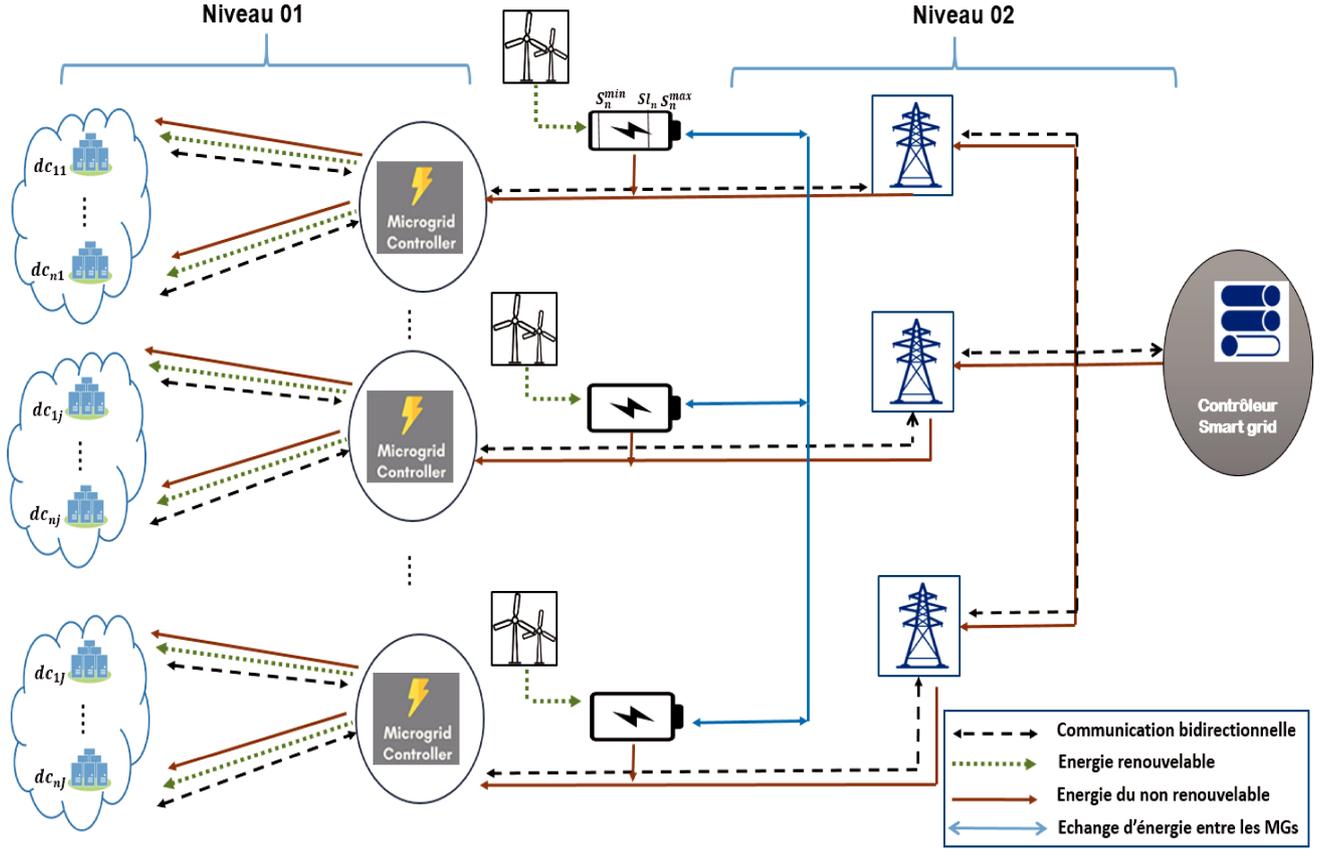


FIGURE 7.1 – Architecture smart microGrid- cloud

7.2.2 Niveau 1 : Description du jeu pour la répartition de l'électricité entre les data centers

Modélisation du jeu non-coopératif

Pour modéliser la répartition d'énergie des data centers comme un jeu non coopératif. Nous devons définir les joueurs, les stratégies et la fonction de gain du jeu. Nous définissons donc le jeu : $G = (N, St_{ij}, Gn_{ij})_{i \in N} \forall j \in J$ où :

- N : est un ensemble de joueurs $dc_{1j}, \dots, dc_{ij}, \dots, dc_{nj}$ où n le nombre des joueurs du jeu.
- St_{ij} ; $\forall i \in N$: représente les stratégies des joueurs dans le jeu. Nous définissons la stratégie dans notre jeu comme la quantité d'énergie qu'un data center dc_{ij} demande au microgrid. Par conséquent, la stratégie est q_{ij} et varie entre zéro et q_{ij}^{max} .

- $Gn_{ij} : St_{ij} \rightarrow \mathbb{R}$ représente la fonction de gain du joueur $dc_{ij} \forall i \in N$. L'objectif de chaque joueur est de maximiser sa fonction de gain, Gn , $0 \leq Gn_{ij} \leq q_{ij}^{max}, \forall i \in N$, afin d'augmenter ses bénéfices.

Afin d'assurer une répartition équitable de l'énergie entre les data centers, le contrôleur du micro-grid répartit l'énergie en fonction des trois paramètres de data center. Par conséquent, la fonction de gain de notre jeu est conçue pour impliquer quatre fonctions : une fonction d'utilité, et une fonction de coût pour chaque métrique de data center. Ces fonctions sont définies comme suit :

1. Fonction d'utilité $Ut_{ij}(q_{ij}^t)$: nous modélisons la fonction d'utilité de manière à ce que plus le data center augmente sa valeur de quantité d'énergie, plus son profit est élevé. Ainsi, nous modélisons la fonction d'utilité d'un joueur dc_{ij} comme suit :

$$ut_{ij}(q_{ij}^t) = \log(q_{ij}^t + 1), \forall i \in N, \forall j \in J, \forall t \in T \quad (7.1)$$

2. Fonction du coût de l'efficacité énergétique : reflète le coût d'un data center en termes de quantité d'énergie demandée q_{ij}^t et sa valeur PUE_{ij}^t . Plus le data center est efficace en termes d'énergie, plus son coût est faible. Et plus son efficacité énergétique est faible, plus son coût est élevé. Nous définissons donc cette fonction de coût comme suit :

$$\begin{aligned} C_{ij}^1(q_{ij}^t, PUE_{ij}^t) &= q_{ij}^t \times \left(1 - \frac{1}{PUE_{ij}^t}\right) \\ &= q_{ij}^t \times \left(1 - \frac{IT_ec_{ij}^t[kwh]}{ec_{ij}^t[kwh]}\right), \\ &\forall i \in N, \forall j \in J, \forall t \in T \end{aligned} \quad (7.2)$$

Où :

$IT_ec_{ij}^t$ est la consommation d'énergie des équipements IT du data center et ec_{ij}^t sa consommation d'énergie totale.

3. Fonction de coût des applications critiques : cette fonction de coût vise à récompenser un joueur lorsqu'il a un nombre élevé d'applications en temps réel par rapport aux autres joueurs, et le pénalise lorsqu'il a un faible taux d'applications en temps réel. Nous définissons donc cette fonction comme suit :

$$\begin{aligned} C_{ij}^2(q_{ij}^t, CAR_{ij}^t) &= q_{ij}^t \times CAR_{ij}^t \\ &= q_{ij}^t \times \left(1 - \frac{Rt_{ij}^t}{\sum_{i=1}^n Rt_{ij}^t}\right), \\ &\forall i \in N, \forall j \in J, \forall t \in T \end{aligned} \quad (7.3)$$

4. Fonction de coût de la bande passante du réseau : cette fonction de coût pénalise le joueur en

fonction de sa demande de quantité de puissance q_{ij}^t et de son utilisation de la bande passante. Plus la bande passante d'un joueur est élevée, plus le coût est faible, et vice-versa. Nous donnons cette fonction de coût comme suit :

$$\begin{aligned}
C_{ij}^3(q_{ij}^t, NBU_{ij}^t) &= q_{ij}^t \times NBU_{ij}^t \\
&= q_{ij}^t \times \left(1 - \frac{B_{ij}^t [Gbps]}{BC_{ij}^t [Gbps]} \right), \\
&\forall i \in N, \forall j \in J, \forall t \in T
\end{aligned} \tag{7.4}$$

Où :

B_{ij}^t et BC_{ij}^t sont, respectivement, la valeur d'utilisation de la bande passante d'un data center et sa capacité de bande passante.

Compte tenu des fonctions détaillées ci-dessus, nous donnons la fonction de gain de notre jeu comme suit :

$$Gn_{ij}^t(q_{ij}^t, q_{-ij}^t) = \alpha Ut_{ij}(q_{ij}^t) - \beta C_{ij}^1(q_{ij}^t, PUE_{ij}^t) - \gamma C_{ij}^2(q_{ij}^t, CAR_{ij}^t) - \delta C_{ij}^3(q_{ij}^t, NBU_{ij}^t) \tag{7.5}$$

Solution du jeu

Pour identifier la solution optimale, nous modélisons un problème d'optimisation non linéaire, et le résolvons à l'aide de la méthode des multiplicateurs de Lagrange.

Dans notre jeu, chaque data center vise à maximiser son profit en maximisant ses demandes d'énergie. Ainsi, nous modélisons notre problème (\mathbf{F}') comme suit :

$$\begin{aligned}
&\underset{q_{ij}^t}{\text{Maximiser}} && Gn_{ij}^t(q_{ij}^t, q_{-ij}^t) \\
&\text{S.C} && q_{ij}^t \geq 0, \\
&&& q_{ij}^t - q_{ij}^{max} \leq 0, \forall i \in N; \forall j \in J.
\end{aligned}$$

Pour résoudre le problème (\mathbf{F}'), nous utilisons deux multiplicateurs de Lagrange ℓ_{ij} et ℓ'_{ij} pour maximiser notre fonction. Nous donnons la fonction de Lagrange comme suit :

$$M_{ij}(q_{ij}^t, \ell_{ij}, \ell'_{ij}) = Gn_{ij}^t(q_{ij}^t, q_{-ij}^t) + \ell_{ij} q_{ij}^t + \ell'_{ij} (q_{ij}^{max} - q_{ij}^t) \tag{7.6}$$

Par conséquent, la fonction de Lagrange répond aux conditions KKT suivantes :

$$\begin{aligned}
\ell_{ij}, \ell'_{ij} &\geq 0 \\
q_{ij}^t &\geq 0 \\
q_{ij}^{max} - q_{ij}^t &\geq 0 \\
\nabla_{q_{ij}^t} Gn_{ij}^t(q_{ij}^t, q_{-ij}^t) + \ell_{ij} \nabla_{q_{ij}^t} q_{ij}^t + \\
\ell'_{ij} \nabla_{q_{ij}^t} (q_{ij}^{max} - q_{ij}^t) &= 0 \\
\ell_{ij} q_{ij}^t + \ell'_{ij} (q_{ij}^{max} - q_{ij}^t) &= 0
\end{aligned} \tag{7.7}$$

Après résolution, nous déterminons la quantité optimale d'énergie (q_{ij}^{t*}) à attribuer pour le data center dc_{ij} , $\forall i \in N, \forall j \in J$ comme suit :

$$q_{ij}^{t*} = \begin{cases} 0 & \text{if cdt 1} \\ q_{ij}^{max} & \text{if cdt 2} \\ \frac{\alpha}{\frac{\beta C_{ij}^1(q_{ij}^t, PUE_{ij}^t) + \gamma C_{ij}^2(q_{ij}^t, CAR_{ij}^t) + \delta C_{ij}^3(q_{ij}^t, NBU_{ij}^t)}{\delta C_{ij}^3(q_{ij}^t, NBU_{ij}^t)} - 1}} & \text{otherwise} \end{cases} \tag{7.8}$$

Où cdt 1 and 2, sont représentées par (09) and (10) :

$$\begin{aligned}
\beta C_{ij}^1(q_{ij}^t, PUE_{ij}^t) + \gamma C_{ij}^2(q_{ij}^t, CAR_{ij}^t) + \\
\delta C_{ij}^3(q_{ij}^t, NBU_{ij}^t) &\geq \alpha
\end{aligned} \tag{7.9}$$

$$\begin{aligned}
\beta C_{ij}^1(q_{ij}^t, PUE_{ij}^t) + \gamma C_{ij}^2(q_{ij}^t, CAR_{ij}^t) + \\
\delta C_{ij}^3(q_{ij}^t, NBU_{ij}^t) &\leq \frac{\alpha}{q_{ij}^{max} + 1}
\end{aligned} \tag{7.10}$$

7.2.3 Niveau 2 : Minimiser le coût énergétique des microgrids

Dans la deuxième étape, le microgrid essaie de minimiser son coût d'énergie en vendant son énergie stockée aux microgrids et en achetant la quantité optimale d'énergie en fonction du résultat du jeu de répartition d'énergie de ses data centers.

Nous indiquons par SEC_j^t le niveau de charge énergétique de chaque microgrid m_j dans un instant t , $\forall t \in T$. Afin de gérer l'énergie stockée dans le microgrid, nous avons défini trois principaux états de charge énergétiques : S_j^{min} , Sl_j , et S_j^{max} .

S_j^{min} définit la quantité minimale d'énergie qui doit être stockée dans chaque microgrid pour une utilisation d'urgence. Sl_j indique le seuil à partir duquel un microgrid peut vendre son énergie. Et S_j^{max} est la quantité maximale d'énergie qu'un microgrid peut stocker.

Modéliser l'achat d'énergie des microgrids

Afin d'alimenter ses data centers, chaque microgrid m_j peut acheter dans chaque tranche de temps t , $\forall t \in T$ une quantité yg_j^t d'électricité du smart grid. Chaque microgrid m_j peut également acheter une quantité $ym_{j,k}^t$ d'énergie renouvelable d'un microgrid k , $k \neq j, \forall j, k \in J$. Par conséquent, nous modélisons l'énergie totale achetée par un microgrid m_j comme suit :

$$y_j^t = yg_j^t + \sum_{k=1}^J ym_{j,k}^t, \forall j, k \in J, k \neq j, \forall t \in T. \quad (7.11)$$

Où

$$yg_j^t \leq y_j^{max} \quad (7.12)$$

Le contrôleur du smart grid et les contrôleurs des microgrids actualisent et échangent périodiquement leurs prix de vente de l'électricité. Nous désignons par Pg_j^t et $Pm_{j,k}^t$ respectivement le prix de vente d'énergie du smart grid et du microgrid m_k à un microgrid m_j où $Pg_j^t, Pm_{j,k}^t > 0$ dans chaque t , $\forall j, k \in J$ and $j \neq k$.

Dans notre système, le smart grid utilise le prix en temps réel (Real Time Price - RTP) comme un système de tarification dynamique [110]. Nous modélisons le prix du smart grid Pg_j^t comme une tarification échelonnée qui prend en compte la valeur de stockage de l'énergie du microgrid comme suit :

$$Pg_j^t \pi_j^t + sp_j \left(\frac{Sl_j}{SEC_j^t} \right) \quad (7.13)$$

Où $\pi_j^t > 0$ représente le prix de base et $sp_j > 0$ est un paramètre de sensibilité. Le terme $sp_j \left(\frac{Sl_j}{SEC_j^t} \right)$ reflète le rapport entre le seuil de stockage d'énergie à partir duquel un microgrid peut vendre son énergie Sl_j et son niveau de charge énergétique SEC_j^t . Le prix du grid Pg_j^t augmente lorsque $SEC_j^t < Sl_j$ et diminue lorsque $SEC_j^t > Sl_j$.

Nous considérons dans notre système que chaque microgrid m_j achète de l'énergie sur la base de la quantité moyenne d'énergie assignée aux data centers dans le créneau temporel précédent ($t - 1$) en utilisant notre approche. Cette quantité d'énergie moyenne est désignée par $Q_Avg_j^t$ et nous utilisons pour définir la contrainte suivante :

$$y_j^t \geq Q_Avg_j^t, \forall j \in J, \forall t \in T \quad (7.14)$$

Où

$$Q_Avg_j^t = \sum_{i=1}^n q_{ij}^{*t-1}, \forall j \in J, \forall t \in T \quad (7.15)$$

Nous définissons donc le coût d'achat de l'énergie des microgrids comme suit :

$$C_Buy_j^t = \left(Pg_j^t \times yg_j^t \right) + \sum_{k=1}^J \left(Pm_{jk}^t \times ym_{jk}^t \times x_{jk} \right),$$

$$\forall j, k \in J, j \neq k, \forall t \in T$$

Où :

x'_{jk} est une variable binaire utilisée pour vérifier si un microgrid m_k a une quantité d'énergie à vendre à un microgrid m_j ou non, et elle est définie comme suit :

$$x_{jk} = \begin{cases} 0 & \text{si } SEC_k^t \leq Sl_k \\ 1 & \text{sinon} \end{cases} \quad (7.16)$$

Modéliser la vente des microgrids

Afin de minimiser son coût énergétique, chaque microgrid m_j vend une quantité ys_j^t de son énergie stockée avec un prix Ps_j^t .

