HAL
open science

Algorithmes et architectures adaptatifs pour les systèmes de communication sans fil efficaces en énergie

Matthieu Gautier

To cite this version:

Matthieu Gautier. Algorithmes et architectures adaptatifs pour les systèmes de communication sans fil efficaces en énergie. Architectures Matérielles [cs.AR]. Université de Rennes 1, 2019. tel-02136001

HAL Id: tel-02136001

https://theses.hal.science/tel-02136001

Submitted on 21 May 2019

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.

ANNÉE 2019

HABILITATION A DIRIGER LES RECHERCHES

présentée devant

L'Université de Rennes 1

Spécialité : Traitement du signal et Télécommunications

par

Matthieu Gautier

Algorithmes et architectures adaptatifs
pour les systèmes de communication sans fil efficaces en énergie

---

soutenue le 10 mai 2019 devant le jury composé de :

Tanguy Risset, Président de jury, Professeur, INSA Lyon

Cécile Belleudy, Rapporteur, Maître de Conférences (HDR), Université de Nice

Laurent Clavier, Rapporteur, Professeur, IMT Lille

Sébastien Roy, Rapporteur, Professeur, Université de Sherbrooke

Christophe Jégo, Examinateur, Professeur, ENSEIRB-MATMECA

Olivier Sentieys, Examinateur, Professeur, Université de Rennes 1

Olivier Berder, Examinateur, Professeur, Université de Rennes 1

Avant propos

Depuis novembre 2011, j'exerce la fonction de maître de conférences au sein de l'Université de Rennes 1. Mes activités d'enseignement sont réalisées au sein du département Réseaux et Télécommunications de l'IUT de Lannion et mes activités de recherche au sein de l'équipe GRANIT du laboratoire IRISA. Je pense avoir acquis lors de ces années une expérience et une autonomie importantes dans mes travaux de recherche qui me poussent aujourd'hui à présenter ma candidature à l'Habilitation à Diriger les Recherches de l'Université de Rennes 1.

Mes implications dans l'encadrement de nombreux stagiaires et doctorants ainsi que dans le montage, la participation et l'animation de projets régionaux, nationaux ou européens m'ont permis de me familiariser avec les différentes facettes de la recherche. Après une période d'adaptation compréhensible lors de l'installation d'un enseignant-chercheur, les recherches que nous menons sont aujourd'hui reconnues et validées par des publications dans des conférences et revues internationales de bon niveau. Au sein de l'équipe GRANIT, mon activité de recherche concerne l'efficacité énergétique des systèmes de communications. J'anime notamment des recherches sur les wake-up radios et sur la radio logicielle dont les premiers résultats sont très prometteurs.

Ce mémoire d'HDR concerne uniquement les travaux effectués depuis mon arrivée à l'IRISA en 2011. Ce choix est dicté par un souci de cohésion scientifique et par le contexte de ces travaux (encadrements de doctorants, projets collaboratifs,...) justifiant directement ma candidature à l'HDR. La totalité de mes activités de recherche sera cependant abordée dans une partie curriculum vitæ et dans la liste de publications.

Ce mémoire est composé de trois parties. La première présente les différents aspects de mes activités de recherche, les encadrements de stagiaires et de doctorants, les projets et collaborations ainsi que la valorisation de tous ces travaux. Une brève présentation de mes activités d'enseignement est également proposée. Dans les parties II et III, la partie scientifique du mémoire est développée autour de deux axes de recherche :

Axe 1 - Les réseaux de capteurs ultra-faible consommation : gestion d'énergie et protocoles de communications

Axe 2 - Les radios logicielles efficaces en énergie : prototypage rapide et exploration de l'espace de conception

Finalement, le bilan scientifique de ces travaux et les perspectives de recherche sont présentés dans une dernière partie.

En vous souhaitant une bonne lecture, je vous prie d'agréer, mesdames et messieurs, l'expression de mes sentiments distingués.

Matthieu Gautier, le 21 mai 2019.

Table des matières

Liste des acronymes **9**

Table des figures **11**

Liste des tableaux **13**

I Introduction et synthèse des travaux **15**

1 Contexte général et contributions **17**
1.1 Vers une efficacité énergétique des objets communicants 17
1.2 Axes de recherche et structure du document 19
1.2.1 Axe 1 : Les réseaux de capteurs ultra-faible consommation : gestion d'énergie et protocoles de communications 19
1.2.2 Axe 2 : Les radios logicielles efficaces en énergie : prototypage rapide et exploration de l'espace de conception 20
1.2.3 Structure du document et contributions 21

2 Synthèse des activités de recherche et d'enseignements **25**
2.1 Synthèse de la carrière 25
2.2 Activités d'enseignements 25
2.3 Activité de recherche scientifique 27
2.3.1 Présentation des thématiques de recherche 27
2.3.2 Publications et production scientifique 27
2.4 Encadrement doctoral et scientifique 28
2.4.1 Direction de thèses 28
2.4.2 Direction de stagiaires et de post-doctorants 29
2.5 Diffusion et rayonnement 30
2.6 Responsabilités scientifiques 31

II Les réseaux de capteurs ultra-faible consommation : gestion d'énergie et protocoles de communications **33**

3 Les réseaux de capteurs à récupération d'énergie **35**
3.1 Vers une autonomie énergétique des réseaux de capteurs 35
3.1.1 Récupération et gestion d'énergie 35
3.1.2 État de l'art des gestionnaires d'énergie 38
3.2 Fuzzyman : Utilisation de la logique floue 39
3.2.1 Conception de Fuzzyman 39
3.2.2 Évaluation de Fuzzyman 42
3.2.3 Conclusion 43
3.3 RLman : Utilisation de l'apprentissage par renforcement 44
3.3.1 Formulation du problème de gestion de l'énergie récupérée 44
3.3.2 Algorithme de RLMan 45
3.3.3 Évaluation de RLMan 46

3.3.4 Conclusion . . . 48
3.4 Autres contributions . . . 48
3.4.1 Gestion d'énergie distribuée pour les réseaux multi-sauts . . . 48
3.4.2 Architecture de récupération d'énergie multi-source . . . 49
3.4.3 Application de la gestion d'énergie à des nœuds LoRa . . . 49
3.5 Conclusion . . . 50

4 Les wake-up radios 51
4.1 Principe des wake-up radios . . . 52
4.1.1 Architecture des wake-up radios . . . 52
4.1.2 Caractéristiques des wake-up radios . . . 53
4.2 Evaluation analytique des wake-up radios . . . 54
4.2.1 Modélisation des protocoles . . . 54
4.2.2 Performances analytiques et comparaisons expérimentales . . . 56
4.2.3 Conclusion . . . 56
4.3 OPWUM : protocole MAC opportuniste . . . 57
4.3.1 Protocole MAC opportuniste utilisant la wake-up-radio . . . 57
4.3.2 Évaluation de OPWUM . . . 58
4.3.3 Application à la récupération d'énergie . . . 59
4.3.4 Conclusion . . . 60
4.4 SNW-MAC : collecte des données dans un réseau en étoile . . . 60
4.4.1 Le protocole Star Network WuRx - MAC . . . 60
4.4.2 Protocole expérimental : combinaison à la récupération d'énergie . . . 61
4.4.3 Résultats expérimentaux . . . 64
4.4.4 Conclusion . . . 65
4.5 Autres contributions . . . 65
4.5.1 Communications hétérogènes courtes et longues portées associant LoRa et WuRx . . . 65
4.6 Conclusion . . . 66

III Les radios logicielles efficaces en énergie : prototypage rapide et exploration de l'espace de conception 69

5 Description haut-niveau des formes d'ondes et prototypage rapide 71
5.1 La radio logicielle et le prototypage rapide . . . 72
5.1.1 La radio logicielle . . . 72
5.1.2 Méthodes de conception . . . 73
5.1.3 La synthèse de haut-niveau . . . 74
5.2 Description des formes d'ondes . . . 75
5.2.1 Langage de description spécifique (DSL) . . . 75
5.2.2 Flot de conception . . . 77
5.3 Prototypage rapide : expérimentation et validation . . . 78
5.3.1 Description DSL de la norme IEEE 802.15.4 . . . 79
5.3.2 Implémentation de la norme IEEE 802.15.4 . . . 80
5.3.3 Exploration de l'espace de conception . . . 81
5.4 Autres contributions . . . 82
5.4.1 Reconfiguration dynamique des formes d'ondes . . . 82
5.4.2 Adaptation dynamique d'un récepteur OFDM en virgule fixe . . . 84
5.5 Conclusion . . . 85

6 Méthodes de conception pour les architectures multi-cœurs hétérogènes 87
6.1 Vue globale du flot de conception pour les architectures HMpSoC . . . 88
6.1.1 Structure des architectures et applications ciblées . . . 88
6.1.2 Flot de conception orienté énergie . . . 90
6.2 Exploration architecturale d'accélérateur . . . 91
6.2.1 Applications, architectures et objectifs du DSE . . . 92

6.2.2 Estimation de l'énergie et du temps d'exécution 94
6.2.3 Exploration de l'espace de conception 96
6.3 Résultats expérimentaux d'exploration 97
6.3.1 Environnement d'expérimentation 98
6.3.2 Extraction des paramètres 99
6.3.3 Performances de l'exploration MILP 100
6.4 Autres contributions 102
6.4.1 Plateforme d'émulation de HMpSoC 102
6.5 Conclusion 102

7 Conclusions et perspectives **105**
7.1 Bilan scientifique 105
7.1.1 Axes de recherche 105
7.1.2 Bilan personnel 106
7.2 Perspectives de recherche 107
7.2.1 Smart wake-up radio 107
7.2.2 Near sensor computing 108
7.2.3 Radio logicielle pour l'interception de signaux compromettants 108
7.2.4 Partitionnement efficace en énergie pour les architectures Cloud-RAN . . . 109

Publications personnelles **111**

Références **119**

Liste des acronymes

ACK Acknowledgment.
ADMM Alternating Direction Method of Multipliers.
AMC Absorbing Markov Chain.
ASIC Application-Specific Integrated Circuit.
ATS About To Send.

BpB Basic power Block.

CCA Clear Channel Assessment.
CSMA/CA Carrier Sense Multiple Access with Collision Avoidance.
CTS Clear To Send.

DDR Double-Data-Rate.
DPR Dynamic Partial Reconfiguration.
DSE Design Space Exploration.
DSL Domain-Specific Language.
DSP Digital Signal Processor.

EM Energy Manager.
ENO Energy Neutral Operation.
EPDG Energy Program Dependence Graph.
EWMA Exponentially Weighted Moving-Average.

FFT Fast Fourier Transform.
FPGA Field Programmable Gate Array.

GPP General Purpose Processor.
GPU Graphics Processing Unit.

HLS High Level Synthesis.
HMpSoC Heterogeneous Multiprocessor System-on-Chip.

IoT Internet of Things.
IP Intellectual property.

MAC Medium Access Control.
MCU Microcontroller Unit.
MDP Markov Decision Process.
MESC Multiple Energy Source Converter.
MILP Mixed integer linear programming.

OFDM Orthogonal Frequency Division Multiplexing.
OOK On-Off Keying.

PDR Packet Delivery Rate.

RAM Random Access Memory.
RF Radio Fréquence.
RFID Radio Frequency IDentification.
RL Reinforcement Learning.
RTL Register Transfer Level.
RTS Ready To Send.

TD Temporal-Difference.

ULP Ultra Low Power.

VHDL VHIC Hardware Description Language.

WSN Wireless Sensor Network.
WuB Wake up Beacon.
WuRx Wake-up Receiver.