Le microgrid peut vendre son énergie stockée au smart grid ou à un autre microgrid m_k . Pour ce faire, un microgrid m_j doit remplir deux conditions : (1) que son niveau de charge énergétique SEC_j^t doit être supérieur au seuil de vente Sl_j , (2) qu'il existe au moins un microgrid m_k avec un prix de vente Pg_k^t qui est supérieur au prix de vente Ps_j^t du microgrid m_j . Nous utilisons deux variables binaires x'_j et b_k pour exprimer ces deux conditions comme suit :

$$x'_j = \begin{cases} 0 & \text{si } SEC_j^t \leq Sl_j \\ 1 & \text{sinon} \end{cases} \quad (7.17)$$

Et

$$b_k = \begin{cases} 0 & \text{si } Pg_k^t < Ps_j^t \\ 1 & \text{sinon} \end{cases} \quad (7.18)$$

Par conséquent, nous modélisons le profit de vente d'énergie d'un microgrid m_j comme suit :

$$C_Sell_j^t = \left(Ps_j^t \times ys_j^t \times x'_j \times \omega_j \right), \forall j \in J, \forall t \in T$$

Où :

ω_j est une variable qui calcule le maximum de l'ensemble V . Cet ensemble contient les valeurs de la variable b_k de toutes les microgrids du système comme $V = \{b_1, \dots, b_k, \dots, b_J\}$, $\forall k \in J$ et $k \neq j$. Ainsi, si $\omega_j = 1$, alors il existe au moins un microgrid qui satisfait la deuxième condition et le microgrid m_j peut vendre son énergie. En revanche, si $\omega_j = 0$, alors le microgrid m_j ne peut vendre son énergie à aucun des microgrids qui ont une meilleure proposition de prix par le smart grid. Nous notons donc :

$$\omega_j = Max\{V\} \quad (7.19)$$

Modélisation du problème de minimisation du coût d'énergie des microgrids

Dans notre système, chaque microgrid vise à minimiser son coût énergétique total en achetant la quantité optimale d'énergie et en vendant l'énergie renouvelable stockée. Ainsi, le coût total de l'énergie d'un microgrid m_j est composé de deux parties : la partie coût d'achat d'énergie et la partie bénéfice de la vente d'énergie. Nous désignons par C_j^t le coût énergétique total d'un microgrid m_j à un intervalle de temps $t, \forall t \in T$, et nous le définissons comme suit :

$$\begin{aligned}
 C_j^t &= C_Buy_j^t - C_Sell_j^t \\
 &= \left(P g_j^t \times y g_j^t \right) + \sum_{k=1}^J \left(P m_{jk}^t \times y m_{jk}^t \times x_{jk} \right) \\
 &\quad - \left(P s_j^t \times y s_j^t \times x'_j \times \omega_j \right), \\
 &\quad \forall j, k \in J, j \neq k, \forall t \in T
 \end{aligned} \tag{7.20}$$

Nous modélisons le problème de minimisation des coûts énergétiques comme un problème d'optimisation linéaire sous contraintes, et le résolvons à l'aide de l'outil d'optimisation de Matlab. Nous donnons la formulation du problème de minimisation des coûts énergétiques d'un microgrid m_j comme suit :

$$\begin{array}{ll}
 \text{Minimiser} & C_j^t \left(y g_j^t, y m_{jk}^t, -y s_j^t \right) \\
 \text{S.C} & (11), (12), (13), (14), (15), (16), (17), (18)
 \end{array}$$

7.3 Résultats des simulations

Cette section présente nos résultats des simulations et étudie les performances de notre approche par rapport à trois travaux existants.

7.3.1 Paramètres des simulations

Nous considérons un ensemble de trois microgrids géographiquement dispersés qui sont chargés d'alimenter respectivement n_1, n_2 et n_3 de data centers homogènes avec leur énergie auto-produite et l'électricité achetée au smart grid. Tous les 2 heures, les microgrids et le smart grid mettent à jour leur prix de vente de l'énergie, et les data centers achètent une quantité d'énergie maximale de $q_{ij}^{max} = 8MWh$ auprès de leur microgrid correspondant.

Nous effectuons notre simulation sur une durée de 24 heures et mesurons la demande d'énergie du data center, le niveau de charge de l'énergie stockée des microgrids et les prix de vente de l'énergie à chaque cycle en deux périodes différentes : aux heures de pointe (de 6 à 10) et aux heures creuses (de 10 à 6). On note que les valeurs des coefficients de la fonction de gain sont fixées par le contrôleur du microgrid en fonction de ses besoins. Nous avons fixé $\alpha = 0,8$ afin d'inciter les data centers à acheter fortement l'énergie du microgrid. Nous avons également défini β, γ, δ avec la même valeur de coefficient

TABLE 7.2 – Paramètres des simulations

Paramètres	Valeurs
Horizon temporel	24h
Heures de pointe	[6am – 10pm]
Heures creuses	[10pm – 6am]
Itérations	1000
Nombre de microgrids (J)	3
Nombre des data centers (n_1, n_2, n_3)	3, 4, 2
q_j^{max}	8 MWh
PUE_j^t	[1.1 – 2]
S_j^{min}	2 MW
S_j^{max}	15 MW
Sl_j	7 MW
Pg_j, Pm_j, Ps_j	[125 – 170] \$/MWh
$\alpha, \beta, \gamma, \delta$	(0.7, 0.1, 0.1, 0.1)
$Wind_{ce}$	22.5(gCO_2e/kwh)
$Grid_{ce}$	968(gCO_2e/kwh)

afin d’avoir une importance égale pour toutes les métriques des data centers. Les paramètres de notre simulation sont détaillés dans le tableau 7.2.

Pour évaluer notre proposition, nous comparons dans un premier temps notre approche avec deux approches de répartition d’énergie :

- *L’approche traditionnelle* : est une méthode de répartition de l’énergie dans laquelle chaque microgrids vise à satisfaire les demandes d’énergie de tous ces data centers.
- *L’approche basée sur le prix* : nous supposons que cette approche tente d’attribuer la quantité d’énergie en fonction du prix de l’énergie du smart grid. Ainsi, nous supposons qu’elle alloue 90% de la quantité d’énergie du smart grid aux heures creuses et 70% aux heures de pointe. Le reste de la quantité d’énergie nécessaire est fourni par l’énergie verte.

Dans cette deuxième étape, nous comparons notre approche de minimisation des coûts énergétiques avec l’approche présentée dans l’un de nos travaux [15]. Cette approche est basée sur une architecture microgrid-cloud où le microgrid répartit équitablement l’énergie entre les data centers en fonction de leur consommation d’énergie verte, de l’utilisation de la bande passante du réseau et de la consommation d’énergie de leurs équipements réseau. Cette proposition ne prend pas en compte la tarification dynamique ou l’échange d’énergie entre les microgrids et nous l’appelons "Basic Game Scheme" **BGS**.

Dans nos simulations, nous fixons globalement les mêmes paramètres que ceux utilisés dans la [15]. Certaines valeurs de paramètres sont choisies de manière aléatoire, mais restent réalistes selon [110] [39]. La moyenne cumulée est donc prise en compte pour obtenir des résultats non biaisés.

7.3.2 Répartition de l’électricité entre les data centers

La figure 7.2 présente les variations des coûts des data centers du microgrid 01, sur une journée. Il faut noter que nous avons choisi le microgrid 01 de manière aléatoire. L’objectif est de mettre en évidence notre approche d’allocation d’énergie. Nous observons sur la figure 7.2-a que le coût de PUE

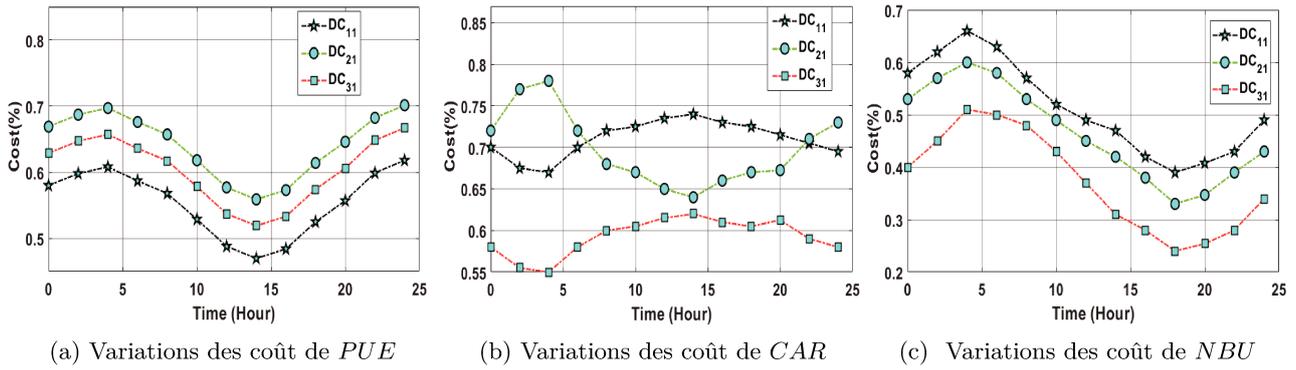


FIGURE 7.2 – Variations des coût des data centers du microgrid 01

des trois data centers est élevé en dehors des heures de pointe, lorsque les data centers sont moins efficaces en termes d'énergie et n'ont pas beaucoup de serveurs actifs en raison de la faible charge de travail durant cette période. Toutefois, le coût de *PUE* est faible aux heures de pointe, lorsque les data centers traitent un nombre plus élevé de demandes de leurs clients, ce qui implique un nombre élevé de serveurs actifs.

Dans notre simulation, nous avons défini un nombre beaucoup plus important d'applications en temps réel aux heures de pointe par rapport au nombre d'applications en cours d'exécution aux heures creuses. Par conséquent, le ratio des applications critiques en cours d'exécution dans les data centers par rapport au nombre total de leurs applications est plus élevé aux heures creuses qu'aux heures de pointe. En outre, nous pouvons observer que les coûts de *CAR* dans la figure 7.2-b varient correctement en fonction du nombre d'applications en temps réel de chaque data center par rapport au nombre total d'applications en temps réel tournant dans tous les data centers. En plus, les datacenters utilisent plus efficacement leur bande passante aux heures de pointe qu'aux heures creuses. Cela explique pourquoi, dans la figure 7.2-c, les coûts des *NBU* sont élevés aux heures creuses et faibles aux heures de pointe.

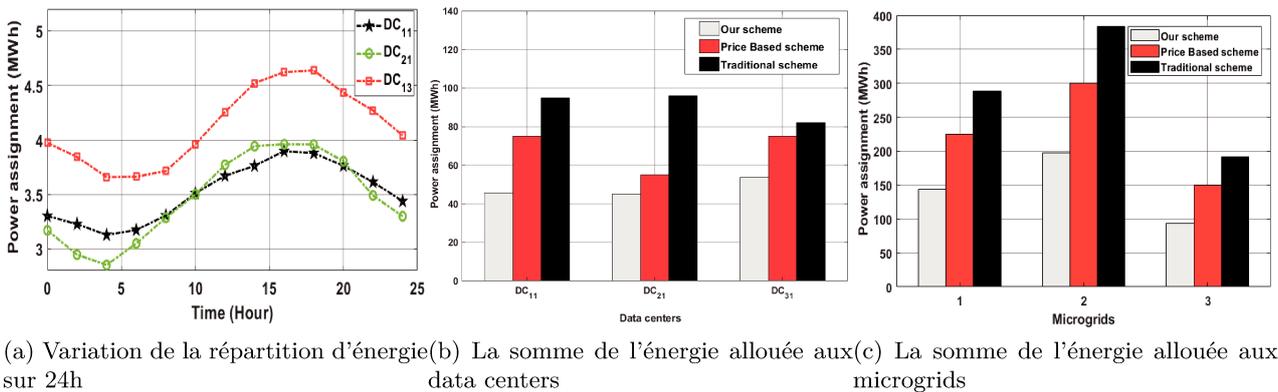


FIGURE 7.3 – Performances de notre approche de répartition d'énergie par rapport à l'approche traditionnelle et à celle basée sur le prix.

La figure 7.3-a donne la quantité d'énergie allouée à chaque data center du microgrid 01 pendant une journée en utilisant notre approche de jeu non coopératif. Nous remarquons que l'énergie assignée

varie dans le temps et qu'elle est faible aux heures creuses et élevée aux heures de pointe, suivant les variations des coût de leurs métriques décrites à la figure 7.2.

Pour mieux évaluer notre approche, nous comparons, dans la figure 7.3-b, la quantité d'énergie totale attribuée à chaque data center pendant une journée en utilisant notre approche avec celles attribuées par les deux autres approches, traditionnelle et basée sur le prix. Nous pouvons remarquer que la quantité d'énergie attribuée en utilisant notre approche est inférieure à celle attribuée en utilisant l'approche traditionnelle et l'approche basée sur le prix. Cela est dû à notre jeu non coopératif où le microgrid répartit équitablement l'énergie entre les data centers en pénalisant les data centers en fonction de leurs différentes métriques. Contrairement à notre approche, l'approche traditionnelle satisfait les premières demandes avec la quantité totale d'énergie demandée, sans tenir compte de la demande globale des data centers et de la capacité du microgrid. Cela peut être démontré en observant la quantité d'énergie qui a été attribuée au data center dc_{13} qui a besoin de plus d'énergie par rapport aux autres data centers. Ce comportement peut surcharger le réseau, provoquer des pannes et des pertes d'énergie. De plus, l'approche basée sur le prix réduit l'énergie allouée aux data centers par rapport à l'approche traditionnelle, mais elle ne tient pas compte des paramètres des data centers, et peut restreindre les autres data centers qui ont besoin de plus d'énergie que les autres. C'est pourquoi la quantité d'énergie qui a été attribuée à dc_{11} est supérieure à celle attribuée à dc_{12} en utilisant l'approche basée sur le prix. Les performances de notre approche sont également illustrées sur la figure 7.3-c qui représente la quantité totale d'énergie attribuée à chaque microgrid durant 24 heures. Nous pouvons voir que notre proposition diminue considérablement la quantité totale d'énergie assignée aux data centers par rapport aux autres algorithmes. De plus, notre méthode basée sur le jeu répartit équitablement l'énergie entre les data centers selon leurs mesures respectives, contrairement aux approches traditionnelles et basées sur le prix qui peuvent priver certains data centers d'avoir suffisamment d'énergie.

7.3.3 Minimiser le coût d'énergie des microgrids

Nous évaluons dans cette sous-section notre approche à deux niveaux de minimisation du coût énergétique, en la comparant avec notre précédente approche *BGS*.

La figure 7.4-a-b-c illustre, respectivement, la variation de l'énergie achetée en une journée par m_1 , m_2 et m_3 du smart grid et des microgrids. Nous pouvons voir que pour les trois microgrids, la quantité d'énergie achetée du réseau principale et des microgrids est faible aux heures creuses et élevée aux heures de pointe, suivant ainsi la répartition d'énergie de leurs data centers. En plus, l'énergie renouvelable achetée des microgrids par m_1 et m_3 est supérieure à celle achetée du réseau principal. Contrairement au microgrid m_2 qui a un nombre plus élevé de data centers à alimenter que m_1 et m_3 qui ont acheté beaucoup plus d'énergie du réseau principal, surtout aux heures de pointe où la demande d'énergie est élevée.

La quantité d'énergie totale achetée et vendue par les trois microgrids est présentée dans la figure 7.4-d. Nous remarquons qu'en général, l'énergie stockée des microgrids a été sollicitée de manière significative en parallèle avec l'énergie du réseau principal. Ceci est dû à notre algorithme de minimisation qui

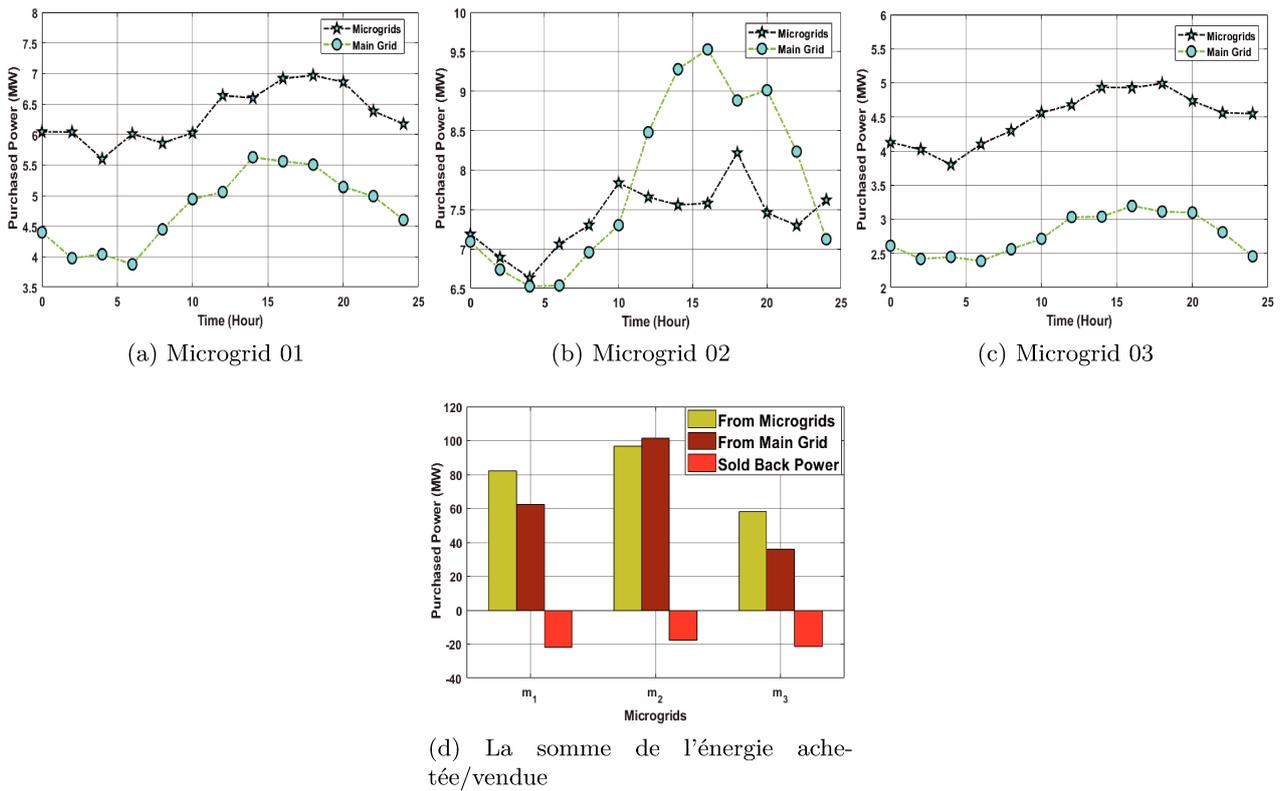


FIGURE 7.4 – Variations de l'énergie achetée du smart grid et des microgrids

prend en compte une variété de sources d'énergie et adopte une politique de tarification dynamique. L'énergie qui a été revendue par les trois microgrids est évidemment minime au regard du seuil fixé pour la vente Sl_j . On remarque que le microgrid m_2 a une quantité vendue d'énergie inférieure à m_1 et m_3 , car il puise rapidement ses réserves en raison de sa forte demande en énergie.

La figure 7.5 compare notre approche d'optimisation en deux étapes, nommée (*TSOA*), avec notre proposition précédente (*BGS*) en termes de coût d'énergie. Nous donnons la variation du coût de l'énergie achetée par chaque microgrid durant une journée en utilisant *TSOA* et *BGS*. Nous pouvons remarquer que le coût énergétique de chaque approche, dans chaque microgrid, est faible aux heures creuses lorsque les prix de l'énergie et les demandes d'énergie sont faibles, et élevé aux heures de pointe lorsque les prix d'énergie et les demandes d'énergie sont élevés. De plus, le coût énergétique de chaque microgrid utilisant notre approche est inférieur à celui de *BGS*. Ce dernier répartit l'énergie de manière optimale entre ses data centers, mais il n'utilise que deux sources d'énergie, son énergie auto générée et, de manière significative, l'énergie du réseau principal. Contrairement à *BGS*, notre approche (*TSOA*) répartit, dans un premier temps, équitablement l'énergie entre les data centers, et dans un deuxième temps, considère plusieurs sources d'énergie avec des prix plus variables. En effet, nous pouvons voir sur la figure 7.5-d que le coût énergétique total de tous les microgrids utilisant notre approche est inférieur de 25% à celui utilisant *BGS*. Cela s'explique par le fait que les prix de vente proposés par les microgrids sont, en général, beaucoup plus intéressants que ceux proposés par

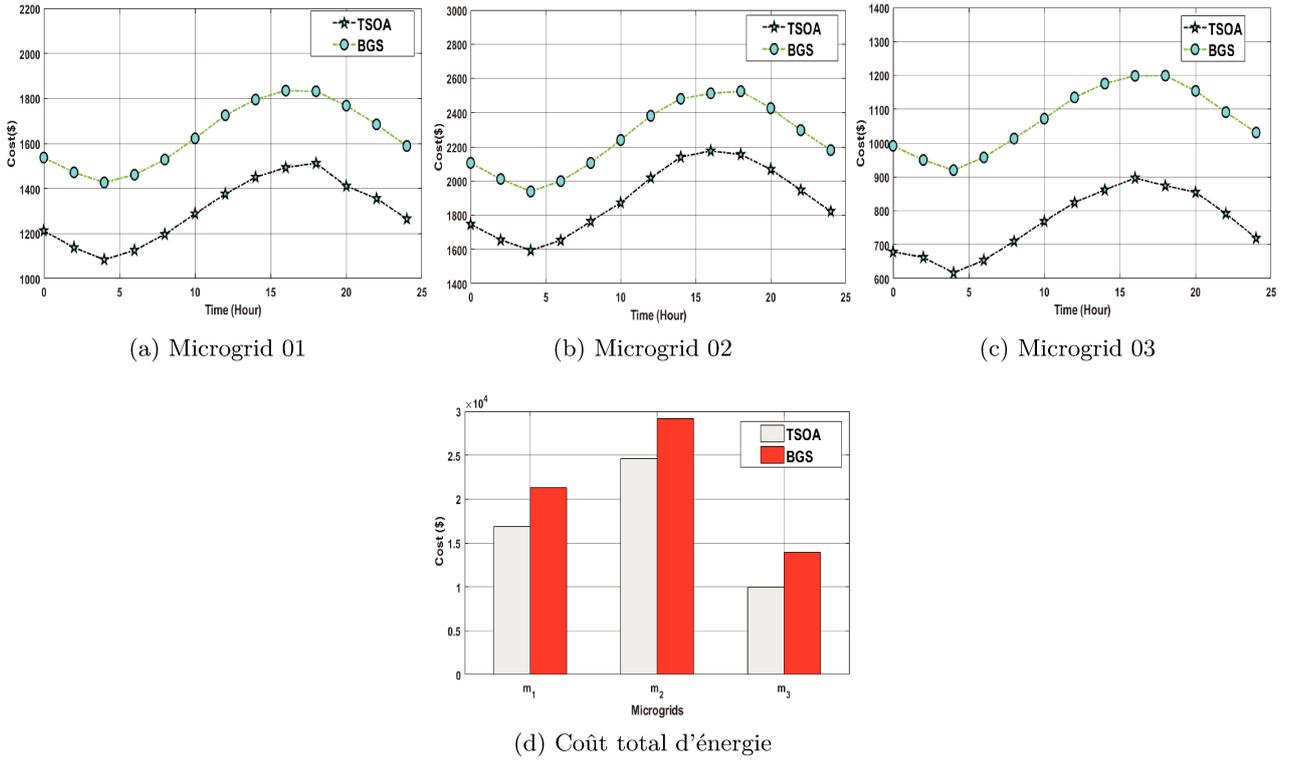


FIGURE 7.5 – Comparer le coût énergétique des microgrids en utilisant notre proposition et BGS.

le réseau principale. Cela favorise l'utilisation des énergies renouvelables tout en minimisant le coût total de l'énergie.

7.3.4 Les performances de notre approche en termes de taux de réduction de charge et des émissions de Co2

Dans cette sous-section, nous considérons deux paramètres pour évaluer notre approche en deux étapes : les émissions de dioxyde de carbone et le taux de charge électrique. Nous calculons l'émission de dioxyde de carbone DE_j de chaque micro-réseau à l'aide de la fonction suivante :

$$DE_j = \sum_{t=1}^T [(y g_j^t * Grid_{ce}) + \sum_{k=1}^m (y m_{jk}^t * Wind_{ce})] \quad (7.21)$$

Où $Grid_{ce}$ et $Wind_{ce}$ sont respectivement le taux d'émission de carbone CER [116] de l'énergie du grid (non renouvelable) et de l'énergie renouvelable des éoliennes présentés dans le tableau 8.2.

Nous présentons sur la figure 7.6-a le taux de charge électrique de chaque microgrid en utilisant notre approche d'optimisation en deux étapes ($TSOA$), BGS et l'approche basée sur le prix. Nous observons que, pour tous les microgrids, ($TSOA$) permet de réduire de manière significative la charge électrique du réseau principal avec un pourcentage de 52% par rapport à BGS , et de 59% par rapport à l'approche basée sur le prix. Ceci est dû à notre proposition qui permet d'acheter en premier lieu l'énergie la moins chère dans le système, favorisant ainsi l'utilisation de l'énergie stockée des microgrids

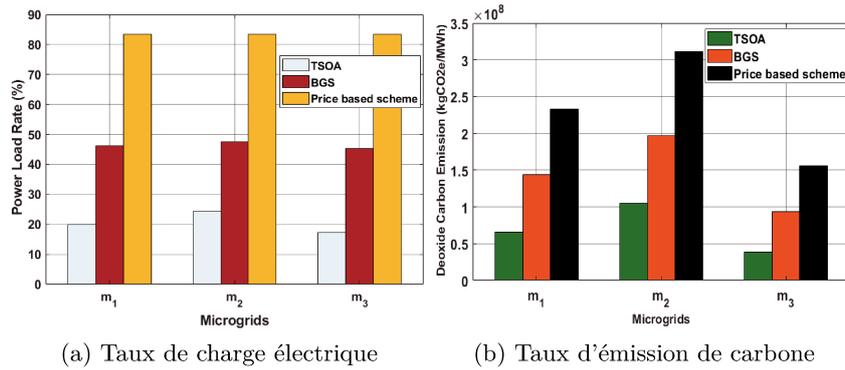


FIGURE 7.6 – Performances de notre approche à deux niveaux en termes des émissions de dioxyde de carbone et de taux de charge électrique

qui est généralement fournie avec des prix intéressants.

Réduire la charge électrique du réseau principal en achetant plus d'énergie verte des microgrids a un impact positif sur l'environnement. C'est ce que montre la figure 7.6-b qui illustre les performances de notre approche en termes d'émissions de dioxyde de carbone. Nous pouvons voir que notre approche est plus performante que *BGS* et l'approche basée sur le prix. Comparant à ces dernières, notre approche réduit considérablement les émissions de dioxyde de carbone avec un pourcentage, respectivement, de 51% et de 66%.

7.4 Conclusion

Nous avons proposé, dans ce chapitre, une approche en deux étapes qui minimise le coût énergétique des microgrids qui alimentent des data centers de cloud. Dans la première étape, les microgrids répartissent l'énergie de manière optimale et équitable entre leurs data centers. Nous modélisons la répartition de l'énergie des microgrids vers les data centers du cloud comme un jeu non coopératif, puis nous calculons la quantité optimale d'énergie à fournir à chaque data center selon ses paramètres définis. En fonction de la répartition de l'énergie entre les data centers, les microgrids essaient, dans la deuxième étape, de minimiser leur coût énergétique en achetant de manière optimale l'énergie des microgrids et du réseau principal, puis en revendant leur énergie stockée. Nous modélisons ce problème de minimisation des coûts énergétiques comme un problème d'optimisation linéaire sous contraintes et le résolvons à l'aide de l'outil d'optimisation de Matlab. Afin de valider notre proposition, nous l'avons comparée avec notre approche appelée "*BGS*" ainsi qu'avec l'approche traditionnelle et l'approche basée sur le prix. Les résultats obtenus ont montré que notre approche d'optimisation en deux étapes est 25% plus efficace en termes de coût énergétique. En plus, notre proposition encourage l'utilisation de l'énergie verte des microgrids qui réduit considérablement le taux de charge électrique et les émissions de gaz à effet de serre avec un pourcentage allant jusqu'à 52% par rapport à l'approche *BGS* et l'approche basée sur le prix.

Classer les nœuds Fog pour la répartition des requêtes des utilisateurs : une approche basée sur la logique floue

Sommaire

8.1	Introduction et motivations	102
8.2	Une approche de classification basée sur des quantificateurs flous	104
8.2.1	Quantificateurs linguistiques	104
8.2.2	Propositions quantifiées floues	104
8.2.3	Notre approche de classification	104
8.3	Classer les nœuds Fog pour la répartition des requêtes des utilisateurs	105
8.3.1	Architecture et modélisation du problème	106
8.3.2	Application de la méthode de classification pour la répartition efficace des requêtes	107
8.4	Simulations et résultats	107
8.4.1	Paramètres des simulations	107
8.4.2	Métriques d'évaluation de notre méthode	110
8.4.3	Niveau de satisfaction des utilisateurs	110
8.4.4	Performances de notre approche en termes de délais et de consommation énergétique	111
8.5	Conclusion	112

8.1 Introduction et motivations

La gestion des données générées par les capteurs et les objets connectés est l'un des plus grands défis à relever lors du déploiement d'un système IoT. Les systèmes IoT traditionnels sont confrontés à l'hétérogénéité et à la latence élevée de certains systèmes Cloud. Une solution consiste à décentraliser les applications, la gestion et l'analyse des données dans le réseau en utilisant un modèle distribué et fédéré. Cette approche est connue sous le nom de « fog computing ».

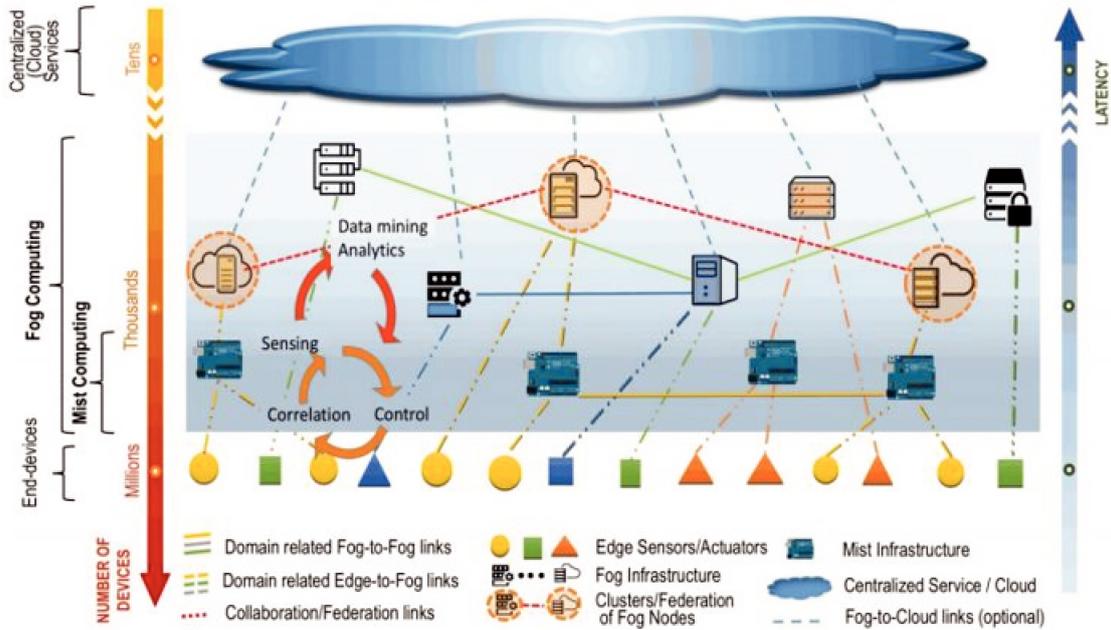


FIGURE 8.1 – Architecture fog-cloud (NIST 2018)

Le fog computing est une extension émergente de l'architecture du cloud computing qui permet de traiter les données au plus proche de leur source et met à disposition les connexions réseau nécessaires pour acheminer les données du nœud fog vers le terminal de l'utilisateur. Chaque nœud fog a des capacités de traitement qui lui permettent de répondre dans les brefs délais aux requêtes des objets qui sont connectés au réseau. Ainsi, en présence de plusieurs nœuds fog proches des extrémités du terminal, il est possible d'obtenir de meilleures performances en termes de délai que le cloud computing [26] [27]. La figure 8.1 représente d'architecture du fog computing vu par NIST [96]. Elle est présentée dans un contexte cloud plus large servant les appareils intelligents finaux. De ce point de vue, le fog computing ne vise pas à remplacer le cloud computing, mais à le compléter dans un nouveau paradigme fog-cloud computing, qui vise à satisfaire les applications des utilisateurs [37].

De plus en plus de terminaux délèguent leurs tâches à des nœuds fog voisins, situés à la périphérie du réseau, afin de réduire le délai de traitement. Toutefois, ce transfert de tâches nécessite une répartition efficace qui tient compte à la fois les préférences des utilisateurs et les exigences des nœuds fog. Une série de travaux de recherche a été menée pour étudier la répartition des tâches des utilisateurs

dans les réseaux fog-cloud. Pham et al [85] ont proposé une heuristique qui permet de minimiser à la fois le délai des requêtes des utilisateurs dans un environnement fog-cloud et aussi le coût monétaire des ressources du cloud. Yin et al [124] présentent un algorithme d'ordonnancement des tâches qui envoie les demandes des équipements industriels intelligents soit au nœud fog, soit au cloud. Le modèle proposé tient compte de la virtualisation des nœuds via des conteneurs. Les auteurs ont proposé un mécanisme qui réduit la durée des tâches tout en améliorant le nombre de tâches à traiter dans les nœuds fog.