Table des figures

1.1 Vision de la 5G (source Nokia [1]) : des gammes d'applications avec des spécifications variées. . . . 17
1.2 Vision technologique : gain énergétique des différents composants de l'IoT [2]. . . . 18
1.3 Architecture d'un nœud de capteurs à récupération d'énergie équipé d'une WuRx. 20
1.4 Flot de conception pour les architectures multi-cœurs hétérogènes. . . . 21

3.1 Seuil du tampon d'énergie. . . . 36
3.2 Exécution périodique de l'EM. . . . 36
3.3 Architecture des gestionnaires d'énergie avec et sans prédiction. . . . 38
3.4 Architecture de Fuzzyman. . . . 40
3.5 Fonctions d'appartenance utilisée par l'unité de floutage. . . . 40
3.6 Performances de Fyzzyman et P-FREEN en fonction de P_{AVG} pour différentes valeurs de σ_D. . . . 43
3.7 Architecture globale de RLMan. . . . 45
3.8 Débit moyen et efficacité énergétique en fonction de la valeur de la capacité pour de la lumière intérieure et du vent extérieur. . . . 48
3.9 Résolution centralisée vs distribuée. du calcul équitable des taux de génération de paquets. . . . 49
3.10 Architecture de commutation de sources d'énergie. . . . 49
3.11 Budget énergétique en fonction de l'état de charge. . . . 50

4.1 Architecture d'un nœud équipé d'une wake-up radio. . . . 52
4.2 Transmission d'un packet avec TI-WuR. . . . 52
4.3 Etat de l'art des WuRx. . . . 53
4.4 Nombre de faux réveils du PIC pendant une durée de 24 h dans un environnement intérieur. . . . 54
4.5 Modélisation générique d'un protocole avec des chaînes de Markov. . . . 55
4.6 Puissance consommée et latence des protocoles évalués : résultats analytiques et expérimentaux. . . . 56
4.7 Communication opportuniste utilisant OPWUM. . . . 58
4.8 Réseau utilisé pour comparer OPWUM et 1-hopMAC. Les flèches indiquent les récepteurs potentiels de chaque nœud. . . . 58
4.9 Energie dépensée et Taux de bonne réception des paquets des protocoles OPWUM et 1-hopMAC. . . . 59
4.10 Application de OPWUM aux réseaux de capteurs à récupération d'énergie. . . . 60
4.11 SNW-MAC : topologie utilisée et description du protocole. . . . 61
4.12 Structure du Wake-up Beacon utilisé par SNW-MAC. . . . 61
4.13 Architecture logicielle. . . . 62
4.14 Les protocoles X-MAC et PW-MAC. . . . 63
4.15 Architecture matérielle du nœud utilisant la récupération d'énergie et une WuRx : (a) description fonctionnelle (b) plateforme PowWow utilisée pour les expérimentations. 63
4.16 Déploiement du réseau en étoile. . . . 64
4.17 Résultats expérimentaux. . . . 65
4.18 Utilisation de plusieurs radios pour les communications hétérogènes : courte et longue portée. . . . 66

5.1 Plateforme d'émission-réception radio logicielle idéale. 72
5.2 Flot de conception HLS. 74
5.3 Flot de conception de la radio logicielle sur FPGA. 75
5.4 Structure de la description d'une forme d'onde avec le DSL proposé. 76
5.5 Structure générique d'une trame de données. 76
5.6 Structure d'interconnexions et de contrôle des IP. 78
5.7 Description de la trame IEEE 802.15.4 avec le DSL. 79
5.8 Implémentation d'un récepteur IEEE 802.15.4 avec le DSL. 80
5.9 Validation expérimentale de l'implémentation IEEE 802.15.4 sur les plateformes Nutaq Perseus 6010. 81
5.10 Exploration de l'espace de conception. 82
5.11 Génération d'un bloc de traitement reconfigurable dynamiquement. 83
5.12 Algorithme et performance du récepteur OFDM adaptatif. 85

6.1 Structure des architectures HMpSoC. 88
6.2 Représentation d'un HMpSoC basée sur la hiérarchie mémoire. 89
6.3 Flot de conception orienté énergie pour les architectures HMpSoC. 90
6.4 Vue globale de l'exploration proposée. 92
6.5 Hiérarchie mémoire de l'architecture Zynq. 93
6.6 Exploration de l'espace de conception pour l'application Matmult. 96
6.7 Traces de mesure de puissance pour l'application Matmult, avec 37.5 % des calculs en HW. 98
6.8 Energie consommée moyenne en fonction de la taille des données (N en Octets) pour différents canaux de communication. 99
6.9 Temps d'exécution et consommation d'énergie en fonction de l'utilisation HW : valeurs mesurées et estimées. 100
6.10 Plateforme d'émulation de HMpSoC. 103

7.1 Interception de signaux dans un contexte TEMPEST. 108
7.2 Topologie d'une architecture C-RAN. 109

Liste des tableaux

2.1 Nombre de publications par année et par types de publications. 27

3.1 Caractéristiques de différentes sources d'énergie (issues de [3]). 37
3.2 Ensemble de règles utilisé par le moteur d'inférence. 41
3.3 Valeurs des paramètres de simulations. Le détail des paramètres de P-FREEN, Fuzzyman et LQ-Tracker est donné dans les publications respectives. 47

4.1 Paramètres utilisés pour les expérimentations. 63

5.1 Estimation des ressources de l'émetteur et du récepteur IEEE 802.15.4. 81
5.2 Performance de la FFT pour la norme LTE avec la *Reconfiguration matérielle* et la *Reconfiguration logicielle*. 84

6.1 Implémentations HW pour l'application Matmult et Stencil. 99
6.2 Configurations obtenues par l'optimisation MILP. 101
6.3 Performances de l'exploration MILP. 101
6.4 Précision de l'exploration MILP : comparaison des performances estimées et mesurées. 102

Première partie

Introduction et synthèse des travaux

Chapitre 1

Contexte général et contributions

1.1 Vers une efficacité énergétique des objets communicants

Spécifications des objets communicants

L'apparition des dispositifs de communication sans fil a complètement modifié notre mode de vie, a permis aux sociétés modernes de fonctionner efficacement et a eu un impact majeur sur la politique, l'économie, l'éducation, la santé, le divertissement, la logistique, les voyages et l'industrie. Leurs performances ont ainsi considérablement augmenté au cours des 30 dernières années, véhiculant de plus en plus de données tout en conservant leur autonomie énergétique le plus longtemps possible. Cependant, le prochain majeur défi auquel ils devront faire face est leur nombre et l'omniprésence des utilisateurs. Cette capacité à être connecté partout de façon continue et transparente, en gardant la même qualité de service quel que soit l'environnement, implique que ces dispositifs puissent gérer différentes normes de communication sans fil, en choisissant pour chacun d'entre eux la configuration la plus efficace en énergie.

La prochaine génération de communication sans fil, la 5G, s'inscrit dans cette tendance et adressera de nouveaux type d'utilisateurs [4,5]. La Figure 1.1 montre les trois grandes familles d'applications envisagées pour la 5G avec l'avènement de nouveaux dispositifs sans fil : les systèmes de communications critiques et le déploiement massif de dispositifs faible énergie. Et donc, en plus des personnes, même les plus petits objets pourront envoyer leurs données sur ce qu'on appelle aussi l'Internet des Objets (IoT), de telle sorte que chaque utilisateur du monde puisse les atteindre. Ainsi, en plus d'une augmentation massive du nombre d'accès et de l'augmentation du débit, de nouvelles spécifications doivent être prises en compte.

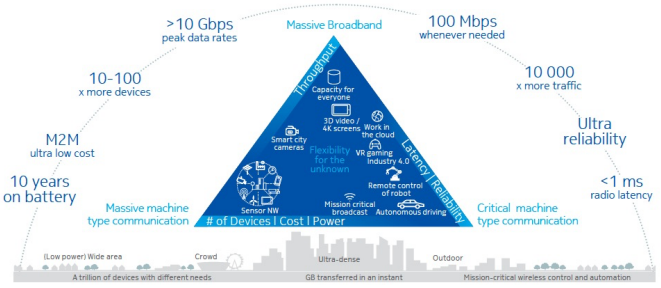


FIGURE 1.1 – Vision de la 5G (source Nokia [1]) : des gammes d'applications avec des spécifications variées.

L'application type des systèmes de communications critiques est le contrôle à distance d'objets réels et virtuels [6]. Son utilisation pour la réalité virtuelle et l'interaction homme-machine repose sur des communications ultra fiables de pilotage avec une infrastructure cloud. La latence requise de ces communications doit également être suffisamment faible pour permettre une communication entre les dispositifs via le réseau en moins de 1 ms.

Dans le cas du déploiement large échelle de dispositifs, la durée de vie de ces dispositifs (e.g. nœuds de réseau de capteurs sans fil, dispositifs IoT) doit considérablement être augmentée pour qu'ils soient déployés et massivement utilisés. Les systèmes cellulaires actuels ne sont pas suffisamment efficaces en énergie à la fois du coté de l'infrastructure réseau et au niveau des dispositifs sans fil. Une raison majeure est le gaspillage d'énergie en communications de contrôle pour la synchronisation et l'attribution des ressources.

La dernière spécification au cœur de la Figure 1.1 est la flexibilité que les objets connectés doivent

avoir pour adresser le nombre important de normes (ou de configurations possibles d'une norme). En effet, afin de répondre aux variations temporelles et spatiales d'un service (i.e. débit), afin de connecter plusieurs classes d'objets et de répondre au problème de saturation des ressources (qualité de service, spectre, énergie, ...), la nouvelle génération de communications sans fil doit être capable de changer ses caractéristiques en temps réel.

Par conséquent, pour relever ces défis (débit, latence, consommation d'énergie et flexibilité), des innovations sont nécessaires aux différentes couches du réseau. Les travaux développés dans ce document se concentrent principalement sur l'amélioration de l'efficacité énergétique des objets communicants pour des applications de types IoT et réseaux de capteurs. Une approche intercouche permettra d'optimiser globalement les systèmes avec d'une part, la prise en compte de spécificités applicatives pour le dimensionnement au plus juste des objets communicants et d'autre part l'impact des innovations technologiques sur les architectures de réseau et les protocoles de communications.

Systèmes embarqués

Les objets communicants sont des systèmes embarqués avec des ressources énergétiques limitées. Un des verrous à l'augmentation du débit des objets communicants et embarqués est donc de pouvoir réaliser un nombre important de traitements sous cette contrainte d'énergie finie. Les objets communicants pouvant être de simples étiquettes RFID ou de puissants smartphones, nos recherches considèrent deux grandes familles d'objets. D'une part, les objets ultra faible consommation ciblent des applications longue durée pour lesquelles le remplacement de la batterie n'est pas forcement requis. Ces objets ont des capacités de traitement limitées pour interagir avec leur environnement. Les exemples d'applications de ce type d'objets sont nombreux, allant de la surveillance environnementale à la réhabilitation fonctionnelle, la durée de vie de l'objet étant l'une des caractéristiques importantes pour ces applications. D'autre part, certains objets ont des capacités de traitement permettant d'interagir avec d'autres types d'objets ou avec leur environnement. L'efficacité énergétique de ces objets est étroitement liée à leur capacité de calcul et leur flexibilité. À titre d'exemple, les véhicules autonomes sont un domaine d'application représentatif qui combine de la vision, des communications sans fil avec d'autres noyaux de calculs intensifs, le tout avec un budget énergétique limité. Pour ces deux familles d'objets, réduire la consommation des traitements numériques embarqués est un challenge important.

Un objet communicant est classiquement constitué de plusieurs composants : une ou plusieurs unités de traitement (ou processeurs), de la mémoire, des capteurs, un ou plusieurs émetteurs-récepteurs et une source d'alimentation. Le processeur contrôle les autres composants et traite les données. Les capteurs sont utilisés pour mesurer les données de l'environnement. L'émetteur-récepteur permet la communication avec les autres objets. Enfin, la source d'alimentation fournit l'énergie nécessaire pour alimenter le système. La Figure 1.2 illustre l'idée de l'optimisation globale de l'efficacité énergétique des composants afin d'améliorer l'autonomie des objets communicants. Les gains technologiques escomptés pour la période 2012-2022 sont donnés.