En plus de prendre en compte le délai et le coût des communications, plusieurs chercheurs sont restés attentifs quant à la consommation d'énergie dans un environnement fog-cloud, et ont proposé des algorithmes de répartition des tâches énergétiquement efficaces. Wan et al [107] ont introduit un système multi-agents pour résoudre un problème d'ordonnancement des tâches dans une usine intelligente. Les nœuds fog sont utilisés dans l'usine pour répondre aux exigences en temps réel des applications industrielles. Les auteurs se sont concentrés sur le problème de la consommation d'énergie des équipements de l'usine et ont proposé une méthode d'équilibrage de charge et d'ordonnancement énergétiquement efficace appelée (ELBS). Tout d'abord, la méthode établit un modèle énergétique dans les nœuds fog utilisé pour calculer la relation entre la consommation d'énergie et la charge de travail de l'équipement industriel. Ensuite, une fonction d'optimisation d'équilibrage de charge a été modélisée et résolue en utilisant l'algorithme d'optimisation par essaims particulaires (PSO). Yang et al [122] ont proposé un algorithme de répartition des tâches dans un réseaux fog appelé (MEETS). Leur proposition vise à répartir les tâches tout en maximisant l'efficacité énergétique des nœuds fog, en tenant compte des contraintes de ressources de calcul des nœuds fog et le spectre inutilisé dans les environnements voisins. Zhang et al [132] ont présenté un algorithme de délégation des tâches appelé (FTO) qui sélectionne de manière équitable le nœud fog chargé de traiter la demande de l'utilisateur tout en maintenant un faible délai de traitement et en réduisant la consommation d'énergie correspondante.

Les approches présentées ci-dessus se concentrent principalement sur la réduction du délai et la consommation énergétique globale du system sans tenir compte des préférences des utilisateurs concernant les contraintes des nœuds fog. Les demandes des utilisateurs sont envoyées aux nœuds fog en fonction des propriétés qui ne concernent que les nœuds fog. Ces méthodes de répartition des tâches sont généralement utilisées dans un environnement fog-Cloud avec des exigences de délai et de tarification [65]. Cependant, l'environnement fog-cloud représente un réseau hétérogène qui traite les requêtes de différents types d'utilisateurs. Ces derniers, doivent exprimer leurs préférences en sélectionnant le meilleur nœud fog pour exécuter leurs demandes [125].

Pour une deuxième contribution, nous considérons la nécessité d'une méthode efficace de répartition des tâches dans une architecture fog-cloud. Notre proposition vise à envoyer les requêtes des utilisateurs au meilleur nœud fog, tout en satisfaisant leurs préférences et les contraintes des nœuds fog. Compte tenu du besoin de scalabilité et du caractère dynamique des prix d'un système fog computing, les nœuds fog et les utilisateurs ont des exigences diverses. Pour résoudre ce problème, nous utilisons la logique floue pour sa capacité à couvrir un large éventail de conditions de fonctionnement et à exprimer les préférences des utilisateurs sous forme de quantificateurs flous.

Dans ce travail, nous avons utilisé des quantificateurs linguistiques et des propositions quantifiées floues regroupant efficacement les préférences des utilisateurs et les contraintes des nœuds de fog afin

de classer les nœuds fog du plus satisfaisant au moins satisfaisant. Pour distinguer les similitudes, nous avons utilisé les paramètres appliqués dans [97], appelés : proportion la moins satisfaisante (lsp) et proportion la plus satisfaisante (gsp).

8.2 Une approche de classification basée sur des quantificateurs flous

Nous résumons dans cette partie les éléments de base de la théorie des ensembles flous et de la quantification floue. Nous présentons également les propositions quantifiées floues de la forme “ $Q X$ sont A ” et notre approche basée sur la décomposition.

8.2.1 Quantificateurs linguistiques

Les quantificateurs linguistiques flous sont des modèles mathématiques qui représentent les quantités linguistiques humaines tels que : *plusieurs, la plupart, beaucoup, peu, assez, grand, faible, élevé, etc.*

Notons Q est un quantificateur linguistique, il est défini comme un sous-ensemble flou de deux types de quantificateurs : **absolu** ou **relatif**. Les quantificateurs absolus tels que “*au moins 9*”, “*presque 2*” et “*environ 5*” sont définis de \mathbb{R} ou \mathbb{N} à $[0; 1]$ comme des ensembles flous avec des fonctions d’appartenance et leurs interprétations ne sont pas liées à la cardinalité de l’ensemble. Les quantificateurs relatifs tels que “*peu*”, “*presque tous*” et “*environ la moitié*” expriment une proportion.

8.2.2 Propositions quantifiées floues

Soit Q un quantificateur linguistique, X un ensemble d’éléments et A, B deux prédicats flous. Il existe deux formes de propositions quantifiées floues : “ $Q X$ sont A ” ou “ $Q B A$ sont A ”. La première signifie qu’il existe des éléments Q de l’ensemble X qui satisfont le prédicat flou A , et la seconde signifie qu’à partir des éléments de X qui satisfont B il existe des éléments Q qui satisfont A . Nous pouvons dire qu’un quantificateur linguistique Q dans une proposition quantifiée floue est croissant (resp. décroissant) si la valeur de vérité de la proposition ne diminue pas (resp. n’augmente pas). Par exemple, “*au moins 9*” est un chiffre croissant, et “*au plus 9*” est un chiffre décroissant.

Dans ce travail, nous utilisons la proposition quantifiée de type “ $Q X$ sont A ”, et selon [19], il existe plusieurs approches exprimant les préférences des utilisateurs qui favorisent l’utilisation de cette forme de proposition quantifiée. On peut citer parmi ces méthodes : La méthode basée sur l’intégrale floue (Choquet/Sugeno), la méthode basée sur OWA, et la méthode basée sur la décomposition. Comme nous considérons un système de communication en temps réel, nous avons choisi d’adopter l’approche basée sur la décomposition [118] en raison de sa complexité polynomiale.

8.2.3 Notre approche de classification

La proposition quantifiée floue ayant la forme “ $P : Q X$ sont A avec un quantificateur croissant Q est évaluée par l’approche basée sur la décomposition. Sa valeur de vérité notée par δ_P est exprimée

comme suit :

$$\delta_P = \begin{cases} \max_{i=1}^n (\min(\mu_A(x_i), \mu_Q(i))) & \text{si Q est absolu} \\ \max_{i=1}^n (\min(\mu_A(x_i), \mu_Q(\frac{i}{n}))) & \text{si Q est relatif} \end{cases} \quad (8.1)$$

Où

$$\mu_A(x_1) \geq \mu_A(x_2) \geq \dots \geq \mu_A(x_n) \quad (8.2)$$

La règle de classification est exprimée comme suit : “L’alternative la mieux classée δ_P sera celle qui est la plus proche de 1 “. Nous définissons la relation de préférence d’ordre total comme suit :

Définition 1 : Supposons respectivement les deux alternatives F_j and $F_{j'}$. Alors :

$$F_j \succeq F_{j'} \iff \delta(F_j) \geq \delta(F_{j'})$$

Classer les cas similaires

Dans le cas où nous avons des propositions quantifiées floues qui ont des valeurs de vérité similaires, tout autre critère supplémentaire n’améliorera pas le degré de satisfaction [23]. Ainsi, nous utilisons l’approche basée sur la décomposition et traduisons les résultats renvoyés comme le degré alternatif suprême pour une alternative x , dans le sous-ensemble de caractéristiques qui satisfait la proposition quantifiée floue et sa valeur correspondante à l’ensemble des caractéristiques. Ainsi, nous pouvons avoir des sous-ensembles de caractéristiques maximum/minimum, qui maximisent/minimisent le degré de satisfaction. Nous définissons les paramètres *la plus grande proportion de satisfaction* et *la plus petite proportion de satisfaction* (désignés respectivement par *gsp* et *lsp*) comme suit :

Définition 2 : Si " $P : Q C_X$ sont A " est une proposition quantifiée floue définie sur des attributs de X . Nous disons que *gsp* (resp. *lsp*) est la proportion du plus grand (resp. du plus petit) sous-ensemble de C_X qui satisfait A . Formellement,

$$gsp = \frac{\max(k : \delta_P = \min(\mu_A(C_k^{\geq}), \mu_Q(\frac{k}{|C_X|})))}{|C_X|} \quad (8.3)$$

$$lsp = \frac{\min(k : \delta_P = \min(\mu_A(C_k^{\geq}), \mu_Q(\frac{k}{|C_X|})))}{|C_X|} \quad (8.4)$$

Notons x_1, x_2 deux éléments de X . On dit que $x_1 \preceq x_2$ si et seulement si $lsp(x_1) < lsp(x_2) \vee (lsp(x_1) = lsp(x_2)) \wedge gsp(x_1) \leq gsp(x_2)$. On dit aussi que $x_1 \succeq x_2$ iff $lsp(x_1) > lsp(x_2) \vee (lsp(x_1) = lsp(x_2)) \wedge (gsp(x_1) \geq gsp(x_2))$.

8.3 Classer les nœuds Fog pour la répartition des requêtes des utilisateurs

Cette section décrit notre système et notre approche de classement utilisés pour la planification des tâches dans un réseau fog-cloud.

8.3.1 Architecture et modélisation du problème

Un système fog-cloud computing a été modélisé sur la base de l'architecture fog computing proposé par Cisco [56]. La figure 8.2 présente notre architecture. Notre système traite avec un fournisseur de fog computing qui dispose d'une partie des ressources louées dans le cloud, qu'il l'utilise pour répondre à ses besoins de calcul et de stockage. Les utilisateurs dotés d'une faible capacité de calcul envoient leurs requêtes de service à plusieurs nœuds fog qui ont des ressources plus performantes. Ces demandes sont acheminées vers le nœud fog via des réseaux sans fil (LoRa, WiFi et Bluetooth, par exemple). Un planificateur de tâches est chargé de collecter les demandes et de les distribuer aux nœuds fog. Il envoie également périodiquement des résumés au cloud, qui à son tour analyse ces données et peut envoyer de nouvelles directives aux nœuds fog pour des fins stratégiques. Nous considérons un ensemble de nœuds

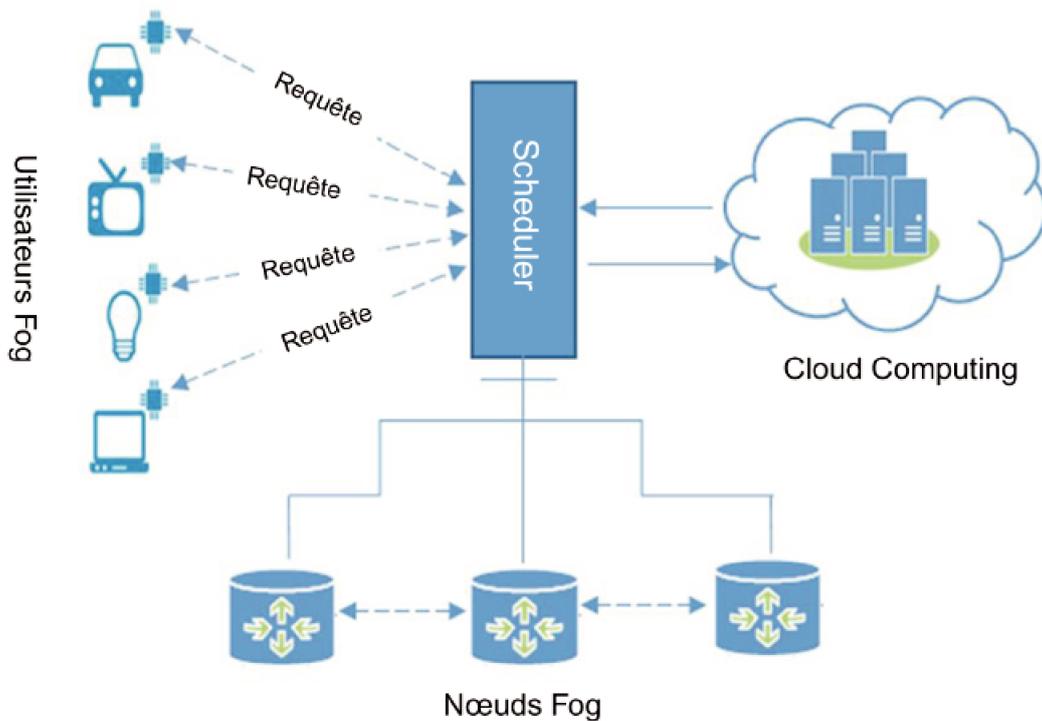


FIGURE 8.2 – Architecture de notre système

Fog J . Chaque nœud fog $F_j, j \in \{1, \dots, |J|\}$ dans le réseau a plusieurs propriétés : (i) la distance entre le nœud fog et l'utilisateur final, (ii) le prix de service, (iii) la latence, (iiii) le débit, (iiiii) le degré de fiabilité.

Considérons un ensemble I d'utilisateurs, chaque utilisateur $U_i, i \in \{1, \dots, |I|\}$, envoie ses demandes au planificateur de tâches, comprenant ses préférences sur les propriétés des nœuds fog. Le planificateur de tâches classe les nœuds fog en fonction des préférences des utilisateurs et des contraintes relatives aux nœuds fog. Il sélectionne ensuite le nœud le plus approprié pour exécuter la tâche de l'utilisateur.

8.3.2 Application de la méthode de classification pour la répartition efficace des requêtes

Nous modélisons un nœud fog F_j comme un ensemble de couples $K (m_{jk}, v_{jk})$, de sorte que $m_{jk}, k \in \{1, \dots, K\}$ est un critère de service de nœud fog, et v_{jk} est sa valeur correspondante. Comme pour le nœud fog, l'utilisateur U_i est modélisé comme un ensemble de couples $L (m_{il}, p_{il})$ où $m_{il}, l \in \{1, \dots, L\}$, est une caractéristique de service, et p_{il} représente la préférence de l'utilisateur exprimée sur cette caractéristique de service (m_{il}). Nous indiquons par $\delta(F_j \setminus i)$ le degré de satisfaction de l'utilisateur à l'égard du service U_i concernant un nœud fog F_j . Il représente la valeur de vérité de la proposition quantifiée floue qui utilise le quantificateur linguistique "Presque tout" comme : "Presque toutes les préférences sont satisfaites". La valeur de vérité est modélisée sur la base de la formule (1) comme suit :

$$\delta(F_j \setminus i) = \max_{l=1}^L (\min(\mu_p(m'_{il}), \mu_{almostall}(\frac{l}{L})) \quad (8.5)$$

Où $\mu_p(m'_{il})$ est le degré de satisfaction de (m'_{il}) exprimé à la préférence p , et le classement ascendant des critères basé sur $\mu_p(m_{il})$ est arrangé comme suit :

$$\mu_p(m'_{i1}) \geq \mu_p(m'_{i2}) \geq \dots \geq \mu_p(m'_{iL}) \quad (8.6)$$

Nous supposons que chaque utilisateur peut être servi en fonction de cinq critères de nœud fog : distance de l'utilisateur, prix du service, latence, bande passante et fiabilité. Ensuite, les préférences de l'utilisateur sont exprimées selon ces critères et traduites en fonctions d'appartenance floue. Les figures 3(a)- 3(e) illustrent les fonctions d'appartenance floue de l'une des trois classes d'utilisateurs considérées dans nos expérimentations.

8.4 Simulations et résultats

Dans cette sous-section, nous évaluons les performances de notre approche de classification par le biais des simulations numériques réalisées avec Matlab. Nous présentons d'abord les paramètres utilisés dans nos expérimentations et les mesures de performance, puis nous discutons les performances de notre méthode de répartition des tâches par rapport à d'autres algorithmes.

8.4.1 Paramètres des simulations

Nous considérons dans nos simulations un réseau avec un nombre maximum $|J| = 20$ de nœuds fog déployés de manière aléatoire. La distance maximale de communication entre un terminal et un nœud fog est de 100 m . Nous effectuons nos simulations avec un nombre constant de tâches d'utilisateurs $nb_{tâches}$ qui ont une taille de paquet constante. Les paquets sont transmis dans un réseau avec une bande passante maximale de 10 $Mbit/s$. La consommation d'énergie en mode veille de l'appareil de l'utilisateur E_{cidle} et la consommation d'énergie d'une tâche dans chaque ms (σ_{ms}) sont définies en se basant sur le travail présenté dans [21]. Les paramètres de notre simulation sont détaillés dans le tableau I.