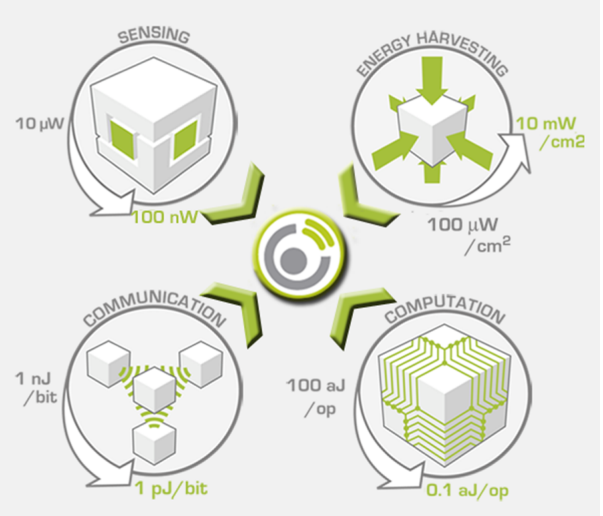


FIGURE 1.2 – Vision technologique : gain énergétique des différents composants de l'IoT [2].

S'ils sont directement influencés par les performances des capteurs et de l'alimentation, nos travaux se concentrent sur la relation entre le calcul et la communication, afin de permettre une gestion dynamique de l'énergie.

Du point de vue de la communication, la réduction énergétique due aux communications radio peut être réalisée selon deux objectifs principaux et complémentaires : (i) minimiser la puissance

d'émission tout en maintenant une qualité de liaison sans fil suffisante et (ii) minimiser les réveils et les écoutes de canaux inutiles tout en restant réactifs. De nouvelles techniques sensibles à la consommation d'énergie doivent ainsi être définies et mises en œuvre afin de pouvoir adapter dynamiquement les paramètres du système (e.g. puissance d'émission, précision des calculs, le protocole, ...) en fonction de l'environnement (conditions de propagation, énergie disponible, topologie du réseau, ...).

Dans ce contexte, la conception d'un nœud radio intelligent demande la définition d'une architecture adaptée à ses nouvelles spécificités. L'efficacité énergétique du nœud est un objectif majeur qui pourra être atteint grâce à des architectures matérielles et logicielles innovantes. La flexibilité de l'architecture doit permettre d'y associer des traitements adaptatifs permettant une réduction supplémentaire de la consommation. Si les circuits intégrés offrant le meilleur rendement énergétique sont les ASIC (Application-Specific Integrated Circuits), ces circuits spécifiques n'offrent pas la flexibilité souhaitée. A l'opposé, les processeurs (GPP - General Purpose Processor, DSP - Digital Signal Processor, GPU - Graphics Processing Unit) sont très flexibles au prix d'une consommation importante. Une alternative émergente consiste en des architectures hétérogènes qui prendront la forme d'un ou plusieurs cœurs de processeur (pour le coté programmabilité et flexibilité) couplés à des accélérateurs matériels reconfigurés dynamiquement (pour le coté capacité de calcul et efficacité énergétique). Un objectif important est donc de comparer l'efficacité énergétique de ces architectures émergentes à celles des architectures plus classiques.

1.2 Axes de recherche et structure du document

Mes activités de recherches sont directement liées aux préoccupations soulevées précédemment, la conception de systèmes de communications sans fil efficace en énergie avec comme approche globale l'adaptativité des systèmes afin d'explorer l'espace entre différentes performances (énergie, débit, latence, qualité de service, ...). Deux questions principales sont soulevées par ces nouvelles spécifications : (i) quelles sont les techniques de traitement du signal qui pourraient améliorer la qualité du lien, l'utilisation des ressources et l'efficacité énergétique ? (ii) quel type de matériel pourrait associer l'efficacité énergétique et le calcul haute performance de ces techniques de traitement du signal ?

Afin de proposer des solutions aux problèmes rencontrés, nous avons développé deux axes de recherche correspondant aux deux familles d'objets communicants définies précédemment : les réseaux de capteurs ultra faible consommation et la radio logicielle efficace en énergie.

1.2.1 Axe 1 : Les réseaux de capteurs ultra-faible consommation : gestion d'énergie et protocoles de communications

Les réseaux de capteurs (WSN - Wireless Sensor Networks) sont composés de plusieurs nœuds de capteurs sans fil qui surveillent un environnement et collectent des données vers un ou plusieurs lieux appelés puits. Ces réseaux constituent une technologie clé de l'IoT et sont essentiels dans de nombreuses applications telles que les villes intelligentes, les usines intelligentes ou l'agriculture de précision. Contrairement aux réseaux cellulaires ou IEEE 802.11, les WSN ne reposent sur aucune infrastructure fixe et, par conséquent, les nœuds s'auto-organisent de manière ad-hoc. Plusieurs topologies de réseau sont possibles pour l'organisation des nœuds, la plus simple étant le réseau en étoile dans laquelle tous les nœuds se trouvent à une distance d'un saut du puits et lui envoient ainsi directement leurs données. Dans le cas où la zone de surveillance est grande et où une partie des nœuds de capteurs est trop éloignée du puits pour une communication à un saut, un réseau multi-saut est utilisé.

Pour déployer de grands WSN, un point important est le coût des nœuds qui doit être maintenu bas. De plus, pour que certaines applications soient intéressantes, les WSN doivent pouvoir fonctionner pendant une longue période et, généralement, le point bloquant de ce type d'application est l'énergie. En effet, les nœuds sans fil traditionnels sont alimentés par des batteries individuelles, qui ne peuvent stocker qu'une quantité d'énergie limitée. Par conséquent, lorsqu'un nœud consomme toute l'énergie initialement stockée dans sa batterie, il s'éteint, et le remplacement des batteries n'est pas toujours possible si le réseau est dense ou déployé dans un environnement hostile. Par

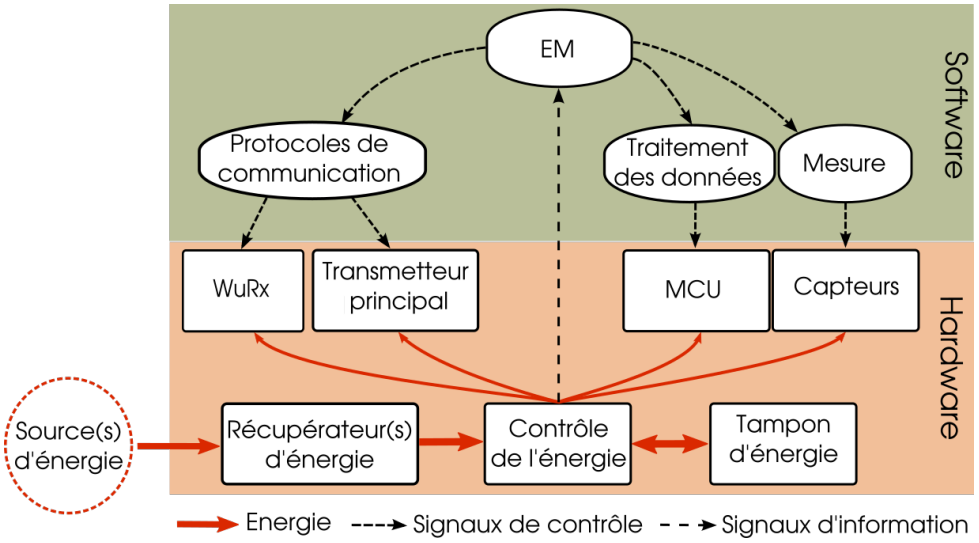


FIGURE 1.3 – Architecture d'un nœud de capteurs à récupération d'énergie équipé d'une WuRx.

conséquent, l'efficacité énergétique est critique dans les WSN car elle est étroitement liée à la durée de vie du réseau. Pour atteindre des WSN durables, la consommation d'énergie des nœuds doit être minimisée.

Une première solution est d'équiper chaque nœud d'un ou plusieurs récupérateurs d'énergie, et ainsi le rendre capable de s'alimenter exclusivement à partir de l'énergie récupérée dans son environnent. Les sources d'énergie pouvant varier au cours du temps, il est nécessaire d'adapter en conséquence la puissance consommée des nœuds. Cette adaptation est réalisée dans nos travaux par un composant logiciel appelé gestionnaire d'énergie (EM - Energy Manager).

De plus, la communication étant souvent la tâche la plus énergivore dans des réseaux de capteurs [7], les Wake-up Receivers (WuRx) sont utilisés dans ces travaux pour réduire le coût des communications. Les WuRx permettent l'écoute continue du canal sans fil tout en maintenant l'émetteur-récepteur principal en état de veille. Le prix d'une consommation d'énergie ultra basse est une faible sensibilité et de faibles débits binaires. Par conséquent, afin de profiter efficacement des WuRx, la conception de protocoles ad-hoc exploitant ces dispositifs émergents est nécessaire.

L'architecture typique des nœuds de capteurs considérés dans ces travaux est illustrée par la Figure 1.3. Le composant EM utilise les informations énergétiques du nœud (e.g. quantité d'énergie stockée, quantité d'énergie récupérée ...) pour contrôler diverses tâches du nœud (e.g. la détection, le traitement des données et les protocoles de communications).

1.2.2 Axe 2 : Les radios logicielles efficaces en énergie : prototypage rapide et exploration de l'espace de conception

La radio logicielle est une technologie qui adresse la mise en œuvre de formes d'ondes hautement flexibles et capables de s'adapter à l'environnement dans lequel elles opèrent. Pour ce faire, la radio logicielle s'appuie sur des plateformes matérielles capables d'assurer les besoins en flexibilité ainsi qu'en puissance de calcul. Principalement utilisées, les implémentations de radio logicielle basées sur des microprocesseurs généralistes présentent cependant certaines limites notamment liées aux performances énergétiques et aux débits atteignables. Dans ce contexte, les Field Programmable Gate Array (FPGA) apparaissent naturellement comme un support pertinent pour la mise en œuvre d'une radio logicielle tant pour leur puissance de calcul que pour leur capacité en matière de reconfiguration. Les FPGA disposent cependant d'un modèle de programmation qui contraste sévèrement avec l'esprit même de la radio logicielle en ce sens qu'ils requièrent l'utilisation de langages de description bas niveau. L'objectif de nos recherches consiste à mettre en avant (i) l'amélioration de la programmabilité par l'utilisation de spécifications de haut niveau exécutables et synthétisables et (ii) la réduction de l'énergie par une gestion efficace de la reconfiguration, très peu prise en compte dans l'état de l'art.

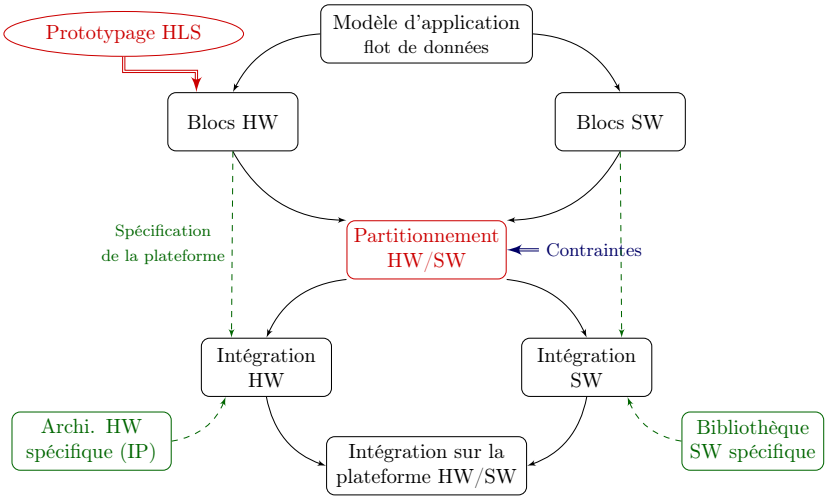


FIGURE 1.4 – Flot de conception pour les architectures multi-cœurs hétérogènes.

Nous avons dans un premier temps traité ce problème en proposant un flot de conception innovant, partant de modèles haut niveau des formes d'ondes jusqu'à leur implantation sur la plateforme ciblée. Ce flot vise à améliorer la programmabilité des traitements par l'utilisation de spécifications de haut niveau exécutables et synthétisables et repose principalement sur la synthèse de haut niveau (HLS - High Level Synthesis) ainsi que sur les architectures de traitement de signal de type flot de données. Cette approche permet une exploration de l'espace de conception, notamment sous contrainte énergétique.

De plus, la plupart des algorithmes utilisés est hautement parallèle et nécessite des performances de calculs importantes qui pourraient tirer profit des nouvelles architectures multi-cœurs hétérogènes (HMpSoC - Heterogeneous Multiprocessor System-on-Chip) combinant des processeurs généralistes (SW) et des accélérateurs matériels (HW). Pour ces architectures hétérogènes, les performances et la consommation d'énergie dépendent d'un large ensemble de paramètres tels que le partitionnement HW/SW, le type d'implémentation HW et le coût de communication entre les organes de calcul HW et SW conduisant ainsi à un immense espace de conception. Des méthodes sont proposées dans ces travaux afin de réduire la complexité de développement et de mise en œuvre d'applications efficaces en énergie sur des architectures multi-cœurs hétérogènes.

La Figure 1.4 présente une vue générale du flot de conception pour les architectures HMpSoC. L'utilisation de la HLS permet de garder un niveau de description élevé de l'application lorsque des blocs sont implémentés sur des cibles HW. La partitionnement HW/SW permet l'optimisation des performances en explorant l'espace de conception.

1.2.3 Structure du document et contributions

Le document est découpé en trois grandes parties : la première (Chapitres 1 et 2) introduit le contexte général et présente une synthèse de travaux de recherche, la deuxième (Chapitres 3 et 4) est dédiée à l'axe 1 : "Les réseaux de capteurs ultra-faible consommation : gestion d'énergie et protocoles de communications" et la troisième (Chapitres 5 et 6) présente l'axe 2 : "Les radios logicielles efficaces en énergie : prototypage rapide et exploration de l'espace de conception". Enfin, le Chapitre 7 dresse un bilan de mes activités de recherche et propose des perspectives à ces travaux.

Cette section introduit les différents chapitres du document. Les différentes contributions sont également données ainsi que les publications associées (les journaux en rouge, les conférences internationales en bleu et les démonstrations en noir).

Chapitre 2 : Synthèse des activités de recherche et d'enseignements Le premier chapitre est une présentation générale de mon parcours où j'aborde, de manière synthétique, l'ensemble de mes activités d'enseignant-chercheur. Ce premier chapitre dresse un bilan de mes implications dans les activités d'encadrements, de valorisations et d'administration de la recherche.

Chapitre 3 : Les réseaux de capteurs à récupération d'énergie La consommation d'énergie d'un nœud, qui est étroitement liée à la qualité de service (taux de détection, débit ...), est classiquement définie au déploiement à une valeur qui garantit la durée de vie requise. Cependant, les batteries ne pouvant stocker qu'une quantité d'énergie finie, le réseau est condamné à mourir. Une solution prometteuse pour augmenter la durée de vie des WSN est de permettre à chaque nœud de récupérer de l'énergie dans son environnement. Dans ce scénario, chaque nœud est équipé d'un ou plusieurs récupérateurs d'énergie, ainsi que d'un tampon d'énergie (batterie ou capacité) pour permettre de stocker une partie de l'énergie récupérée pour une utilisation future pendant les périodes de pénurie d'énergie. Diverses sources d'énergie sont possibles, telles que la lumière, le vent, le mouvement, la chaleur, ... [3, 8]. Les sources d'énergie étant généralement dynamiques et incontrôlées, il est nécessaire d'adapter dynamiquement la consommation électrique des nœuds, en ajustant leur qualité de service afin d'éviter les pannes de courant tout en maximisant l'efficacité énergétique et en garantissant le respect des spécifications applicatives. Cette tâche est réalisée par le gestionnaire d'énergie EM, qui est en charge de l'adaptation dynamique de la consommation des nœuds, et donc de leur qualité de service. Ce chapitre aborde la conception logicielle et matérielle de nœuds de capteurs à récupération d'énergie.

Les contributions associées à ce chapitre sont :

- un gestionnaire d'énergie utilisant la logique floue [aitaoudia16icc] [aitaoudia16dasip, legentil17sips],
- un gestionnaire d'énergie utilisant l'apprentissage par renforcement [aitaoudia18gcn] [aitaoudia17icc],
- une gestion de l'énergie distribuée pour les réseaux multi-saut [aitaoudia17wcl],
- un générateur de traces énergétiques diurnes pour la comparaison des gestionnaires d'énergie [aitaoudia15pimrc],
- une architecture de récupération d'énergie multi-source [gleonec16ogc, kazdoghli18vtc],
- l'application de la gestion d'énergie au standard LoRa [gleonec18ict,le18ict] [gleonec18ict_demo].

Chapitre 4 : Les wake-up radios De nombreuses techniques ont été proposées pour améliorer l'efficacité énergétique des protocoles de communication pour les WSN. En particulier, de nombreux travaux ont été réalisés au niveau des protocoles MAC (Medium Access Control), qui gèrent les communications point-à-point et contrôlent l'activité de l'émetteur-récepteur. Ils ont donc un fort impact sur l'efficacité énergétique des communications [9]. Un moyen efficace de réduire la consommation électrique des nœuds est le *duty-cycling*, qui consiste simplement à éteindre les émetteurs-récepteurs lorsqu'ils ne sont pas utilisés. Cependant, pour éviter la surdité, les nœuds doivent écouter périodiquement le canal pour recevoir des paquets [10]. Des progrès récents au niveau des circuits ont rendu possible la réalisation de WuRx à très faible consommation capables de "réveiller" efficacement un nœud lorsqu'un signal spécifique, appelé Wake-up Beacon (WuB), est envoyé par un nœud voisin [11–14]. De plus, de nombreuses WuRx intègrent des fonctionnalités de détection d'adresse, qui permettent aux nœuds de ne réveiller qu'un seul nœud spécifique et non tous leurs voisins [15, 16]. La consommation d'énergie ultra basse est obtenue au prix d'une faible sensibilité et d'un faible débit binaire, ce qui rend l'envoi des WuB coûteux en énergie. Par conséquent, la WuRx ne remplace pas le récepteur principal, mais est un dispositif supplémentaire à l'émetteur-récepteur traditionnel. Nos travaux consistent en l'évaluation des gains potentiels de l'utilisation des WuRx et en la proposition de protocoles ad-hoc exploitant ces dispositifs émergents.

Les contributions associées à ce chapitre sont :

- une modélisation des protocoles MAC utilisant les chaînes de Markov absorbantes [aitaoudia17ton] permettant une évaluation des gains potentiels des WuRx [aitaoudia16globecom],
- des protocoles permettant le choix opportuniste d'un relai dans un réseau multi-saut [aitaoudia16js] [djidi18icecs],
- un protocole pour la récolte d'information dans un réseau en étoile [aitaoudia16scube], protocole évalué dans une expérimentation alliant pour la première fois récupération d'énergie et WuRx [aitaoudia18sensors],
- une architecture de réseau alliant WuRx et le standard LoRa [aitaoudia17micpro] [aitaoudia16dsd] et la plateforme matérielle associée [magno17date].

Chapitre 5 : Description haut-niveau des formes d'ondes et prototypage rapide Ce chapitre aborde la définition et la mise en œuvre d'un flot de conception d'une application télécom (e.g. une couche physique, une forme d'onde, un protocole) s'exécutant sur des cibles matérielles de types FPGA. Ce problème est traité en proposant un flot de conception partant de modèles haut niveau de l'application jusqu'à son implantation sur la plateforme ciblée. Ce flot contourne les limites liées au modèle de programmation des FPGA en employant la HLS. En effet, la HLS est une approche qui permet de programmer les FPGA à partir de langages haut niveau comme le C/C++, relâchant ainsi la contrainte liée à son modèle de programmation. Afin de bénéficier de cette nouvelle approche dans le contexte de la radio logicielle, nous avons défini au dessus de la HLS un langage à usage spécifique (DSL - Domain-Specific Language) dans l'optique de faciliter le prototypage sur ce type de plateformes. La reconfigurabilité de l'architecture matérielle doit permettre également une flexibilité de l'application impliquant une modification de l'architecture. Cette flexibilité peut être obtenue, d'une part, à la conception avec le prototypage rapide permis par la HLS. D'autre part, la modification de l'application en cours d'exécution est également une caractéristique importante qui est abordée dans nos travaux par l'utilisation de la reconfiguration dynamique partielle.

Les contributions associées à ce chapitre sont :

- la proposition d'un langage de description spécifique aux formes d'ondes [ouedraogo14crowncom] et l'outil associé de génération automatique d'un code synthétisable [ouedraogo14jasp],
- l'exploration de l'espace de conception de formes d'ondes avec les outils de synthèse de haut niveau [gautier14dsd, tran16crowncom],
- des prototypages rapides sur des plateformes matérielles pour des formes d'ondes de type Zigbee [bhatnagar13vtc] et LTE [tran16dasip],
- une méthode d'adaptation de la dynamique des données d'un récepteur OFDM pour réduire sa consommation d'énergie [cladera15eusipco, cladera15isvlsi].

Chapitre 6 : Méthodes de conception pour les architectures multi-cœurs hétérogènes Ce chapitre aborde la réduction de la complexité de développement et de mise en œuvre d'applications efficaces en énergie sur HMpSoC. Si de nombreux travaux proposent des outils d'exploration de l'espace de conception (DSE - Design Space Exploration) avec des objectifs énergétiques, l'impact des communications entre les cœurs sur la consommation globale a peu été pris en compte. Tout d'abord, une définition formelle de la structure HMpSoC est introduite ainsi qu'une méthode de représentation générique axée sur la hiérarchie mémoire. Ensuite, un outil de modélisation et d'estimation rapide de l'énergie est proposé et validé sur plusieurs applications. Ce modèle énergétique sépare les sources d'énergie en trois catégories (calcul statique, dynamique et communications) et calcule leurs contributions sur la consommation globale de manière indépendante. Basée sur une étude précise des communications, cette approche calcule rapidement la consommation d'énergie pour une répartition donnée de l'application sur un HMpSoC. Dans un deuxième temps, nous proposons une méthodologie permettant l'exploration énergétique d'accélérateurs sur HMpSoC. Cette méthode s'appuie sur le modèle de consommation précédent couplé à une formulation de programmation linéaire en nombre entier mixte (MILP - Mixed Integer Linear Programming). Ce problème d'optimisation permet de sélectionner efficacement les types d'accélérateurs HW et le partitionnement HW/SW et ainsi d'obtenir une implémentation efficace en énergie pour une application tuilée. Les expériences réalisées ont montré la complexité du processus de validation des algorithmes de DSE sur une large gamme d'applications et d'architectures.

Les contributions associées à ce chapitre sont :

- une modélisation de la consommation d'énergie des HMpSoC basée sur les communications [roux16mcsoc] et validée sur une architecture Zynq,
- une méthode de partitionnement HW/SW d'applications reposant sur une estimation rapide de la consommation d'énergie [roux17fccm],
- une plateforme d'évaluation des HMpSoC combinant un générateur exhaustif d'applications et une plateforme SystemC d'émulation des architectures.

Autres contributions : D'autres contributions ont été réalisées depuis 2011, principalement dans le cadre d'encadrements d'étudiants en Master. Par soucis de cohésion, elles ne sont pas abordées dans ce document.