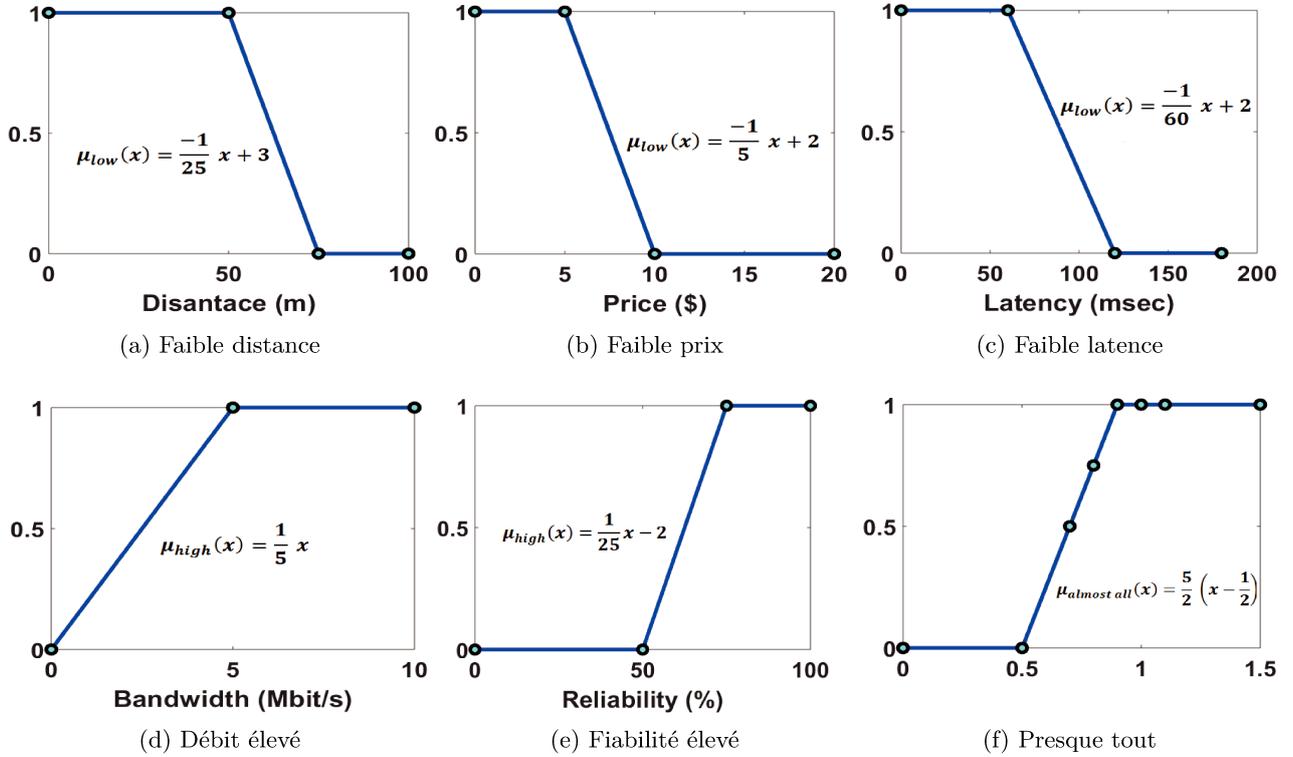


FIGURE 8.3 – Exemples de fonctions d'appartenance floue

TABLE 8.1 – Paramètres des simulations

Paramètres	Valeurs
Noeuds fog	20
Nombre de paquets (nb_{tasks})	500
Taille des paquets $Mbits$	[0.01, 0.03]
$E_{c_{idle}}(mW)$	0.4
$\sigma_{ms}(mW/ms)$	[0.05, 0.07]
Distance (de l'utilisateur au noeud fog) (m)	[0, 100]
Débit ($Mbit/s$)	[0, 10]

TABLE 8.2 – Les ensembles fog et les valeurs de leurs critères

Les ensembles fog	Nombre de noeuds fog	Distance (m)	Prix de service (\$)	Latency (ms)	Débit (Mbit/s)	Fiabilité (%)
Ensemble 01	10	[30,80]	[4,12]	[40,100]	2-10	[30,100]
Ensemble 02	15	[40,70]	[3,15]	[40,200]	2-10	[30,100]
Ensemble 03	20	[20,100]	[7,18]	[50,180]	2-10	[30,100]

TABLE 8.3 – Fonctions d'appartenances des classes des utilisateurs

Classe utilisateur	Distance(m)	Prix (\$)	Latence (ms)	Débit (Mbit/s)	Fiabilité (%)
Classe 01	$\begin{cases} 0 & \text{si } x \geq 75 \\ 1 & \text{si } x \leq 50 \\ \frac{-1}{25}x + 3 & \text{sinon} \end{cases}$	$\begin{cases} 0 & \text{si } x \geq 10 \\ 1 & \text{si } x \leq 5 \\ \frac{-1}{5}x + 2 & \text{sinon} \end{cases}$	$\begin{cases} 0 & \text{si } x \geq 120 \\ 1 & \text{si } x \leq 60 \\ \frac{-1}{60}x + 2 & \text{sinon} \end{cases}$	$\begin{cases} 0 & \text{si } x = 0 \\ 1 & \text{si } x \leq 5 \\ \frac{1}{5}x & \text{sinon} \end{cases}$	$\begin{cases} 0 & \text{si } x \leq 50 \\ 1 & \text{si } x \geq 75 \\ \frac{1}{25}x - 2 & \text{sinon} \end{cases}$
Classe 02	$\begin{cases} 0 & \text{si } x \geq 60 \\ 1 & \text{si } x \leq 40 \\ \frac{-1}{20}x + 3 & \text{sinon} \end{cases}$	$\begin{cases} 0 & \text{si } x \geq 12 \\ 1 & \text{si } x \leq 6 \\ \frac{-1}{6}x + 2 & \text{sinon} \end{cases}$	$\begin{cases} 0 & \text{si } x \geq 100 \\ 1 & \text{si } x \leq 50 \\ \frac{-1}{50}x + 2 & \text{sinon} \end{cases}$	$\begin{cases} 0 & \text{si } x = 0 \\ 1 & \text{si } x \leq 4 \\ \frac{1}{4}x & \text{sinon} \end{cases}$	$\begin{cases} 0 & \text{si } x \leq 40 \\ 1 & \text{si } x \geq 60 \\ \frac{1}{20}x - 2 & \text{sinon} \end{cases}$
Classe 03	$\begin{cases} 0 & \text{si } x \geq 75 \\ 1 & \text{si } x \leq 50 \\ \frac{-1}{25}x + 3 & \text{sinon} \end{cases}$	$\begin{cases} 0 & \text{si } x \geq 8 \\ 1 & \text{si } x \leq 4 \\ \frac{-1}{4}x + 2 & \text{sinon} \end{cases}$	$\begin{cases} 0 & \text{si } x \geq 140 \\ 1 & \text{si } x \leq 70 \\ \frac{-1}{70}x + 2 & \text{sinon} \end{cases}$	$\begin{cases} 0 & \text{si } x = 0 \\ 1 & \text{si } x \leq 8 \\ \frac{1}{8}x & \text{sinon} \end{cases}$	$\begin{cases} 0 & \text{si } x \leq 35 \\ 1 & \text{si } x \geq 70 \\ \frac{1}{35}x - 1 & \text{sinon} \end{cases}$

Pour montrer l'efficacité de notre approche, nous exécutons nos simulations en utilisant différents cas de propriétés de nœuds fog. Nous modélisons ces cas sous forme d'ensembles dont nous définissons trois : *Ensemble 1*, *Ensemble 2* et *Ensemble 3*. Chaque ensemble a un nombre différent de nœuds fog et des valeurs de critères différentes. Le tableau II énumère toutes les valeurs correspondants aux critères des trois ensembles de nœuds fog.

Du côté des utilisateurs, nous définissons trois classes d'utilisateurs différentes : *Classe 1*, *Classe 2* et *Classe 3*. Chaque classe d'utilisateur a différentes valeurs de préférence exprimées sous forme de fonctions d'appartenance sur les cinq critères des nœuds fog : "*Distance, Prix de service, Débit, Latence et Fiabilité*". Le tableau III détaille les fonctions d'appartenance de ces trois classes d'utilisateurs.

Nous avons comparé notre méthode de classification des nœuds fog pour répartir les tâches des utilisateurs avec trois autres méthodes de répartition des tâches.

- **Une approche basée sur la distance** : sélectionne le nœud fog le plus proche de l'utilisateur.
- **Une approche basée sur le prix** : sélectionne le nœud fog le moins cher.
- **Une approche basée sur la latence** : sélectionne le nœud fog qui a la plus faible latence.

8.4.2 Métriques d'évaluation de notre méthode

Afin d'évaluer notre approche de classification et de répartition des tâches des utilisateurs, nous avons pris en compte trois paramètres : (i) la satisfaction de l'utilisateur (ii) le temps d'exécution, et (iii) la consommation d'énergie de l'utilisateur. Ils sont définis comme suit :

1. **La satisfaction de l'utilisateur** : calcule le degré de satisfaction moyen de l'utilisateur sur tous les critères du nœud fog sélectionné. Nous définissons \overline{U}_{si} la satisfaction moyenne d'un utilisateur i comme suit :

$$\overline{U}_{si} = \sum_{l=1}^L \frac{\mu_p(m'_{il})}{L} \quad (8.7)$$

2. **Temps d'exécution** : noté E_x , mesure le temps entre l'envoi d'une requête et la réception de la réponse. Ainsi, le temps d'exécution est la somme du temps de calcul $T_{compute}$ et du temps de transmission de la requête $T_{transmission}$:

$$E_x = T_{compute} + T_{transmission} \quad (8.8)$$

Où :

$$T_{transmission} = \left(\frac{Packetsize_{req} + Packetsize_{res}(bit)}{Bandepassante(\frac{bit}{s})} \right) \quad (8.9)$$

Où $Packetsize_{req}$ et $Packetsize_{res}$ représentent la taille du paquet envoyé et reçu, respectivement.

3. **Consommation électrique** : mesure la consommation électrique d'un appareil de l'utilisateur dans l'exécution d'une tâche, nous la modélisons comme suit :

$$E_c(mW) = E_{cidle} + (E_x \times \sigma_{ms}) \quad (8.10)$$

Où E_{cidle} est la consommation d'énergie d'un dispositif utilisateur en veille et σ_{ms} représente la consommation d'énergie d'une tâche dans chaque ms pendant le temps de son exécution.

8.4.3 Niveau de satisfaction des utilisateurs

La figure 8.4 compare la satisfaction moyenne des utilisateurs obtenue avec notre approche par rapport aux autres approches de comparaison dans chaque classe d'utilisateurs, tout en changeant les ensembles de nœuds fog. Les résultats montrent que notre approche de classification satisfait davantage les utilisateurs en comparant avec les algorithmes, car elle tient compte de tous les critères de préférence des utilisateurs. De plus, en présence d'un nombre plus élevé de nœuds fog, les utilisateurs se trouvent face à davantage de caractéristiques de nœuds fog plus flexibles qui pourraient correspondre leurs préférences. C'est pourquoi la satisfaction des utilisateurs en utilisant toutes les approches dans chaque classe d'utilisateurs s'améliore à chaque fois que nous changeons les ensembles de nœuds fog. Cependant, cela implique un temps d'exécution plus élevé pour toutes les approches utilisées, comme le montre la figure 8.5.

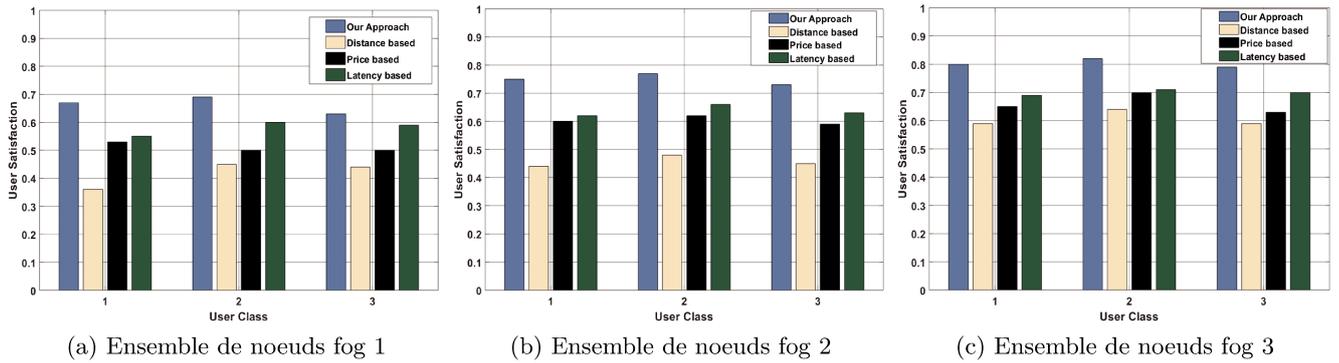


FIGURE 8.4 – Satisfaction des utilisateurs en utilisant notre approche par rapport aux approches basé sur la distance, le prix et à la latence dans chaque ensemble de nœuds fog et pour chaque classe d’utilisateurs

8.4.4 Performances de notre approche en termes de délais et de consommation énergétique

La figure 8.5 illustre le temps d’exécution moyen des quatre approches de répartition des tâches définis dans chaque classe d’utilisateurs, tout en faisant varier les ensembles de nœuds fog. Elle montre que les approches basées sur la latence et la distance permettent d’obtenir un temps d’exécution plus faible dans chaque ensemble de nœuds fog par rapport aux deux autres approches. Cela peut s’expliquer par le fait qu’elles sélectionnent le nœud fog ayant, respectivement, les valeurs de latence et de distance les plus faibles, améliorant ainsi le temps d’exécution moyen. Par exemple, notre proposition prend un peu plus de temps d’exécution parce qu’elle nécessite des fonctions de calcul supplémentaires. Cependant, notre approche surpasse l’approche basée sur le prix en termes de temps d’exécution parce qu’elle tient compte des critères de distance et de délai, alors que l’approche basée sur le prix ne les prend pas en considération. En outre, nous observons que dans chaque ensemble de nœuds fog, la classe d’utilisateur 2 offre un meilleur temps d’exécution que les classes d’utilisateur 1 et 3, car elle a des préférences plus adéquates concernant les valeurs des critères des trois ensembles de nœuds fog. Nous avons choisi au hasard un terminal d’utilisateur et avons tracé sur la figure 8.6 sa consommation

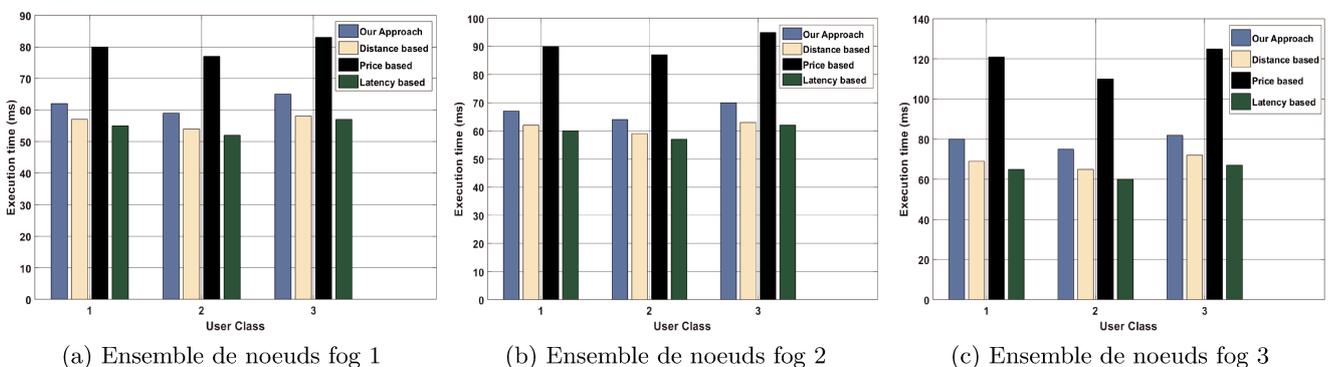


FIGURE 8.5 – Temps d’exécution de notre approche par rapport aux autres approche tout en variant la classe d’utilisateurs et l’ensemble des nœuds fog

d'énergie pendant son temps d'exécution. Nous avons comparé la consommation d'énergie des quatre approches tout en variant les ensembles de nœuds fog. Il convient de noter que l'approche basée sur la latence est la plus efficace en termes de consommation d'énergie par rapport aux autres approches, en raison de son faible temps de calcul et de transmission. Notre approche comporte des étapes supplémentaires, ce qui consomme plus d'énergie en comparant aux autres approches. Comme dans la figure 8.5, la consommation d'énergie en utilisant toutes les approches augmente au fur et à mesure que le nombre des nœuds fog augmentent.

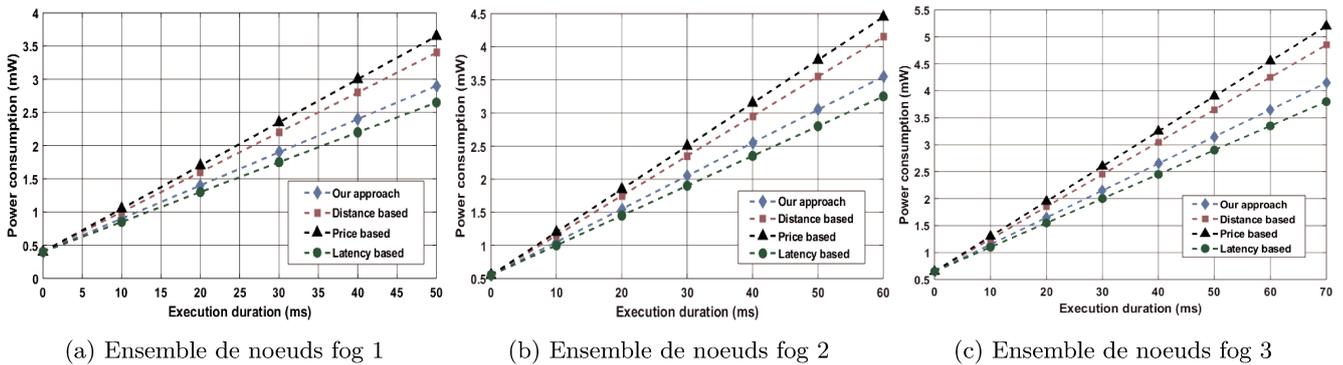


FIGURE 8.6 – Performances de notre approche en termes de consommation d'énergie par rapport aux autres approches

En général, notre approche de classification des nœuds fog pour répartir les tâches des utilisateurs dans un environnement fog-cloud donne une bonne satisfaction aux utilisateurs, mais les performances sont moins bonnes en termes de temps d'exécution et de consommation d'énergie. Cependant, notre proposition offre une meilleure solution de compromis entre la satisfaction moyenne des utilisateurs, le temps d'exécution et la consommation électrique, par rapport aux autres approches de comparaison.