- **Communication optique :** des algorithmes aveugles de compensation des défauts d'un récepteur ont été proposés et appliqués aux communications optiques [nguyen16globecom]. Une validation expérimentale a été réalisée [nguyen16icc] en collaboration avec le laboratoire FOTON.
- **Réseaux de capteurs corporels :** des protocoles de communications ont été proposés pour ce type de réseau et permettent la réduction de l'énergie en adaptant le débit [nguyen16crowncom,nguyen14omnet] ou l'accès aux canaux fréquentiels [nhan14crowncom]. Un algorithme coopératif de localisation a également été proposé [hoang15icc].

Chapitre 2

Synthèse des activités de recherche et d'enseignements

2.1 Synthèse de la carrière

Depuis novembre 2011, je suis maître de conférences à l'Université de Rennes 1. Je suis rattaché au département Réseaux & Télécoms de l'IUT de Lannion pour mes activités d'enseignement. J'enseigne principalement en électronique et en télécommunications. Pour mes activités de recherche, je suis rattaché à l'équipe GRANIT du laboratoire IRISA (UMR 6074). Je suis membre titulaire du CNU en section 61 depuis 2015 et chef du département Réseaux & Télécoms depuis septembre 2018.

Synthèse des postes occupés :

De janv. 15 à ...	**Chercheur** au Laboratoire IRISA Membre de l'équipe GRANIT `http://www-granit.irisa.fr`
De nov. 11 à janv. 15	**Chercheur** au Laboratoire IRISA - INRIA Membre de l'équipe projet CAIRN `https://team.inria.fr/cairn/`
De nov. 11 à ...	**Enseignant** à l'IUT de Lannion Rattaché au département Réseaux & Télécoms
De oct. 09 à oct. 11	**Ingénieur de Recherche** au Laboratoire ANP du CEA-LETI Algorithmes et implémentations matérielles d'une radio intelligente (projets ANR Infop et IST QoSMOS)
De juin 08 à sept. 09	**Postdoctorat** au Laboratoire INRIA - CITI (INSA de Lyon) Architectures de récepteurs multi-antennes, multi-standards (CRE avec Orange Labs)
De sept.06 à mars 08	**Postdoctorat** à France Télécom R&D - Orange Labs à Grenoble Architectures de récepteurs large bande, multi-standards et reconfigurables (ANR Versanum)
De nov. 01 à avril 06	**Thèse de doctorat** au Laboratoire des Images et des Signaux de l'INPG (Grenoble) Algorithmes et architectures de récepteurs pour les systèmes multiporteuses par paquets d'ondelettes, soutenue le 27 avril 2006

2.2 Activités d'enseignements

Depuis ma nomination en tant que maître de conférences, mes activités d'enseignement sont effectuées essentiellement au département Réseaux & Télécoms de l'IUT de Lannion. J'assure également régulièrement des heures de cours à l'Ecole Nationale Supérieure de Sciences Appliquées et de Technologie (ENSSAT). Au niveau international, je suis intervenu deux années dans le Master ICT (Information and Communications Technology) de l'Université des Sciences et Techniques de Hanoi au Vietnam. L'ensemble des cours que j'assure de façon régulière depuis plusieurs années peut se répartir en trois grands volets : l'électronique analogique et numérique, le traitement du signal et les communications sans fil, certains de mes enseignements se situant bien sûr à l'interface de ces domaines.

Activités passées (2001 - 2011) : Auparavant, j'ai enseigné à l'école d'ingénieur Polytech'Grenoble (Université Joseph Fourier) en tant que moniteur de 2001 à 2004, puis dans le cadre d'un demi-poste d'Attaché Temporaire d'Enseignement et de Recherche (ATER) pendant l'année 2004-2005. Puis, j'ai effectué un autre demi-poste d'ATER à l'UFR de physique de l'Université Joseph Fourier durant l'année 2005-2006. Sur la période 2008-2011, j'ai effectué des vacations aux départements Génie Electrique et Télécoms de l'INSA de Lyon.
Enfin, en 2010 et 2011, j'ai effectué 4 h de cours dans le cadre du Master Recherche SIC & ESA de l'ENSEA à Cergy-Pontoise.
Sur la période 2001-2011, j'avais effectué un total d'enseignement de 461 heures équivalent TD.

Volume d'enseignements : Le tableau suivant synthétise la charge d'enseignement effectuée sur la période 2011-2018.

Année	**Service**	**Heures para.**[1]	**Vacations**[2]	**Total**
2017-2018	260 h	45 h	10 h	315 h
2016-2017	245 h	50 h	0 h	295 h
2015-2016	220 h	43 h	12 h	275 h
2014-2015	210 h	44 h	37 h	291 h
2013-2014	200 h	27 h	25 h	252 h
2012-2013[3]	182 h	35 h	0 h	217 h
2011-2012[4]	158 h	23 h	0 h	181 h

[1] Heures para-pédagogiques : tutorat étudiants, responsabilité communication, suivi de projet tutorés et de stages, entretiens de recrutement ...
[2] Heures d'enseignements hors-service effectuées à l'ENSSAT et l'USTH
[3] Décharge de 48 h. Heures supplémentaires non payées afin d'assurer les enseignements.
[4] Pas de décharge d'enseignement mais un service réduit dû à mon recrutement en novembre.

Présentation synthétique des enseignements : Le tableau suivant synthétise les enseignements effectués depuis 2011. Je suis responsable des deux premiers modules d'enseignement du tableau.

Matière	**Etablissement**	**Année(s)**	**Niveau**[1]	**Type**[2]	**Volume**[3]
Numérisation	IUT R&T	2011-...	1A	CM-TD-TP	30 h
Emetteur-Récepteur RF	IUT R&T	2011-...	1A	CM-TD-TP	30 h
Modulations numériques	IUT R&T	2012-...	1A	TP	24 h
Architectures des ordinateurs	IUT R&T	2013-...	1A	TD-TP	24 h
Signaux et systèmes	IUT R&T	2011-...	1A	TP	24 h
Modulations large bande	IUT R&T	2013-...	2A	CM-TD-TP	30 h
Electronique analogique	ENSSAT	2014-2016	1A	TP	12 h
Objets connectés	ENSSAT	2017-...	3A	CM-Projet	10 h
Digital Signal Processing	USTH	2013-2015	Master	CM-TD-TP	25 h

[1] 1A : Première année d'études dans l'établissement, soit bac+3 pour ENSSAT, et bac+1 pour l'IUT.
[2] CM : Cours Magistral, TP : Travaux Pratiques, TD : Travaux Dirigés.
[3] Le volume horaire est donné en heures équivalent TD.

Tâches parapédagogiques : De 2013 à 2018, j'ai été **Responsable de la communication** du département R&T, les principales missions sont les organisations de la journée portes ouvertes, de la remise des diplômes, des visites dans les lycées et les salons étudiants ainsi que la rédaction et mise en ligne d'actualités sur le site internet de l'IUT.
Depuis septembre 2018, je suis **Chef de département** du département R&T.

2.3 Activité de recherche scientifique

2.3.1 Présentation des thématiques de recherche

Comme présenté en introduction de ce mémoire, mes travaux de recherche s'articulent autour de 2 axes thématiques :

Axe 1 - Les réseaux de capteurs ultra-faible consommation : gestion d'énergie et protocoles de communications

Axe 2 - Les radios logicielles efficaces en énergie : prototypage rapide et exploration de l'espace de conception

Une grande partie de ma production scientifique et mon encadrement doctoral est directement associée à ces axes de recherche.

2.3.2 Publications et production scientifique

Je ne présente dans cette section qu'une vue synthétique de mon activité de publication depuis le début de ma carrière. La liste détaillée est présentée à la fin de ce document dans le Chapitre "Publications personnelles". La tableau 2.1 présente le nombre de publications par année et par type de publications. On notera une activité croissante avec un effort particulier sur les revues depuis 2016. Nos travaux étant très souvent validés sur des plateformes matérielles, un autre point important est la participation depuis 2016 à 4 sessions de démonstrations dans des conférences internationales.

L'ensemble des publications est disponible et téléchargeable sous HAL :

`https://hal.inria.fr/search/index/?q=%2A&authIdHal_s=matthieu-gautier`,

et sur ma page personnelle :

`http://people.irisa.fr/Matthieu.Gautier/publications/`.

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	**Total**
Revues Int.	-	1	-	-	2	1	-	-	1	-	1	3	2	**11**
Conf. Int.	1	1	1	4	1	4	1	2	4	4	10	2	4	**39**
Conf. Nat.	-	1	-	2	-	-	-	1	-	2	1	1	1	**9**
Colloque/Démos	1	-	1	1	-	-	-	1	-	-	6	3	1	**14**
Brevets	-	-	-	-	1	-	-	2	-	-	-	-	-	**3**
Chapitres	-	-	-	-	-	2	-	-	1	-	-	-	-	**3**

Revues Int. : Revues internationales avec comité de lecture et référencées JCR
Conf. Int. : Conférences internationales avec actes et comité de lecture
Conf. Nat. : Communications nationales avec actes et comité de lecture
Colloques/Démos : colloques nationaux / posters et démonstrations dans des conf. internationales
Brevets : Dépôts de brevets issus de mes travaux à Orange et au CEA-LETI
Chapitres : Chapitres de livres

Table 2.1 – Nombre de publications par année et par types de publications.

2.4 Encadrement doctoral et scientifique

Depuis mon intégration au sein de l'équipe CAIRN puis au travers de mes activités au sein de l'équipe GRANIT, j'ai eu l'occasion d'encadrer un certain nombre de Master Recherche et de doctorants (les encadrements d'étudiants ingénieurs et autres ne sont pas répertoriés).

Au 21 mai 2019, je suis impliqué dans l'encadrement de 7 thèses :

Axe 1 "réseaux de capteurs" :

- 1 thèse soutenue en 2017 (co-encadrée à 50% avec O. Berder)
- 2 thèses en cours (1 co-encadrée à 50% avec O. Berder, l'autre co-encadrée à 33% avec O. Berder et A. Courtay)

Axe 2 "radio logicielle" :

- 3 thèses soutenues en 2014, 2017 et 2018 (1 co-encadrée à 50% avec O. Sentieys, 1 à 70% avec O. Sentieys et 1 à 50% avec E. Casseau)
- 1 thèse en cours (co-encadrée à 33% avec O. Berder et R. Gerzaguet)

Depuis 2011, j'ai également participé à l'encadrement de 8 stages de masters et de 1 projet de fin d'étude (orienté recherche).

2.4.1 Direction de thèses

Liste des thèses soutenues : 4 thèses - 220 %

- Mai-Thanh Tran

Titre : Towards hardware synthesis of a flexible radio from a high-level language,
Commencée en Oct. 2013 et soutenue le 13 novembre 2018,
Co-encadrée à 50% avec E. Casseau,
Production : 1 conférence internationale (Crowncom [tran16crowncom]) et une démonstration (DASIP [tran16dasip]),
Actuellement Ingénieur à M3 Systems, Toulouse.

- Baptiste Roux

Titre : Methodology and tools for energy-aware task mapping on heterogeneous multiprocessor architectures,
Commencée en Oct. 2014 et soutenue le 23 novembre 2017,
Co-encadrée à 70% avec O. Sentieys,
Production : 2 conférences internationales (McSoC [roux16mcsoc], FCCM [roux17fccm]),
Actuellement Ingénieur de Recherche à Asygn, Grenoble.

- Faycal Ait-Aoudia

Titre : Energy harvesting wireless sensor networks leveraging wake-up receivers : energy managers and MAC protocols,
Commencée en Oct. 2014 et soutenue le 28 septembre 2017,
Co-encadrée à 50% avec O. Berder,
Production : 6 revues (Sensors [aitaoudia18sensors], IEEE TGCN [aitaoudia18gcn], Elsevier MICPRO [aitaoudia17micpro], IEEE WCL [aitaoudia17wcl], IEEE TON [aitaoudia17ton], Hindawi JoS [aitaoudia16js]), 7 conférences internationales (DATE [magno17date], PIMRC [aitaoudia15pimrc], ICCx2 [aitaoudia17icc] [aitaoudia16icc], Globecom [aitaoudia16globecom], DSD [aitaoudia16dsd], S-Cube [aitaoudia16scube]).
Actuellement Chercheur à Nokia Bell Labs, Nozay.