8.5 Conclusion

Dans ce travail, nous avons proposé une méthode de classification quantitative floue pour la répartition des tâches dans les réseaux fog-cloud. Notre approche effectue la répartition des tâches en classant les nœuds fog en fonction de leurs exigences, puis en sélectionnant le nœud fog qui satisfait le plus les préférences de l'utilisateur. Des quantificateurs linguistiques et des propositions quantifiées floues ont été utilisés pour agréger en un temps polynomial les préférences de l'utilisateur et les caractéristiques des nœuds fog. Les résultats de la simulation ont montré que notre proposition permet de répartir les tâches tout en satisfaisant les préférences de l'utilisateur. Notre approche fournit également une solution de compromis entre la satisfaction moyenne de l'utilisateur, le délai d'exécution et la consommation d'énergie.

Conclusion générale et perspectives

Les chercheurs et les industriels portent un grand intérêt à la réduction des coûts d'exploitation et de consommation d'énergie ainsi que les émissions de carbone des data centers. Avec le développement du smart grid, ces efforts ont pris une autre dimension en intégrant des concepts visant à améliorer l'efficacité énergétique des data centers. Nous avons, tout le long de cette thèse, étudié le comportement énergétique des data centers du cloud et répondu aux besoins identifiés en gestion d'énergie des data centers dans un système smart grid.

Nous avons, au début de ce manuscrit, résumé les principes fondamentaux et l'architecture du smart grid. Ensuite, nous avons détaillé l'architecture de communication du smart grid et expliqué ses principales caractéristiques.

Nous avons expliqué dans le chapitre 3 la gestion d'énergie des data centers du cloud. Nous avons donné dans un premier temps les principes fondamentaux du cloud computing et avons mis l'accent sur le rôle des data centers dans cette technologie. Puis, nous avons présenté les principaux composants des data centers et leurs comportements énergétiques. Nous avons expliqué à la fin de ce chapitre l'interaction des data centers du cloud avec le smart grid.

Nous avons identifié dans le chapitre 4 les besoins et les solutions existantes dans la gestion d'énergie des data centers. Nous avons présenté les différents travaux qui ont été menés dans ce contexte en termes d'architectures, algorithmes et outils d'aide à la décision.

Notre première contribution a été l'objet du chapitre 5 où nous avons présenté notre approche de répartition équitable d'énergie entre les data centers dans un système smart grid - cloud. Une extension de ce travail a été présentée dans le chapitre 6. Nous avons considéré dans ce chapitre les énergies renouvelables pour répartir équitablement l'électricité entre les data centers du cloud dans un système smart grid - cloud. La répartition a été faite en fonction de trois critères : la consommation d'énergie renouvelable, le nombre d'applications critiques en cours d'exécution et la charge de travail des data centers. Nous avons comparé notre proposition avec un travail de la littérature, et nous avons montré les performances de notre approche en termes d'équité, réduction de taux de charge d'électricité et de taux d'émission de CO₂.

Pour mettre en avant le rôle des microgrids dans notre système, nous avons présenté dans le chapitre 7 une approche de minimisation à deux niveaux du coût d'énergie des microgrids dans une architecture

microgrid - cloud. Nous avons montré l'efficacité et les performances de notre approche à deux niveaux en la comparant avec l'approche présentée dans le chapitre précédent.

Enfin, nous avons considéré dans le chapitre 8 la nécessité d'une méthode efficace de répartition des tâches dans une architecture fog-cloud. Nous avons présenté dans ce chapitre, notre architecture fog - cloud et nous avons présenté les éléments de base de la théorie des ensembles flous et de la quantification floue. Nous avons détaillé dans ce chapitre notre approche basée sur la décomposition qui sélectionne le noeud fog qui satisfait le mieux les préférences des utilisateurs et les contraintes des noeuds fog.

Les résultats présentés dans cette thèse ouvrent la voie à de nouvelles pistes à étudier qui peuvent encore améliorer le travail réalisé. Nous citons dans ce qui suit quatre propositions qui peuvent être considérées dans les travaux à venir :

Perspective 1 : *Combiner les approches classiques et nouvelles dans un cadre global d'efficacité énergétique*

Plusieurs approches proposées dans la littérature ont été traitées indépendamment du contexte smart grid - cloud. De ce fait, ces approches peuvent entrer en conflit les unes avec les autres lorsqu'elles sont intégrées dans de véritables systèmes smart grid - cloud, ce qui les rend inefficaces. Par conséquent, l'une des orientations futures consiste à combiner les stratégies d'efficacité énergétique à tous les niveaux (utilisateurs, edge computing, cloud computing, data centers et smart grid) et à former un cadre global d'efficacité énergétique.

Perspective 2 : *Interagir avec d'autres systèmes pour réduire la charge électrique*

En raison de leur énorme consommation d'énergie, les data centers devraient avoir un impact majeur sur le réseau électrique en augmentant considérablement la charge aux endroits où ils sont installés. Cependant, les data centers offrent également la possibilité d'aider le réseau dans sa robustesse et l'équilibrage de la charge électrique. Par exemple, comme les data centers sont des charges électriques importantes et flexibles, ils peuvent être de bons candidats pour offrir des services auxiliaires, tels que la réduction volontaire de la charge au smart grid. En outre, les data centers peuvent mieux stabiliser le prix de l'énergie et réduire le coût d'énergie en exploitant la diversité des prix sur le marché.

Perspective 3 : *Considérer la perte d'énergie dans un système smart grid - cloud*

Un troisième point à prendre en considération est la perte d'énergie. Plusieurs types de stockage d'énergie pourraient en outre être combinés et distribués en fonction de leurs propres caractéristiques, ce qui pourrait contribuer à réduire les pertes de transmission d'énergie et à mieux réguler la puissance des data centers.

Perspective 4 : *Interopérabilité, sécurité et décentralisation*

Enfin, plusieurs points stratégiques restent à étudier et à développer dans un système smart grid - cloud notamment :

- **L'interopérabilité** : pour assurer la convergence des protocoles réseaux et de transmissions en vue d'une coopération et d'une communication renforcée.
- **La fiabilité et la sécurité** : pour une fourniture de services fiable et résiliente.
- **L'architecture décentralisée et auto-organisée** : pour une flexibilité accrue dans le contrôle et la gestion du réseau.

Démonstration du jeu non-coopératif

Avant de calculer la solution de notre jeu, il est nécessaire de démontrer que notre jeu non-coopératif a une solution. Par conséquent, nous prouvons dans cette section l'existence d'un équilibre de Nash (EN), et démontrons son unicité.

A.1 L'existence de l'équilibre de Nash

Dans la théorie des jeux, l'équilibre de Nash est un concept fondamental, communément utilisé pour prédire le comportement futur de tous les joueurs et déterminer l'état permanent. Cela signifie que chaque joueur du jeu n'a aucun intérêt à modifier sa stratégie et effectue la même action. Par conséquent, si un EN existe, un jeu non coopératif a une solution. Dans notre jeu $G=(K, St_k, Gn_k)_{k \in K}$ un ensemble de stratégies représente un EN si aucun joueur ne peut augmenter son gain en changeant de stratégie. Formellement, un EN est un N-tuple $st^* = [q_1^*, \dots, q_k^*, \dots, q_n^*]$, $st^* \in St$ représente un EN si aucun joueur ne peut augmenter son gain en changeant de stratégie. Formellement, un EN est un N-tuple $\{q_k^*\}$ satisfaisant :

$$Gn_k(q_k^*, q_{-k}^*) \geq Gn_k(q_k, q_{-k}^*), \forall q_k^*, q_k \in St_k, q_k \neq q_k^* \forall k \in K \quad (\text{A.1})$$

Nous utilisons le théorème de Nikaido-Isodra [82] pour démontrer qu'il existe au moins un EN :

theorem 1 (Nikaiido-Isoda). Le jeu $G=(K, St_k, Gn_k)_{k \in K}$ a au moins un EN si et seulement si, $\forall k \in K$, St_k est compact et convexe, et $Gn_k(q_k, q_{-k})$ est concave dans St_k et continu dans les stratégies de profil $st \in St$.

Démonstration. — L'ensemble de stratégie de tous les joueurs est $St = \prod_{k=1}^n St_k$, where $0 \leq St_k \leq q_k^{max}$, $\forall k \in K$. As $St_k = [0, q_k^{max}]$, l'ensemble des stratégies de chaque joueur est fermé et délimité. Ainsi, l'ensemble St_k est compact, $\forall k \in K$.

Soit $z_1, z_2 \in St_k$ deux points et $\epsilon = [0, 1]$. Par conséquent, nous avons $0 \leq \epsilon z_1 + (1 - \epsilon) z_2 \leq q_k^{max}$, ce qui signifie que le point $\epsilon z_1 + (1 - \epsilon) z_2 \in St_k$. Ainsi, on peut dire que St_k est convexe ; $\forall k \in K$.

— Pour prouver que la fonction de gain est concave, nous calculons la matrice hessienne de $Gn(s)$,

avec $st = \{q_k\} \forall k \in K$, comme suit :

$$H(st) = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix} \quad (\text{A.2})$$

Où $A_{kl} = \left(\frac{\partial^2 G_{n_k}}{\partial q_k \partial q_l} \right), \forall k, l \in K$.

$$A_{kl} = \frac{\partial}{\partial q_l} \left[\frac{\partial G_n}{\partial q_k}(q_l, q_k) \right] = \frac{\partial}{\partial q_l} \left[\frac{\alpha}{(q_k+1)} - \beta DCp_k - \psi CAp_k - \omega TAp_k - \varpi GECp_k - \vartheta REPP_k \right]$$

Donc, on obtient :

$$A_{kl} = \begin{cases} -\frac{\alpha}{(q_k+1)^2} < 0 & \text{if } k = l; \forall k, l \in K \\ 0 & \text{if } k \neq l; \forall k, l \in K \end{cases} \quad (\text{A.3})$$

En ce qui concerne le principal mineur de $H(st)$, on remarque que pour toutes les stratégies $st \in St$, $A(st)$ est négatif, donc, $G_{n_k}(q_k, q_{-k})$ est strictement concave en St_k pour tout $k \in K$. Compte tenu de ces conditions et sur la base du théorème de Nikaido-Isodra [82], il existe au moins un EN dans le jeu G . □

A.2 unicité de l'équilibre de Nash

Nous utilisons le théorème de Rosen [91] pour prouver l'unicité du EN. Le théorème de Rosen stipule que si la concavité stricte en diagonale est respectée, alors l'unicité est garantie. Nous notons un vecteur aléatoire de paramètres positifs $v = (v_1, v_2, \dots, v_n)$ et définissons la somme positive pondérée de la fonction de gain $G_{n_k}(q_k, q_{-k}), \forall k \in K$ comme suit :

$$\varsigma(q_k, q_{-k}; v) = \sum_{k=1}^n v_k G_{n_k}(q_k, q_{-k}), v_k \geq 0. \quad (\text{A.4})$$

Où le pseudo-gradient de $\varsigma(q_k, q_{-k}; v)$ est donné par :

$$f(q_k, q_{-k}; v) = \begin{bmatrix} v_1 \nabla G_{n_1}(q_1, q_{-1}) \\ v_2 \nabla G_{n_2}(q_2, q_{-2}) \\ \vdots \\ v_n \nabla G_{n_L}(q_n, q_{-n}) \end{bmatrix} \quad (\text{A.5})$$

Nous donnons :

$$\nabla G_{n_k}(q_k, q_{-k}) = \frac{\alpha}{(q_k+1)} - \beta \theta_k(q_k, DCp_k) - \psi \phi_k(q_k, CAp_k) - \omega \rho_k(q_k, TAp_k) - \varpi \varphi_k(q_k, GECp_k) - \vartheta \varrho_k(q_k, REPP_k)$$

De plus, nous définissons la matrice Jacobienne $JA(q_k, q_{-k}, v)$ de f comme suit :

$$JA(q_k, q_{-k}, v) = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{bmatrix} \quad (\text{A.6})$$

Où $C_{kl} = v_k A_{kl}; \forall k, l \in K$.

Par conséquent, on remarque que la matrice symétrique $[JA + JA^T]$ est négative pour tout les $q_k, q_{-k} \in St$. Et selon le théorème de Rosen [91], la fonction $\varsigma(q_k, q_{-k}; v)$ est diagonalement strictement concave. Par conséquent, le jeu G a un EN unique.

Liste des publications

Publications dans des journaux internationaux référencés :

1. Mohammed Anis BENBLIDIA, Bouziane BRIK, Moez ESSEGHIR, Leila MERGHEM-BOULAHIA, "A Renewable Energy-Aware Power Allocation for Cloud Data Centers : A Game Theory Approach", Elsevier : Computer communications, soumis.
2. Mohammed Anis BENBLIDIA, Bouziane BRIK, Moez ESSEGHIR, Leila MERGHEM-BOULAHIA, "Power Allocation And Energy Cost Minimization In Cloud Data Centers Microgrids : A Two-Stage Optimization Approach", IEEE Systems Journal, soumis.

Publications dans des congrès internationaux avec articles étendus :

1. Mohammed Anis BENBLIDIA, Bouziane BRIK, Moez ESSEGHIR, Leila MERGHEM-BOULAHIA, "Power Dispatching in Cloud Data Centers Using Smart Microgrids : A Game Theory Approach", IEEE Global Communications Conference (GLOBECOM), Waikoloa, HI, USA, 2019, pp. 1-6.
2. Mohammed Anis BENBLIDIA, Bouziane BRIK, Leila MERGHEM-BOULAHIA, Moez ESSEGHIR, "Ranking Fog nodes for Task Scheduling in Fog-Cloud Environments : A Fuzzy Logic Approach", IEEE 15th International Wireless Communications and Mobile Computing Conference (IWCMC), Tangier, Marocco, 2019, pp. 1451-1457.
3. Mohammed Anis BENBLIDIA, Bouziane BRIK, Moez ESSEGHIR, Leila MERGHEM-BOULAHIA, "A Game Based Power Allocation in Cloud Computing Data Centers", IEEE 14th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), Limassol, 2018, pp. 1-7.

Autres produits scientifiques : chapitre de livre

Mohammed Anis BENBLIDIA, Leila MERGHEM-BOULAHIA, Moez ESSEGHIR, Bouziane BRIK
Chapitre de livre–ISTE Editions : "Gestion intelligente des ressources dans un système Smart Grid-Cloud pour une meilleure efficacité énergétique", 2019.

Bibliographie

- [1] URL : <http://www.smartgrids-cre.fr>.
- [2] " 2018 <https://www.juniperresearch.com/press/press-releases/iot-connectionsto-grow-140-to-hit-50-BILLION> "Iot connections to grow 140computing accelerates roi. *Iot connections growth*. 2018.
- [3] S. ABOLFAZLI et al. "Cloud-Based Augmentation for Mobile Devices : Motivation, Taxonomies, and Open Challenges". In : *IEEE Communications Surveys Tutorials* 16.1 (2014), p. 337-368. DOI : 10.1109/SURV.2013.070813.00285.
- [4] B. AKSANLI, A. S. AKYUREK et T. ROSING. "Minimizing the effects of data centers on microgrid stability". In : *2015 Sixth International Green and Sustainable Computing Conference (IGSC)*. 2015, p. 1-9.
- [5] H. ALLCOTT. In : *Harvard University* (2009).
- [6] *Amazon, Google, Microsoft : Here's Who Has the Greenest Cloud*. URL : <https://www.wired.com/story/amazon-google-microsoft-green-clouds-and-hyperscale-data-centers/> (Accessed%20=%2007-12-2020).
- [7] P. ASMUS. "Data centers and advanced microgrids". In : *NAVIGANT Research Report* (2017).
- [8] *AWS and Sustainability*. URL : <https://aws.amazon.com/about-aws/sustainability/> (Accessed%20=%2007-12-2020).
- [9] A.K. AZAD et al. "Analysis of Wind Energy Prospect for Power Generation by Three Weibull Distribution Methods". In : *Energy Procedia* 75 (2015). Clean, Efficient and Affordable Energy for a Sustainable Future : The 7th International Conference on Applied Energy (ICAE2015), p. 722-727. ISSN : 1876-6102. DOI : <https://doi.org/10.1016/j.egypro.2015.07.499>. URL : <http://www.sciencedirect.com/science/article/pii/S1876610215012679>.
- [10] P. G. BALAJI et D. SRINIVASAN. "An Introduction to Multi-Agent Systems". In : *Innovations in Multi-Agent Systems and Applications - 1*. Sous la dir. de Dipti SRINIVASAN et Lakhmi C. JAIN. Berlin, Heidelberg : Springer Berlin Heidelberg, 2010, p. 1-27. ISBN : 978-3-642-14435-6. DOI : 10.1007/978-3-642-14435-6_1. URL : https://doi.org/10.1007/978-3-642-14435-6_1.

- [11] J. BALIGA et al. “Green Cloud Computing : Balancing Energy in Processing, Storage, and Transport”. In : *Proceedings of the IEEE* 99.1 (2011), p. 149-167. DOI : 10.1109/JPROC.2010.2060451.
- [12] Luiz Andr BARROSO, Jimmy CLIDARAS et Urs HLZLE. *The Datacenter as a Computer : An Introduction to the Design of Warehouse-Scale Machines*. 2nd. Morgan Claypool Publishers, 2013. ISBN : 1627050094.
- [13] A. BELOGLAZOV et R. BUYYA. “Energy Efficient Resource Management in Virtualized Cloud Data Centers”. In : *2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing*. 2010, p. 826-831. DOI : 10.1109/CCGRID.2010.46.
- [14] M. A. BENBLIDIA et al. “A Game Based Power Allocation in Cloud Computing Data Centers”. In : *2018 14th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*. 2018, p. 1-7. DOI : 10.1109/WiMOB.2018.8589177.
- [15] M. A. BENBLIDIA et al. “Power Dispatching in Cloud Data Centers Using Smart Microgrids : A Game Theory Approach”. In : *2019 IEEE Global Communications Conference (GLOBECOM)*. 2019, p. 1-6.
- [16] S. BERA, S. MISRA et J. RODRIGUES. “Cloud Computing Applications for Smart Grid : A Survey”. In : *IEEE Transactions on Parallel and Distributed Systems* 26 (2015), p. 1477-1494.
- [17] Claire BERGAENTZLÉ et Cédric CLASTRES. *Tarifications dynamiques et efficacité énergétique : l'apport des Smart Grids*. Post-Print halshs-00822731. HAL, fév. 2013. URL : <https://ideas.repec.org/p/hal/journal/halshs-00822731.html>.
- [18] Stuart BORLASE. *Smart grids : infrastructure, technology, and solutions*. CRC press, 2016.
- [19] Patrick BOSC, Ludovic LIETARD et Olivier PIVERT. “On the interpretation of set-oriented fuzzy quantified queries and their evaluation in a database management system”. In : *Methodologies for Intelligent Systems*. Sous la dir. de Jan KOMOROWSKI et Zbigniew W. RAŚ. Berlin, Heidelberg : Springer Berlin Heidelberg, 1993, p. 209-218. ISBN : 978-3-540-47750-1.
- [20] Rajkumar BUYYA, Chee Shin YEO et Srikumar VENUGOPAL. “Market-oriented cloud computing : Vision, hype, and reality for delivering it services as computing utilities”. In : *2008 10th IEEE international conference on high performance computing and communications*. Ieee. 2008, p. 5-13.
- [21] Aaron CARROLL et Gernot HEISER. “An Analysis of Power Consumption in a Smartphone”. In : *USENIXATC'10* (2010), p. 21-21. URL : <http://dl.acm.org/citation.cfm?id=1855840.1855861>.
- [22] D. ÇAVDAR et F. ALAGOZ. “A survey of research on greening data centers”. In : *2012 IEEE Global Communications Conference (GLOBECOM)*. 2012, p. 3237-3242. DOI : 10.1109/GLOCOM.2012.6503613.

- [23] Soraya CHACHOUA et Nouredine TAMANI. “Hybrid Ranking Method for E-Learning Platform Selection : A Flexible Approach”. In : *Computers Supported Education*. Sous la dir. de Gennaro COSTAGLIOLA et al. Cham : Springer International Publishing, 2017, p. 214-236. ISBN : 978-3-319-63184-4.
- [24] S. CHEN, Y. WANG et M. PEDRAM. “Resource allocation optimization in a data center with energy storage devices”. In : *IECON 2014 - 40th Annual Conference of the IEEE Industrial Electronics Society*. 2014, p. 2604-2610. DOI : 10.1109/IECON.2014.7048873.
- [25] D. CHENG, C. JIANG et X. ZHOU. “Heterogeneity-Aware Workload Placement and Migration in Distributed Sustainable Datacenters”. In : (2014), p. 307-316. DOI : 10.1109/IPDPS.2014.41.
- [26] M. CHIANG et T. ZHANG. “Fog and IoT : An Overview of Research Opportunities”. In : *IEEE Internet of Things Journal* 3.6 (2016), p. 854-864. DOI : 10.1109/JIOT.2016.2584538.
- [27] M. CHIANG et al. “Clarifying Fog Computing and Networking : 10 Questions and Answers”. In : *IEEE Communications Magazine* 55.4 (2017), p. 18-20. DOI : 10.1109/MCOM.2017.7901470.
- [28] Hybrid CLOUD. “The NIST Definition of Cloud Computing”. In : *National Institute of Science and Technology, Special Publication, 800 145* (2011).
- [29] <https://www.huawei.com/uk/about-uawei/publications/COMMUNICATE/85/CLOUD-DATA-TOPIC=CENTER-IN-5G-ERA>. *Cloud data centers in the 5g era*.
- [30] X. DAI, J. M. WANG et B. BENSOU. “Energy-Efficient Virtual Machines Scheduling in Multi-Tenant Data Centers”. In : *IEEE Transactions on Cloud Computing* 4.2 (2016), p. 210-221. DOI : 10.1109/TCC.2015.2481401.
- [31] X. DAI, J. M. WANG et B. BENSOU. “Energy-Efficient Virtual Machines Scheduling in Multi-Tenant Data Centers”. In : *IEEE Transactions on Cloud Computing* 4.2 (2016), p. 210-221. DOI : 10.1109/TCC.2015.2481401.
- [32] R. DENG et al. “A Survey on Demand Response in Smart Grids : Mathematical Models and Approaches”. In : *IEEE Transactions on Industrial Informatics* 11.3 (2015), p. 570-582. DOI : 10.1109/TII.2015.2414719.
- [33] W. DENG et al. “Harnessing renewable energy in cloud datacenters : opportunities and challenges”. In : *IEEE Network* 28.1 (2014), p. 48-55. DOI : 10.1109/MNET.2014.6724106.
- [34] Z. DING et al. “Emission-Aware Stochastic Resource Planning Scheme for Data Center Microgrid Considering Batch Workload Scheduling and Risk Management”. In : *IEEE Transactions on Industry Applications* 54.6 (2018), p. 5599-5608. DOI : 10.1109/TIA.2018.2851516.
- [35] Z. DING et al. “Integrated Stochastic Energy Management for Data Center Microgrid Considering Waste Heat Recovery”. In : *IEEE Transactions on Industry Applications* 55.3 (2019), p. 2198-2207.
- [36] T. H. DUONG-BA et al. “A Dynamic Virtual Machine Placement and Migration Scheme For Data Centers”. In : *IEEE Transactions on Services Computing* (2018), p. 1-1. DOI : 10.1109/TSC.2018.2817208.

- [37] E. ELMROTH et al. “Connecting Fog and Cloud Computing”. In : *IEEE Cloud Computing* 4.2 (2017), p. 22-25. DOI : 10.1109/MCC.2017.29.
- [38] Hiroshi ENDO et al. “Effect of climatic conditions on energy consumption in direct fresh-air container data centers”. In : *Sustainable Computing : Informatics and Systems* 6 (2015). Special Issue on Selected Papers from 2013 International Green Computing Conference (IGCC), p. 17-25. ISSN : 2210-5379. DOI : <https://doi.org/10.1016/j.suscom.2014.03.003>. URL : <http://www.sciencedirect.com/science/article/pii/S2210537914000067>.
- [39] *Energy Market Operational Data*. URL : [https://www.nyiso.com/energy-market-operational-data%20\(Accessed=%2018-09-2020\)](https://www.nyiso.com/energy-market-operational-data%20(Accessed=%2018-09-2020)).
- [40] M. EROL-KANTARCI et H. T. MOUFTAH. “Energy-Efficient Information and Communication Infrastructures in the Smart Grid : A Survey on Interactions and Open Issues”. In : *IEEE Communications Surveys Tutorials* 17.1 (Firstquarter 2015), p. 179-197. ISSN : 1553-877X. DOI : 10.1109/COMST.2014.2341600.
- [41] H. FARHANGI. “The path of the smart grid”. In : *IEEE Power and Energy Magazine* 8.1 (2010), p. 18-28. DOI : 10.1109/MPE.2009.934876.
- [42] N. MAIZI G.GUERASSIMOF. In : *Smart Grids : Au-dela du concept, comment rendre les réseaux plus intelligents, Presses de l'Ecole des mines*. 2013.
- [43] A. Z. GABR, A. A. HELAL et N. H. ABBASY. “Dynamic pricing ; different schemes, related research survey and evaluation”. In : *2018 9th International Renewable Energy Congress (IREC)*. 2018, p. 1-7. DOI : 10.1109/IREC.2018.8362562.
- [44] Anshul GANDHI et al. “Optimal Power Allocation in Server Farms”. In : *Proceedings of the Eleventh International Joint Conference on Measurement and Modeling of Computer Systems. SIGMETRICS '09*. Seattle, WA, USA : Association for Computing Machinery, 2009, p. 157-168. ISBN : 9781605585116. DOI : 10.1145/1555349.1555368. URL : <https://doi.org/10.1145/1555349.1555368>.
- [45] Peter Xiang GAO et al. “It’s Not Easy Being Green”. In : *SIGCOMM '12* (2012), p. 211-222. DOI : 10.1145/2342356.2342398. URL : <http://doi.acm.org/10.1145/2342356.2342398>.
- [46] Maurizio GIACOBBE et al. “Towards energy management in Cloud federation : A survey in the perspective of future sustainable and cost-saving strategies”. In : *Computer Networks* 91 (2015), p. 438-452. ISSN : 1389-1286. DOI : <https://doi.org/10.1016/j.comnet.2015.08.031>. URL : <http://www.sciencedirect.com/science/article/pii/S1389128615002911>.
- [47] C. GU et al. “Lowering Down the Cost for Green Cloud Data Centers by Using ESDs and Energy Trading”. In : (août 2016), p. 1508-1515. DOI : 10.1109/TrustCom.2016.0235.
- [48] Chonglin GU et al. “Greening cloud data centers in an economical way by energy trading with power grid”. In : *Future Generation Computer Systems* 78 (2018), p. 89-101. ISSN : 0167-739X. DOI : <https://doi.org/10.1016/j.future.2016.12.029>. URL : <http://www.sciencedirect.com/science/article/pii/S0167739X16308330>.

- [49] G. GUERASSIMOF. “Microgrids : pourquoi, pour qui ?” In : *Presses de l’Ecole des mines* (2017).
- [50] V. C. GUNGOR et al. “Smart Grid Technologies : Communication Technologies and Standards”. In : *IEEE Transactions on Industrial Informatics* 7.4 (2011), p. 529-539. DOI : 10.1109/TII.2011.2166794.
- [51] K. HAGSHENAS et al. “Infrastructure Aware Heterogeneous-Workloads Scheduling for Data Center Energy Cost Minimization”. In : *IEEE Transactions on Cloud Computing* (2020), p. 1-1.
- [52] M. E. EL-HAWARY. “The Smart Grid—State-of-the-art and future trends”. In : *2016 Eighteenth International Middle East Power Systems Conference (MEPCON)*. 2016, p. xvii-xvii. DOI : 10.1109/MEPCON.2016.7836856.
- [53] Brian HAYES. “Cloud Computing”. In : *Commun. ACM* 51.7 (juil. 2008), p. 9-11. ISSN : 0001-0782. DOI : 10.1145/1364782.1364786. URL : <https://doi.org/10.1145/1364782.1364786>.
- [54] [HTTP://WWW.NIST.GOV/PUBLICAFFAIRS/RELEASES/UPLOAD/SMARTGRID_INTEROPERABILITY_FINAL.PDF](http://www.nist.gov/publicaffairs/releases/upload/smartgrid_interoperability_final.pdf). *NIST Framework and Roadmap for Smart Grid Interoperability Standards, Release 1.0. NIST Special Publication 1108*,
- [55] [HTTPS://DOCUMENTS.SWISSCOM.COM/PRODUCT/1000174-INTERNET/DOCUMENTS/CLOUD/CLOUD-FUER-DIGITALE-TRANSFORMATION_fr.pdf](https://documents.swisscom.com/product/1000174-INTERNET/DOCUMENTS/CLOUD/CLOUD-FUER-DIGITALE-TRANSFORMATION_fr.pdf). *Le cloud comme accélérateur de votre transformation numérique*.
- [56] [HTTPS://WWW.CISCO.COM](https://www.cisco.com). *Fog Computing and the Internet of Things : Extend the Cloud to Where the Things Are*. 2015.
- [57] [HTTPS://WWW.CISCO.COM/C/EN/US/SOLUTIONS/COLLATERAL/EXECUTIVE-PERSPECTIVES/ANNUAL-INTERNET-REPORT/WHITE-PAPER-C11-741490.PDF](https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.pdf). *Cisco annual Internet report 2018-2023*. 2018.
- [58] [HTTPS://WWW.CONNAISSANCEDESENERGIES.ORG/VERS-UNE-HAUSSE-MAJEURE-DE-LA-CONSOMMATION-MONDIALE-DENERGIE-160519](https://www.connaissancedesenergies.org/vers-une-hausse-majeure-de-la-consommation-mondiale-denergie-160519). *Vers une hausse majeure de la consommation mondiale d’énergie ?*
- [59] [HTTPS://WWW.EDF.FR/ENTREPRISES/LE-MAG/LE-MAG-ENTREPRISES/DECRYPTAGE-DU-MARCHE-DE-L-ENERGIE/AUGMENTATION-DE-57-DE-LA-CONSOMMATION-ELECTRIQUE-MONDIALE-D-ICI-2050](https://www.edf.fr/entreprises/le-mag/le-mag-entreprises/decryptage-du-marche-de-l-energie/augmentation-de-57-de-la-consommation-electrique-mondiale-d-ici-2050). *Augmentation de 57% de la consommation électrique mondiale d’ici 2050* », EDF France, sept. 14, 2018.
- [60] [HTTPS://WWW.ENGIE.FR/ACTUALITES/MICROGRIDS-SMART-GRID-EXPERIENCE/](https://www.engie.fr/actualites/microgrids-smart-grid-experience/). *Les microgrids : une tendance d’avenir*.
- [61] [HTTPS://WWW.GARTNER.COM/EN/NEWSROOM/PRESS-RELEASES/2019-04-02-GARTNER-FORECASTS-WORLDWIDE-PUBLIC-CLOUD-REVENUE-TO-G](https://www.gartner.com/en/newsroom/press-releases/2019-04-02-gartner-forecasts-worldwide-public-cloud-revenue-to-grow-17.5-percent-in-2019). *Gartner Forecasts Worldwide Public Cloud Revenue to Grow 17.5 Percent in 2019*.
- [62] J.KLAIMI. In : *Gestion multi-agents des smart grids intégrant un système de stockage : cas résidentiel*. 2017.

- [63] Xibo JIN et al. “Green Data Centers : A Survey, Perspectives, and Future Directions”. In : *ArXiv* abs/1608.00687 (2016).
- [64] Yogendra JOSHI et Pramod KUMAR. *Energy Efficient Thermal Management of Data Centers*. Jan. 2012. ISBN : 978-1-4419-7123-4. DOI : 10.1007/978-1-4419-7124-1.
- [65] S. JOŠILO et G. DÁN. “Decentralized Algorithm for Randomized Task Allocation in Fog Computing Systems”. In : *IEEE/ACM Transactions on Networking* 27.1 (2019), p. 85-97. DOI : 10.1109/TNET.2018.2880874.
- [66] C. JU et al. “A Two-Layer Energy Management System for Microgrids With Hybrid Energy Storage Considering Degradation Costs”. In : *IEEE Transactions on Smart Grid* 9.6 (2018), p. 6047-6057.
- [67] Yasin KABALCI et Ersan KABALCI. “Modeling and analysis of a smart grid monitoring system for renewable energy sources”. In : *Solar Energy* 153 (2017), p. 262-275. ISSN : 0038-092X. DOI : <https://doi.org/10.1016/j.solener.2017.05.063>. URL : <http://www.sciencedirect.com/science/article/pii/S0038092X17304589>.
- [68] A. A. KHAN, M. H. REHMANI et M. REISSLEIN. “Cognitive Radio for Smart Grids : Survey of Architectures, Spectrum Sensing Mechanisms, and Networking Protocols”. In : *IEEE Communications Surveys Tutorials* 18.1 (2016), p. 860-898. DOI : 10.1109/COMST.2015.2481722.
- [69] A. KIANI et N. ANSARI. “Profit Maximization for Geographically Dispersed Green Data Centers”. In : *IEEE Transactions on Smart Grid* 9.2 (2018), p. 703-711. DOI : 10.1109/TSG.2016.2562565.
- [70] D. KLIÁZOVICH et al. “GreenCloud : A Packet-Level Simulator of Energy-Aware Cloud Computing Data Centers”. In : *2010 IEEE Global Telecommunications Conference GLOBECOM 2010*. 2010, p. 1-5. DOI : 10.1109/GLOCOM.2010.5683561.
- [71] Sonja KLINGERT. “Mapping Data Centre Business Types with Power Management Strategies to Identify Demand Response Candidates”. In : New York, NY, USA : Association for Computing Machinery, 2018. ISBN : 9781450357678. DOI : 10.1145/3208903.3213521. URL : <https://doi.org/10.1145/3208903.3213521>.
- [72] P. M. Avgerinou L.CASTELLAZZI. “Trends in data centre energy consumption under the european code of conduct for data centre energy efficiency”. In : *On energies MDPI* (2017).
- [73] M. LEVY, D. RAVIV et J. O. HALLSTROM. “Data Center Modeling Using a Cyber-Physical Systems Lens”. In : (jan. 2019), p. 0146-0155. DOI : 10.1109/CCWC.2019.8666523.
- [74] L. LIU et al. “Exploring Customizable Heterogeneous Power Distribution and Management for Datacenter”. In : *IEEE Transactions on Parallel and Distributed Systems* 29.12 (2018), p. 2798-2813. DOI : 10.1109/TPDS.2018.2841405.
- [75] X. LU et al. “Bulk Savings for Bulk Transfers : Minimizing the Energy-Cost for Geo-Distributed Data Centers”. In : *IEEE Transactions on Cloud Computing* 8.1 (2020), p. 73-85.
- [76] R. MA et al. “Smart Grid Communication : Its Challenges and Opportunities”. In : *IEEE Transactions on Smart Grid* 4.1 (2013), p. 36-46. DOI : 10.1109/TSG.2012.2225851.