- Ganda-Stéphane Ouedraogo

Titre : Automatic synthesis of hardware accelerators from high-level specifications of physical layer waveform for flexible radio,
Commencée en octobre 2011 et soutenue le 10 décembre 2014,
Co-encadrée à 50% avec O. Sentieys,
Production : 1 revue (Eurasip JASP [ouedraogo14jasp]), 3 conférences internationales (IEEE VTC [bhatnagar13vtc], DSD [gautier14dsd], Crowncom [ouedraogo14crowncom]) et 1 chapitre de livre [gautier11crcpress],
Actuellement Ingénieur de Recherche à NestWave, Paris.

Liste des thèses en cours : 3 thèses - 116%

- Nour El Hoda Djidi

Titre : Energy management and protocol design for heterogeneous IoT networks leveraging wake-up radio,
Débutée en Oct. 2018,
Co-encadrée à à 33% avec O. Berder et A. Courtay.

- Corentin Lavaud

Titre : Systèmes reconfigurables pour l'interception de signaux sporadiques compromettants,
Débutée en Oct. 2018,
Co-encadrée à à 33% avec O. Berder et R. Gerzaguet.

- Philip-Dylan Gleonec

Titre : Design and implementation of power management strategies for long range radio modules with energy harvesting,
Débutée en Oct. 2015,
Co-encadrée à 50% avec O. Berder,
Production : 2 conférences internationales (IEEE OnlineGreenCom [gleonec16ogc], ICT [gleonec18ict]) et une démonstration (ICT [gleonec18ict_demo]).

2.4.2 Direction de stagiaires et de post-doctorants

Pour les 7 étudiants du Master SISEA de l'Université de Rennes 1, l'encadrement consiste au suivi de l'étude bibliographique réalisée par l'étudiant d'octobre à février (1 jour par semaine) et en la direction du stage de mars à septembre.

Nour El Hoda Djidi, Protocol design for wake-up radio nodes in heterogeneous IoT networks, Master en Systèmes Embarqués et Traitement de l'Information (SETI), Université Paris-Saclay, 2018, co-encadrée à 33% avec A. Courtay et O. Berder (1 conférence internationale : IEEE ICECS [djidi18icecs]).

Corentin Lavaud, Adaptive near sensor compressing for energy-constrained applications in Wireless Sensor Networks, Master SISEA, 2017-2018, co-encadré à 33% avec A.courtay et O. Berder (1 conférence soumise : IEEE ICC 2019).

Marwa Kazdoghli Lagha, Feature extraction and selection in multi-source energy harvesting Wireless Sensor Networks, École Polytechnique de Tunisie, 2017, co-encadrée à 50% avec O. Berder (1 conférence internationale : IEEE VTC-spring [kazdoghli18vtc].

Ti-Ti Nguyen, I/Q imbalance and blind compensation in digital communications, Master SISEA, 2014-2015, co-encadré à 50% avec P. Scalart (1 conférence internationale : IEEE Globecom [nguyen16globecom]).

Duc-Hung Luong, Evaluation of distributed beamforming for physical layer security in Wireless Sensor Networks, Master SISEA, 2013-2014, co-encadré à 50% avec O. Berder.

Gia-Minh Hoang, Localization via Ultra Wide Band Radio for Wireless Body Area Sensor Networks, Master SISEA, 2013-2014, co-encadré à 50% avec A. Courtay (1 conférence internationale : IEEE ICC [hoang15icc]).

Fernando Cladera, Dynamic Precision Scaling for low-power wireless Receivers based on OFDM, 2012-2013, co-encadré à 50% avec O. Sentieys (2 conférences internationales : Eusipco [cladera15eusipco] et IEEE ISVLSI [cladera15isvlsi], best paper award à ISVLSI).

Nhat-Quang Nhan, Spectrum agility for energy-efficient Wireless Body Area Sensor Networks, 2012-2013, co-encadré à 50% avec O. Berder (1 conférence internationale : Crowncom [nhan14crowncom]).

Xuan-Chien Le, Cooperative relay selection through game theory approach, 2012-2013, co-encadré à 50% avec O. Berder.

J'ai également participé à l'encadrement de **Xuan-Chien Le**, post-doctorant de mai 2017 à mai 2018 dans le cadre du projet PME ALAMO. Il a travaillé sur les performances théoriques du standard LoRa (1 conférence internationale : ICT [le18ict]).

2.5 Diffusion et rayonnement

Prix et distinctions scientifiques :

- **EURASIP best paper award 2016** pour le journal [noguet11jwcn] intitulé "Advances in opportunistic radio technologies for TVWS" publié avec Dominique Noguet et Vincent Berg dans Eurasip Journal on Wireless Communications and Networking en 2011
- **IEEE ISVLSI best paper award 2015** pour l'article de conférence [cladera15isvlsi] intitulé "Energy-Aware Computing via Adaptive Precision under Performance Constraints in OFDM Wireless Receivers" publié avec Fernando Cladera et Olivier Sentieys à la conférence IEEE Computer Society Annual Symposium on VLSI à Montpellier en juillet 2015
- **IARIA COCORA best paper award 2011** pour l'article de conférence [gautier11cocora] intitulé "Cyclostationarity Detection of DVB-T Signal : Testbed and Measurement", publié avec Marc Laugeois et Philippe Hostiou à la conférence IARIA International Conference on Advances in Cognitive Radio à Budapest, Hongrie, en 2011

Conférence invitée :

- La référence [aitaoudia16scube] est un article de conférence invité de la conférence EAI S-cube

Participation à un jury de thèse :

- Examinateur de la thèse de Marwa Charfi, Centrale Supélec, soutenue le 7 octobre 2016 et encadrée par J. Palicot et R. Gribonval

Participation à des comités de sélection :

- Membre externe du comité de sélection pour le poste 63MCF0175 au département Systèmes et Réseaux de Communications (SRC) de l'INSA de Rennes, 2018
- Membre externe du comité de sélection pour le poste d'enseignant-chercheur contractuel (CDD 3 ans) au département SRC de l'INSA de Rennes, 2016
- Membre externe du comité de sélection pour le poste 61MCF4048 au département Électronique et Informatique Industrielle (EII) de l'INSA de Rennes, 2015

Membre du Comité de Programme Technique des conférences (depuis 2011) :

- ICT 2018
- IEEE PIMRC 2013, 2014, 2015, 2016, 2017, 2018
- IEEE WCNC 2013, 2014, 2016, 2019
- IARIA AICT 2018
- IARIA SENSORCOM 2017, 2018
- IARIA COCORA 2013, 2014, 2015, 2016
- EAI CROWNCOM 2016, 2017, 2018
- ISWCS 2015, 2016, 2017, 2018
- IEEE ICCVE 2013, 2014
- IEEE Percom 2013 (Workshop on Cognitive Computing and Communications)

Relecteurs d'article de revues (depuis 2011) :

- IET The Journal of Engineering (x1)
- EURASIP Journal on Wireless Communications and Networking (x1)
- MDPI Sensors (x6)
- IEEE Transactions on Communication (x2)
- IEEE Transactions on Signal and Information Processing over Networks (x1)
- IEEE Wireless communication letters (x1)

Présidence de 6 sessions de conférences internationales :
- IEEE WCNC 09 : Session on “Cognitive Radio Networks”
- IEEE WCNC 09 : Poster session on “Selected topics in Wireless Systems”
- IEEE VTC 2013 : Poster session on “Cognitive Radio and Spectrum Sensing”
- IEEE VTC 2013 : Session on “Spectrum Sensing”
- EAI Crowncom 2016 : Session on “Networking Protocols for Cognitive Radio”
- IEEE ICC 2017 : Session on “Ad Hoc Networks”

Rayonnement et activités internationales :
- En 2014 et 2015, j’ai réalisé 2 séjours de 10 jours à **Hanoi (Vietnam) à l’USTH**. L’objectif premier est une intervention en première année du Master USTH avec un cours de 25 h en traitement numérique du signal. L’objectif second est le développement d’activités de recherche communes : présentation de travaux autour des réseaux de capteurs et participation au groupe de travail du consortium USTH pour créer un laboratoire commun.
- Depuis 2015, je travaille en étroite collaboration avec **Michele Magno de l’ETH Zurich** autour de la thématique émergente de la wake-up radio. Nous avons accueilli 2 fois Michele en 2014 et 2016 pour des séjours de 1 et 2 semaines. Mon doctorant, Faycal Ait Aoudia, a effectué un séjour de 6 mois à l’ETH. De nombreuses publications résultent de cette collaboration.

Implication dans les organismes d’animation de la recherche : J’essaie de participer activement aux instances d’animation de la communauté au travers des réunions et des colloques des GDR et d’autres structures d’animation. Ces lieux d’échanges permettent régulièrement l’émergence d’idées nouvelles et de projets nationaux.
- Intervention à la journée thématique “Near Sensor Computing” du GDR SoC2 : “Gestion de l’énergie dans les réseaux de capteurs multi-sauts à récupération d’énergie : traitement distribué et protocole opportuniste”, novembre 2017.
- Intervention à la journée thématique “Plateformes d’évaluations radio” du GDR ISIS : “High performance software radios : concepts, tools and applications”, mars 2017.
- Pour la demande de renouvellement du GDR (2019-2023), j’assurerai l’animation de l’action “Architectures et applications” du thème D (voir site du GDR : `http://www.gdr-isis.fr/uploads/documents/Renouvellement-ISIS-2019-2023.pdf`).

2.6 Responsabilités scientifiques

Responsabilités et mandats nationaux
- Je suis actuellement **membre élu titulaire au CNU 61** (2015-2019).
- J’ai été membre élu de la **Commission recherche** de l’IUT de Lannion (2013-2018).

Contrats de recherche collaboratif : Depuis mon recrutement en 2011, j’ai participé et monté des collaborations autour de projets et de contrats de recherche. Ces projets ont évidemment permis de renforcer mes axes de recherche personnels, mais ont aussi permis de développer de nouveaux axes de travail.
J’ai ainsi été impliqué dans 5 projets de recherche collaboratifs. De nombreux autres ont été soumis, je souhaiterais juste mentionner mon implication dans le montage du projet “Guardian Angels for a Smarter Life”, projet FET Flagship Pilot de la commission européenne. Malheureusement non-sélectionné, cet immense projet couvrait un spectre très large autour des réseaux de capteurs. J’étais impliqué (workshop, meeting) dans la roadmap communication afin de proposer les axes de recherche que l’IRISA souhaitait développer dans le projet.

Projet en cours :

- **Projet ANR Wake-up (2017-2021)**
  Le projet Wake-up (Comment l’apparition de la wake-up radio va révolutionner les réseaux M2M hétérogènes) a été accepté pour financement à l’appel à projet générique de l’ANR en

2017. Le consortium est composé du CEA-LETI, des Universités de Strasbourg et Rennes 1, et d'une PME (Wi6labs). L'objectif est d'étudier l'impact de la wake-up radio sur la conception de réseaux hétérogènes (courte et longue portées) efficaces en énergie et faibles latences. Mon implication dans ce projet est la participation (avec O. Berder qui coordonne le projet) au montage du projet avec les partenaires et la responsabilité technique du work-package 4. Budget : 160 k€ avec 6 mois d'ingénieur et 1 thèse.

- **Projet PME "Images & Réseaux" – Alamo (2016-2018)**
  Le projet Alamo (Autonomous Long rAnge radio Modules for cOnnected farms) est réalisé en collaboration avec 2 PME de Lannion. L'objectif est le transfert technologique des méthodes de gestion d'énergie développée dans l'axe "réseaux de capteurs ultra-faible consommation" décrit précédemment et la montée en compétence sur les technologies de radio longue portée. Mon implication dans ce projet est la participation (avec O. Berder) au montage du projet avec les PME. J'assure également le suivi et la réalisation technique avec un ingénieur de recherche et un post-doctorant recrutés pour ce projet.
  Budget : 150 k€ avec 6 mois d'ingénieur et 12 mois de Post-doc.

Projets terminés :

- **Projet PME "Images & Réseaux" – Embrace (2014-2016)**
  Le projet Embrace (Embedded Radio Accelerator) a été réalisé en collaboration avec 2 PME bretonnes. L'objectif était le transfert technologique des méthodes de conception que nous développons dans l'axe "radio logicielle efficace en énergie" décrit précédemment.
  Mon implication dans ce projet était très forte car directement lié à mes activités de recherche. J'ai effectué le montage du projet (100% de la partie IRISA) avec les PME et assuré le suivi et la réalisation technique avec un ingénieur de recherche et un post-doctorant recrutés pour ce projet.
  Budget : 170 k€ avec 18 mois d'ingénieur et 12 mois de Post-doc.

- **Equipex FIT Future Internet of Things (2011-2014)**
  Participation à l'Equipex FIT porté par le laboratoire CITI de l'INSA de Lyon, nous avons réalisé des développements FPGA sur la plateforme NutaqPerseus 6010. Mon rôle était le suivi technique de l'ingénieur de recherche recruté pour ce projet.
  Budget : 24 mois d'ingénieur sur des supports INRIA.

- **Projet ANR FAON Frequency based Access Optical Networks (2012-2015)**
  Mon rôle était le suivi technique de l'ingénieur de recherche recruté pour ce projet et la rédaction des livrables.
  Budget : 24 mois d'ingénieur.

Autres financements :

- **Défis scientifiques émergents de l'Université de Rennes 1 (2016)** :
  Aide financière pour la création de nouvelle thématique, ce projet a permis d'appuyer nos travaux autour de la technologie émergente de la wake-up radio. Il a permis de renforcer notre collaboration avec l'ETH Zurich en financement une partie du séjour du doctorant Fayçal Ait Aoudia à Zurich.
  Budget : 8 k€

- **Contrat d'encadrement doctoral avec Wi6labs (2015-2018)** :
  Ce projet est un contrat d'accompagnement pour l'encadrement de Philip Dylan Gléonec qui est en thèse CIFRE dans la start-up rennaise Wi6labs.
  Budget : 30 k€

Deuxième partie

Les réseaux de capteurs ultra-faible consommation : gestion d'énergie et protocoles de communications

Chapitre 3

Les réseaux de capteurs à récupération d'énergie

Ce chapitre présente nos activités autour de la récupération d'énergie. La récupération continue d'énergie permet le fonctionnement à long terme des nœuds de capteurs sans fil mais elle doit être associée à des algorithmes de gestion d'énergie (EM - Energy Manager) afin de garantir ce fonctionnement. Le rôle de l'EM est d'adapter dynamiquement l'énergie consommée en fonction des fluctuations de l'énergie récupérée. Les contributions présentées dans ce chapitre visent principalement à concevoir et à implémenter de nouveaux EM capables de s'adapter aux contraintes environnementales et matérielles.

La première section (Section 3.1) présente la problématique de la gestion de l'énergie dans les réseaux de capteurs à récupération d'énergie, ainsi qu'un état de l'art des gestionnaires. Les Sections 3.2 et 3.3 présentent nos deux contributions principales de gestionnaires : le premier EM reposant sur la théorie de la logique floue, le second sur le principe de l'apprentissage par renforcement (RL - Reinforcement Learning). Enfin, la dernière section résume nos autres activités autour de la récupération d'énergie, notamment l'application à un cas concret : le standard de radio longue portée LoRa.

Sommaire

3.1 Vers une autonomie énergétique des réseaux de capteurs 35
3.1.1 Récupération et gestion d'énergie 35
3.1.2 État de l'art des gestionnaires d'énergie 38
3.2 Fuzzyman : Utilisation de la logique floue 39
3.2.1 Conception de Fuzzyman 39
3.2.2 Évaluation de Fuzzyman 42
3.2.3 Conclusion 43
3.3 RLman : Utilisation de l'apprentissage par renforcement 44
3.3.1 Formulation du problème de gestion de l'énergie récupérée 44
3.3.2 Algorithme de RLMan 45
3.3.3 Évaluation de RLMan 46
3.3.4 Conclusion 48
3.4 Autres contributions 48
3.4.1 Gestion d'énergie distribuée pour les réseaux multi-sauts 48
3.4.2 Architecture de récupération d'énergie multi-source 49
3.4.3 Application de la gestion d'énergie à des nœuds LoRa 49
3.5 Conclusion 50

3.1 Vers une autonomie énergétique des réseaux de capteurs

3.1.1 Récupération et gestion d'énergie

Dans les réseaux de capteurs à récupération d'énergie, chaque nœud est équipé d'un ou plusieurs dispositifs de récupération d'énergie, permettant aux nœuds d'être entièrement alimentés

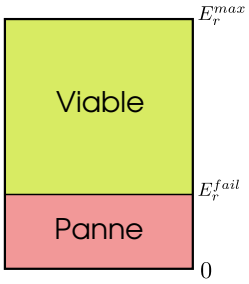


FIGURE 3.1 – Seuil du tampon d'énergie.

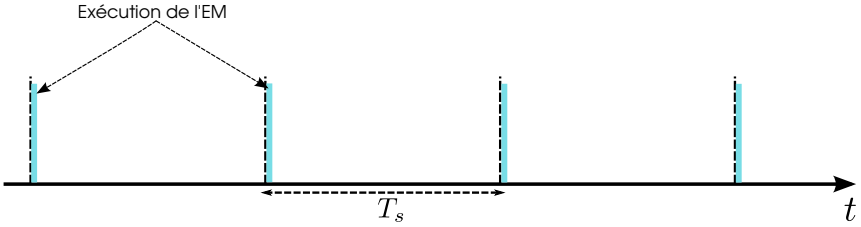


FIGURE 3.2 – Exécution périodique de l'EM.

par l'énergie récupérée dans leur environnement. Afin de stocker une partie de l'énergie récupérée pour une future utilisation, chaque nœud intègre également un dispositif de stockage d'énergie, qui peut être une super-capacité ou une batterie. Ce tampon d'énergie a une capacité finie notée E_r^{max} et un seuil de panne noté E_r^{fail}, comme illustré sur la Figure 3.1. Le seuil de panne correspond à la quantité minimale d'énergie requise pour que le nœud fonctionne, et une panne survient si l'énergie résiduelle passe en dessous de ce seuil. La capacité E_r^{max} est la quantité maximale d'énergie que le tampon d'énergie peut stocker. Cette capacité finie est à l'origine du gaspillage d'énergie. En effet, si l'énergie n'est pas immédiatement utilisée après sa récupération, elle est stockée dans le tampon d'énergie. Cependant, si le tampon d'énergie est plein, alors l'énergie récupérée ne sera pas stockée, et par conséquent gaspillée.

Différentes sources d'énergie peuvent être envisagées en fonction de l'environnement dans lequel les nœuds sont déployés et de la quantité d'énergie requise par l'application. La puissance récupérée dépend fortement du type de source d'énergie, ainsi que du rendement du récupérateur, i.e. le rapport entre la puissance extraite de la source par la récupérateur et la puissance de la source. La Table 3.1 donne les caractéristiques de quelques sources d'énergie (issues de [3]). Comme on peut le voir, la puissance récupérée de la lumière ambiante extérieure est 10^5 supérieure à celle obtenue avec la récupération d'énergie RF (Radio Fréquence), montrant la forte variation d'énergie disponible d'une source à l'autre. De plus, les sources présentent classiquement de fortes variations temporelles. Ce comportement motive la nécessité d'une adaptation dynamique de la performance des nœuds. En effet, l'augmentation de la qualité de service d'une application WSN (e.g. le taux de mesure, le débit...) demande une consommation d'énergie plus élevée. Une solution naïve pour obtenir un WSN durable consiste à paramétrer les nœuds à une qualité de service constamment faible, avec ainsi une consommation électrique suffisamment faible pour éviter les pannes de courant. Cette solution présente deux inconvénients principaux : (i) elle conduit à de mauvaises performances car une quantité importante d'énergie est gaspillée et (ii) une connaissance précise de la source d'énergie est nécessaire pour garantir qu'aucune panne de courant ne se produira. Une autre solution consiste à effectuer une adaptation dynamique de la performance, permettant ainsi une maximisation de la qualité de service tout en évitant les pannes de courant.

En pratique, l'EM est un programme logicielle qui est embarqué dans chaque nœud et qui est exécuté périodiquement comme illustrée par la Figure 3.2. Le durée entre deux exécutions de l'EM est notée T_s. Ainsi, l'énergie résiduelle $e_r[k]$ à l'instant k est définie par :

$$e_r[k] = e_r[0] + \sum_{i=0}^{k-1} e_h[i] - \sum_{i=0}^{k-1} e_c[i] - \sum_{i=0}^{k-1} e_w[i], \;\; k = 1, 2, \ldots \tag{3.1}$$

avec $e_r[0] \in [0, E_r^{max}]$ l'énergie résiduelle initiale, et $e_h[k]$, $e_c[k]$ et $e_w[k]$ sont respectivement l'éner-

Source	Puissance de la source	Puissance récupérée
Lumière ambiante		
Indoor	0.1 mW/cm^2	10 μW/cm^2
Outdoor	100 mW/cm^2	10 mW/cm^2
Vibration/Mouvement		
Humain	0.5 m à 1 Hz 1 m/s^2 à 50 Hz	4 μW/cm^2
Industriel	1 m à 5 Hz 10 m/s^2 à 1 kHz	100 μW/cm^2
Énergie thermique		
Humain	20 mW/cm^2	30 μW/cm^2
Industriel	100 mW/cm^2	1 − 10 mW/cm^2
RF		
Station de base GSM	0.3 μW/cm^2	0.1 μW/cm^2

Table 3.1 – Caractéristiques de différentes sources d'énergie (issues de [3]).

gie récupérée, l'énergie consommée et l'énergie gaspillée pendant l'intervalle de temps k.

En supposant que la qualité de service de l'application peut être mesurée par une fonction d'utilité $\mathcal{U}(e_c[k],\ k = 0, 1, \dots)$, qui dépend de l'énergie consommée (e.g. taux de paquets, efficacité énergétique...), le problème de gestion d'énergie peut être formulé par :

$$\begin{aligned} \underset{e_c[k],\ k=0,1,\dots}{\text{maximize}} \quad & \mathcal{U}(e_c[0], e_c[1], \dots) \\ \text{subject to :} \quad & e_r[k] \leq E_r^{max},\ k = 1, 2, \dots \\ & e_r[k] \geq E_r^{fail},\ k = 1, 2, \dots \end{aligned} \tag{3.2}$$

Dans ce problème, la première contrainte est due à a capacité finie du tampon d'énergie. La deuxième permet d'éviter les pannes d'énergie, et satisfaire cette contrainte permet d'attendre des opérations neutres en énergie (ENO - Energy Neutral Operation) [17]. Afin de maximiser l'efficacité énergétique, les e_c doivent être choisies afin qu'aucune énergie ne soit gaspillée, ou plus formellement :

$$e_w[k] = 0,\ k = 1, 2, \dots \tag{3.3}$$

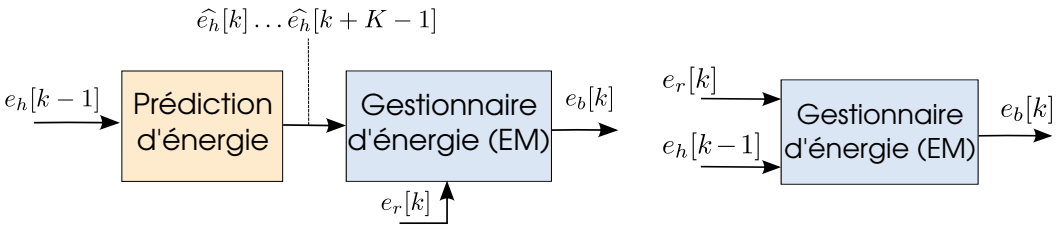
Satisfaire cette contrainte revient à atteindre l'ENO-MAX [18].