- [77] Dragan S. MARKOVIC et al. “Smart power grid and cloud computing”. In : *Renewable and Sustainable Energy Reviews* 24 (2013), p. 566-577. ISSN : 1364-0321. DOI : <https://doi.org/10.1016/j.rser.2013.03.068>. URL : <http://www.sciencedirect.com/science/article/pii/S136403211300227X>.
- [78] Sahbi MARROUCHI et Souad Ben SABER. “A Comparative Study of Fuzzy Logic, Genetic Algorithm, and Gradient-Genetic Algorithm Optimization Methods for Solving the Unit Commitment Problem”. In : *Mathematical Problems in Engineering* 2014 (2014), p. 1-14.
- [79] *Microsoft will be carbon negative by 2030*. URL : <https://blogs.microsoft.com/blog/2020/01/16/microsoft-will-be-carbon-negative-by-2030/> (Accessed%20=%2007-12-2020).
- [80] M. M. MOGHADDAM et al. “On Data Center Demand Response : A Cloud Federation Approach”. In : *IEEE Access* 7 (2019), p. 101829-101843.
- [81] A. MOHSENIAN-RAD et al. “Autonomous Demand-Side Management Based on Game-Theoretic Energy Consumption Scheduling for the Future Smart Grid”. In : *IEEE Transactions on Smart Grid* 1.3 (2010), p. 320-331. DOI : 10.1109/TSG.2010.2089069.
- [82] Hukukane NIKAIDÔ et Kazuo ISODA. “Note on non-cooperative convex games”. In : *Pacific J. Math.* 5 (1955), p. 807-815.
- [83] NRDC. “Data Center Efficiency Assessment”. In : (2014).
- [84] RPS BEDI P. CHOPRA. “Application of Fuzzy logic in Cloud Computing : A Review”. In : *International Journal of Scientific Research Engineering Technology (IJSRET)* (2017).
- [85] Xuan-Qui PHAM et Eui-Nam HUH. “Towards task scheduling in a cloud-fog computing system”. In : *2016 18th Asia-Pacific Network Operations and Management Symposium (APNOMS)*. Oct. 2016, p. 1-4. DOI : 10.1109/APNOMS.2016.7737240.
- [86] A. S. PRASAD et S. RAO. “A Mechanism Design Approach to Resource Procurement in Cloud Computing”. In : *IEEE Transactions on Computers* 63.1 (2014), p. 17-30. DOI : 10.1109/TC.2013.106.
- [87] A. u. R. KHAN et al. “A Survey of Mobile Cloud Computing Application Models”. In : *IEEE Communications Surveys Tutorials* 16.1 (2014), p. 393-413. DOI : 10.1109/SURV.2013.062613.00160.
- [88] A. RAHMAN, X. LIU et F. KONG. “A Survey on Geographic Load Balancing Based Data Center Power Management in the Smart Grid Environment”. In : *IEEE Communications Surveys Tutorials* 16.1 (2014), p. 214-233. DOI : 10.1109/SURV.2013.070813.00183.
- [89] L. RAO et al. “Minimizing Electricity Cost : Optimization of Distributed Internet Data Centers in a Multi-Electricity-Market Environment”. In : *2010 Proceedings IEEE INFOCOM*. 2010, p. 1-9. DOI : 10.1109/INFOCOM.2010.5461933.
- [90] “Report to congress on server and data center energy efficiency”. In : (2007).
- [91] J. B. ROSEN. “Existence and Uniqueness of Equilibrium Points for Concave N-Person Games”. In : *Econometrica* 33 (1965), p. 520-534.

- [92] P. SAMADI et al. “Advanced Demand Side Management for the Future Smart Grid Using Mechanism Design”. In : *IEEE Transactions on Smart Grid* 3.3 (2012), p. 1170-1180. DOI : 10.1109/TSG.2012.2203341.
- [93] N. K. SHARMA et G. R. M. REDDY. “Multi-Objective Energy Efficient Virtual Machines Allocation at the Cloud Data Center”. In : *IEEE Transactions on Services Computing* 12.1 (2019), p. 158-171. DOI : 10.1109/TSC.2016.2596289.
- [94] H. SHI et al. “Fairness in Wireless Networks :Issues, Measures and Challenges”. In : *IEEE Communications Surveys Tutorials* 16.1 (2014), p. 5-24. DOI : 10.1109/SURV.2013.050113.00015.
- [95] Ankit SOMANI et Yogendra JOSHI. “Data Center Cooling Optimization : Ambient Intelligence Based Load Management (AILM)”. In : t. 1. Jan. 2009. DOI : 10.1115/HT2009-88228.
- [96] National Institute of STANDARDS et TECHNOLOGY. *Fog Computing Conceptual Model : NiST SP 500-325 500-291*. North Charleston, SC, USA : CreateSpace Independent Publishing Platform, 2018. ISBN : 1986596958.
- [97] N. TAMANI et al. “Vehicular Cloud Service Provider Selection : A Flexible Approach”. In : *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*. Déc. 2017, p. 1-6. DOI : 10.1109/GLOCOM.2017.8253924.
- [98] V. TERZIJA et al. “Wide-Area Monitoring, Protection, and Control of Future Electric Power Networks”. In : *Proceedings of the IEEE* 99.1 (2011), p. 80-93. DOI : 10.1109/JPROC.2010.2060450.
- [99] *The Future of Data Centers – Renewable Energy*. URL : <https://www.colocationamerica.com/blog/renewable-energy-data-centers> (Accessed%20=%2007-12-2020).
- [100] *The nexus between data centres, efficiency and renewables : a role model for the energy transition*. URL : <https://energypost.eu/the-nexus-between-data-centres-efficiency-and-renewables-a-role-model-for-the-energy-transition/> (Accessed%20=%2007-12-2020).
- [101] *The Organisation for Economic Co-operation and Development (OECD), Nomura Research Institute (NRI) Presentation*. URL : <http://www.oecd.org/sti/ind/45009540.pdf> %20(Accessed%20=%2022-09-2020).
- [102] N. H. TRAN et al. “How Geo-Distributed Data Centers Do Demand Response : A Game-Theoretic Approach”. In : *IEEE Transactions on Smart Grid* 7.2 (2016), p. 937-947. DOI : 10.1109/TSG.2015.2421286.
- [103] *Tsunami of data could consume one fifth of global electricity by 2025*. URL : <https://www.climatechangenews.com/2017/12/11/%20tsunami-data-consume-one-fifth-global-electricity-2025>.
- [104] E. HOSSAIN Z.HAN V.POOR. “Communication architectures and models for Smart Grid”. In : *Cambridge University Press 978-1-107-01413-8 - Smart Grid Communications and Networking* (2012).

- [105] J. S. VARDAKAS, N. ZORBA et C. V. VERIKOUKIS. “A Survey on Demand Response Programs in Smart Grids : Pricing Methods and Optimization Algorithms”. In : *IEEE Communications Surveys Tutorials* 17.1 (2015), p. 152-178. DOI : 10.1109/COMST.2014.2341586.
- [106] S. WALIA et K. S. SANDHU. “Capacity factor of wind turbine system based on different power curves and Weibull distribution parameters”. In : *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*. Mar. 2019, p. 1135-1138. DOI : 10.1109/ICCMC.2019.8819668.
- [107] J. WAN et al. “Fog Computing for Energy-aware Load Balancing and Scheduling in Smart Factory”. In : *IEEE Transactions on Industrial Informatics* (2018), p. 1-1. ISSN : 1551-3203. DOI : 10.1109/TII.2018.2818932.
- [108] Di WANG, Chuangang REN et Anand SIVASUBRAMANIAM. “Virtualizing Power Distribution in Datacenters”. In : *Proceedings of the 40th Annual International Symposium on Computer Architecture*. ISCA '13. New York, NY, USA : Association for Computing Machinery, 2013, p. 595-606. ISBN : 9781450320795. DOI : 10.1145/2485922.2485973. URL : <https://doi.org/10.1145/2485922.2485973>.
- [109] H. WANG et Z. YE. “Renewable energy-aware demand response for distributed data centers in smart grid”. In : (nov. 2016), p. 1-8. DOI : 10.1109/IGESC.2016.7790076.
- [110] H. WANG et al. “Proactive Demand Response for Data Centers : A Win-Win Solution”. In : *IEEE Transactions on Smart Grid* 7.3 (2016), p. 1584-1596. DOI : 10.1109/TSG.2015.2501808.
- [111] L. WANG et G. S. G. S. KUO. “Mathematical Modeling for Network Selection in Heterogeneous Wireless Networks x2014 ; A Tutorial”. In : *IEEE Communications Surveys Tutorials* 15.1 (First 2013), p. 271-292. ISSN : 1553-877X.
- [112] Y. WANG, X. LIN et M. PEDRAM. “A Stackelberg Game-Based Optimization Framework of the Smart Grid With Distributed PV Power Generations and Data Centers”. In : *IEEE Transactions on Energy Conversion* 29.4 (2014), p. 978-987. DOI : 10.1109/TEC.2014.2363048.
- [113] Y. WANG, X. LIN et M. PEDRAM. “Coordination of the smart grid and distributed data centers : A nested game-based optimization framework”. In : *ISGT 2014*. 2014, p. 1-5. DOI : 10.1109/ISGT.2014.6816386.
- [114] Y. WANG et al. “A Game-Theoretic Approach to Energy Trading in the Smart Grid”. In : *IEEE Transactions on Smart Grid* 5.3 (2014), p. 1439-1450. DOI : 10.1109/TSG.2013.2284664.
- [115] Aaron WEISS. “Computing in the Clouds”. In : *NetWorker* 11.4 (déc. 2007), p. 16-25. ISSN : 1091-3556. DOI : 10.1145/1327512.1327513. URL : <https://doi.org/10.1145/1327512.1327513>.
- [116] *Wind vs. Solar — Which Power Source Is Better?* URL : <https://greenfuture.io/solar/wind-vs-solar-energy/> (Accessed%20=%2001-10-2020).
- [117] D. XU, X. LIU et B. FAN. “Efficient Server Provisioning and Offloading Policies for Internet Data Centers with Dynamic Load-Demand”. In : *IEEE Transactions on Computers* 64.3 (2015), p. 682-697. DOI : 10.1109/TC.2013.2295797.

- [118] Ronald R. YAGER. “General multiple-objective decision functions and linguistically quantified statements”. In : *International Journal of Man-Machine Studies* 21.5 (1984), p. 389-400. ISSN : 0020-7373. DOI : [https://doi.org/10.1016/S0020-7373\(84\)80066-8](https://doi.org/10.1016/S0020-7373(84)80066-8). URL : <http://www.sciencedirect.com/science/article/pii/S0020737384800668>.
- [119] Y. YAN et al. “A Survey on Smart Grid Communication Infrastructures : Motivations, Requirements and Challenges”. In : *IEEE Communications Surveys Tutorials* 15.1 (2013), p. 5-20. DOI : 10.1109/SURV.2012.021312.00034.
- [120] Y. YAN et al. “A Survey on Smart Grid Communication Infrastructures : Motivations, Requirements and Challenges”. In : *IEEE Communications Surveys Tutorials* 15.1 (2013), p. 5-20. DOI : 10.1109/SURV.2012.021312.00034.
- [121] J. YANG et al. “Online Learning-Based Server Provisioning for Electricity Cost Reduction in Data Center”. In : *IEEE Transactions on Control Systems Technology* 25.3 (2017), p. 1044-1051. DOI : 10.1109/TCST.2016.2575801.
- [122] Y. YANG et al. “MEETS : Maximal Energy Efficient Task Scheduling in Homogeneous Fog Networks”. In : *IEEE Internet of Things Journal* (2018), p. 1-1. ISSN : 2327-4662. DOI : 10.1109/JIOT.2018.2846644.
- [123] MURAT YILDIZOGLU. “Introduction à la théorie des jeux”. In : *Dunod* (2011).
- [124] L. YIN, J. LUO et H. LUO. “Tasks Scheduling and Resource Allocation in Fog Computing Based on Containers for Smart Manufacture”. In : *IEEE Transactions on Industrial Informatics* (2018), p. 1-1. ISSN : 1551-3203. DOI : 10.1109/TII.2018.2851241.
- [125] L. YIN, J. LUO et H. LUO. “Tasks Scheduling and Resource Allocation in Fog Computing Based on Containers for Smart Manufacturing”. In : *IEEE Transactions on Industrial Informatics* 14.10 (2018), p. 4712-4721. DOI : 10.1109/TII.2018.2851241.
- [126] L. YU, T. JIANG et Y. CAO. “Energy Cost Minimization for Distributed Internet Data Centers in Smart Microgrids Considering Power Outages”. In : *IEEE Transactions on Parallel and Distributed Systems* 26.1 (2015), p. 120-130. DOI : 10.1109/TPDS.2014.2308223.
- [127] L. YU, T. JIANG et Y. ZOU. “Real-Time Energy Management for Cloud Data Centers in Smart Microgrids”. In : *IEEE Access* 4 (2016), p. 941-950. DOI : 10.1109/ACCESS.2016.2539369.
- [128] L. YU et al. “Carbon-Aware Energy Cost Minimization for Distributed Internet Data Centers in Smart Microgrids”. In : *IEEE Internet of Things Journal* 1.3 (2014), p. 255-264. DOI : 10.1109/JIOT.2014.2322606.
- [129] H. YUAN et al. “Revenue and Energy Cost-Optimized Biobjective Task Scheduling for Green Cloud Data Centers”. In : *IEEE Transactions on Automation Science and Engineering* (2020), p. 1-14.
- [130] L. ZADEH. *Fuzzy sets,* *Information and Control*. 1965.
- [131] Hammad ZAFAR et al. “An Effective Fairness Scheme for Named Data Networking”. In : *Electronics* 9.5 (2020), p. 749.

- [132] G. ZHANG et al. “Fair Task Offloading among Fog Nodes in Fog Computing Networks”. In : *2018 IEEE International Conference on Communications (ICC)*. Mai 2018, p. 1-6. DOI : 10.1109/ICC.2018.8422316.

Mohammed Anis BENBLIDIA

Doctorat : Systèmes SocioTechniques

Année 2021

Pour une meilleure efficacité énergétique dans un système *Smart Grid* – *Cloud*

Dans cette thèse, nous étudions l'efficacité énergétique des infrastructures informatiques dans un système *smart grid - cloud*. Nous nous intéressons plus particulièrement aux réseaux de communication et aux *data centers* du cloud. Nous nous focalisons sur ces derniers à cause de leur grande consommation d'énergie et du rôle vital qu'ils jouent dans un monde connecté en pleine expansion, les positionnant, ainsi, comme des éléments importants dans un système *smart grid - cloud*. De ce fait, les travaux de cette thèse s'inscrivent dans le cadre d'un seul *framework* intégrant le *smart grid*, le *microgrid*, le *cloud*, les *data centers* et les utilisateurs. Nous avons, en effet, étudié l'interaction entre les *data centers* du *cloud* et le fournisseur d'énergie du *smart grid* et nous avons proposé des solutions d'allocation d'énergie et de minimisation du coût d'énergie en utilisant deux architectures : (1) une architecture *smart grid-cloud* et (2) une architecture *microgrid-cloud*. Par ailleurs, nous avons porté une attention particulière à l'exécution des requêtes des utilisateurs tout en leur garantissant un niveau de qualité de service satisfaisant dans une architecture *fog - cloud*. En comparaison avec les travaux de l'état de l'art, les résultats de nos contributions ont montré qu'ils répondent aux enjeux identifiés, notamment en réduisant les émissions de gaz à effet de serre et le coût d'énergie des *data centers*.

Mots clés : réseaux électriques intelligents – centres de traitement informatique – informatique dans les nuages – internet des objets – théorie des jeux.

For Better Energy Efficiency in a Smart Grid – Cloud System

This thesis considers the energy efficiency of information and communication infrastructures in a smart grid - cloud system. It especially deals with communication networks and cloud data centers due to their high energy consumption, which confers them an important role in the network. The contributions of this thesis are implemented on the same framework integrating the smart grid, microgrid, cloud, data centers and users. Indeed, we have studied the interaction between the cloud data centers and the smart grid provider and we have proposed energy efficient power allocation solutions and an energy cost minimization scheme using two architectures: a smart grid-cloud architecture and a microgrid-cloud architecture. In addition, we paid close attention to execute user requests while ensuring a good quality of service in a fog-cloud architecture. In comparison with state-of-the-art works, the results of our contributions have shown that they respond to the identified challenges, particularly in terms of reducing carbon emissions and energy costs of cloud data centers.

Keywords: smart power grids – data centers – cloud computing – internet of things – game theory.

Thèse réalisée en partenariat entre :



Ecole Doctorale "Sciences pour l'Ingénieur"