Cependant, pour certaines applications, atteindre l'ENO-MAX n'est pas un objectif approprié. Dans les réseaux multi-sauts par exemple, où chaque nœud génère et relaie des paquets, un nœud peut avoir un taux de génération de paquets limité, non pas par la quantité d'énergie récupérée mais par ses relais qui ne peuvent pas transmettre les paquets générés. Dans un tel scénario, gaspiller de l'énergie peut être nécessaire si le tampon d'énergie n'est pas assez grand pour stocker l'énergie récupérée en excès. Dans cette étude, l'accent est mis sur les réseaux en étoile, et dans de tels scénarios, l'efficacité énergétique, i.e. la capacité d'un EM à minimiser l'énergie gaspillée, est pertinente pour évaluer l'EM. Par conséquent, l'efficacité énergétique, définie formellement par :

$$\xi = 1.0 - \frac{\sum_k e_w[k]}{\sum_k e_h[k] + e_r[0]}, \tag{3.4}$$

sera utilisée comme métrique pour évaluer les stratégies de gestion d'énergie. On considèrera également le taux de panne D_R, qui est défini comme le rapport entre le temps passé en état de panne et la durée totale de l'application.

Challenges pour la conception des EM Un des challenges est dû à l'incapacité de connaitre (ou mesurer) l'énergie récupérée e_h. Comme nous le verrons dans la partie suivante, certaines approches conçoivent des EM utilisant des prédicteurs qui visent à prédire l'énergie récupérée future



(a) Gestionnaires d'énergie avec prédiction. (b) Gestionnaires d'énergie sans prédiction.

FIGURE 3.3 – Architecture des gestionnaires d'énergie avec et sans prédiction.

sur une fenêtre de temps finie. Cependant, l'utilisation de prédicteurs nécessite des mesures de la quantité d'énergie récupérée aux instants précédents $e_h[k-1], e_h[k-2], \dots$, qui sont difficiles à obtenir en pratique. Par conséquent, certaines approches reposent uniquement sur l'énergie résiduelle pour prendre une décision. Une autre difficulté de la gestion de l'énergie est le contrôle précis de la consommation électrique du nœud. En effet, certaines tâches exécutées par un nœud n'ont pas de coûts énergétiques déterministes. Les tâches de communication ont par exemple un coût en énergie qui dépend fortement de la qualité du canal, une échec dans les transmissions de paquets conduisant généralement à des retransmissions. Certains EM ne considèrent pas ce problème : ils calculent uniquement un budget d'énergie $e_b[k]$, correspondant à la quantité d'énergie que le nœud peut dépenser pendant l'intervalle de temps k ; ils ne s'intéressent pas à la façon dont ce budget énergétique est réparti entre les différentes tâches exécutées par le nœud. Ces EM ne fixent donc pas directement la qualité de service, mais la contrôlent indirectement via le budget énergétique. Dans ce cas, un autre composant logiciel est chargé de définir la qualité du service (e.g. le taux de détection) à partir du budget énergétique. D'autres approches définissent directement certains paramètres de qualité de service, et contrôlent donc indirectement la consommation d'énergie du nœud. Dans nos travaux, le taux de génération des paquets, notée χ_g, est souvent le paramètre de contrôle de la consommation d'énergie. Enfin, un autre défi est que les nœuds de capteurs sans fil sont des systèmes très contraints en ressources. En effet, de tels dispositifs ont généralement très peu de puissance de calcul (fréquence typiquement inférieure à 16 MHz), de mémoire (typiquement 2 ko) et stockage (typiquement moins de 100 ko), et les algorithmes proposés doivent donc être adaptés à ces types de plateformes.

3.1.2 État de l'art des gestionnaires d'énergie

De nombreux algorithmes de gestion d'énergie [17, 18, 18–25] ont été proposés ces dernières années afin de proposer des algorithmes d'adaptation efficaces, adaptés aux ressources limitées fournies par les nœuds. Les différents EM peuvent être classés en fonction de leurs besoins ou non de prédictions sur la quantité d'énergie qui pourra être récupérée dans le futur, i.e. avec prédiction et sans prédiction.

Gestionnaires d'énergie avec prédiction

L'architecture générale d'un EM basé sur des prédictions est représentée sur la Figure 3.3a. Elle repose sur un prédicteur d'énergie fournissant à l'EM des prédictions de l'énergie pouvant être récupérée dans le futur. Pour un intervalle de temps k, le prédicteur d'énergie utilise la valeur de $e_h[k-1]$ pour calculer les prédictions de l'énergie récupérée $\widehat{e}_h[k] \dots \widehat{e}_h[k+K-1]$ sur une fenêtre de K intervalles. L'EM utilise ensuite ces prédictions, ainsi que l'énergie résiduelle actuelle, pour prendre une décision sur le budget énergétique $e_b[k]$.

Le premier EM utilisant l'approche par prédiction a été introduit en 2007 par *Kansal et al.* [17]. Dans ce schéma, un filtre moyenneur à lissage exponentiel (EWMA - Exponentially Weighted Moving Average) est utilisé pour prédire la quantité future d'énergie récupérée, et un duty-cycle est calculé en fonction de l'écart entre les énergies prévues et observées. *Castagnetti et al.* ont introduit Closed-Loop Power Manager (CL-PM) dans [19] qui utilise deux stratégies de gestion d'énergie distinctes, une pour les périodes où l'énergie est disponible dans l'environnement, et une pour les périodes pendant lesquelles l'énergie récupérée est inférieure à un seuil fixe. *Le et*

al. ont proposé Wake-up Variation Reduction Power Manager (WVR-PM) [20], une variante de CL-PM qui permet à un nœud de stocker plus d'énergie lorsque l'énergie récupérée est importante afin d'obtenir une qualité de service similaire pendant les périodes de carences énergétiques que celle atteinte lorsque l'énergie est disponible. *Renner et al.* ont proposé un algorithme basé sur la prédiction [21] pour les nœuds de capteurs utilisant une super-capacité comme dispositif de stockage d'énergie. L'algorithme consiste à fournir une prévision de l'énergie résiduelle pour une consommation donnée et une prévision de récupération donnée. L'algorithme a été implémenté sur un réseau de douze nœuds avec une topologie multi-saut, et l'expérimentation a duré quatre semaines.

La majorité des algorithmes proposés suppose la possibilité de pouvoir prédire l'énergie récupérée [26], la conception de prédicteurs précis étant un domaine de recherche actif [27]. Cependant, les prédicteurs sont généralement spécifiques à un type de source d'énergie (e.g. RF [28], solaire [29], vent [30], ...) et certaines sources sont difficilement prédictibles. Ainsi, des techniques sans prédiction sont également proposées.

Gestionnaires d'énergie sans prédiction

La quantité d'énergie qu'un capteur peut récupérer pouvant montrer de grandes variations et être difficile à prévoir, les prédicteurs d'énergie peuvent engendrer des erreurs significatives entraînant une sur-utilisation ou une sous-utilisation de l'énergie récupérée [22]. Ainsi, des techniques permettent de calculer un budget énergétique en utilisant aucunes prédictions ou aucuns modèles de la source d'énergie. La Figure 3.3b donne la structure de ce type d'EM. LQ-Tracker [18] utilise Linear-Quadratic Tracking, une technique issue la théorie du contrôle adaptatif, afin d'adapter le duty-cycle et n'utilise que l'énergie résiduelle comme entrée. De manière similaire, *Le et al.* ont proposé d'utiliser un contrôleur Proportional Integral Derivative (PID) [23]. Avec P-FREEN [22], *Peng et al.* ont conçu un EM qui maximise le duty-cycle d'un nœud en présence d'inefficacités de stockage de l'énergie. Les auteurs ont formulé le problème de maximisation du duty-cycle en problème de programmation non linéaire et ont proposé un ensemble de règles d'attribution de budget énergétique en utilisant à la fois l'énergie récupérée (actuellement observée) et l'énergie résiduelle.

Ces approches ont montré qu'il est possible de réaliser une gestion efficace de l'énergie sans le surcoût engendré par les mécanismes de prédiction en termes de mémoire et de quantité de calcul. Par conséquent, nos travaux proposent principalement des schémas sans prédiction. Le récent EM P-FREEN est similaire à une politique du "tout ou rien", i.e. le budget énergétique est soit à une valeur minimum, soit à une valeur calculée en fonction de l'énergie récupérée (afin de limiter le gaspillage d'énergie). Cette approche peut conduire à des variations brutales du budget alloué, généralement non adaptées aux applications WSN. Dans la section suivante (Section 3.2), Fuzzyman [aitaoudia16icc], un EM basée sur la théorie du contrôle flou, est présenté et permet une gestion douce de l'énergie. En plus de l'énergie résiduelle, Fuzzyman utilise l'énergie récupérée pour déterminer le budget énergétique. Cependant, cette valeur est difficile à mesurer en pratique. Par conséquent, la Section 3.3 présente RLman [aitaoudia17icc], un EM basé sur l'apprentissage par renforcement dans des espaces continus d'états et d'actions. Cet EM nécessite seulement l'énergie résiduelle comme entrée et est comparé à LQ-Tracker (en plus de P-FREEN et Fuzzyman), car il s'agit également d'un schéma sans prédiction utilisant seulement l'énergie résiduelle comme entrée.

3.2 Fuzzyman : Utilisation de la logique floue

3.2.1 Conception de Fuzzyman

Architecture de Fuzzyman

L'objectif de Fuzzyman est de calculer un budget énergétique $e_b[k]$ que le nœud doit dépenser durant l'intervalle de temps k en fonction de l'énergie résiduelle à la fin de l'intervalle précédent $e_r[k-1]$ et l'énergie récupérée pendant l'intervalle de temps précédent $e_h[k-1]$. Fuzzyman est donc exécuté à la fin de chaque intervalle. Fuzzyman est composée de trois composants [31] : l'unité de floutage, le moteur d'inférence et l'unité de défloutage comme montré en Figure 3.4. Le floutage consiste à transformer les entrées physiques dans un ensemble flou compatible avec le

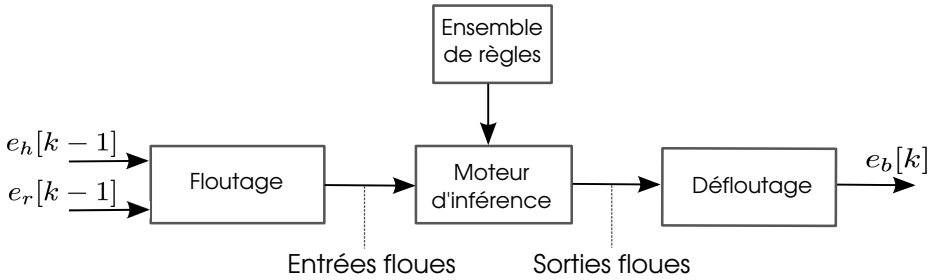


FIGURE 3.4 – Architecture de Fuzzyman.

moteur d'inférences. Un ensemble flou est un intervalle des valeurs des données d'entrée et une fonction d'appartenance associée. Le moteur d'inférence définit des règles d'actions en fonction des entrées floutées. Enfin, le défloutage permet d'associer une valeur physique aux sorties du moteur.

Floutage des entrées du contrôleur

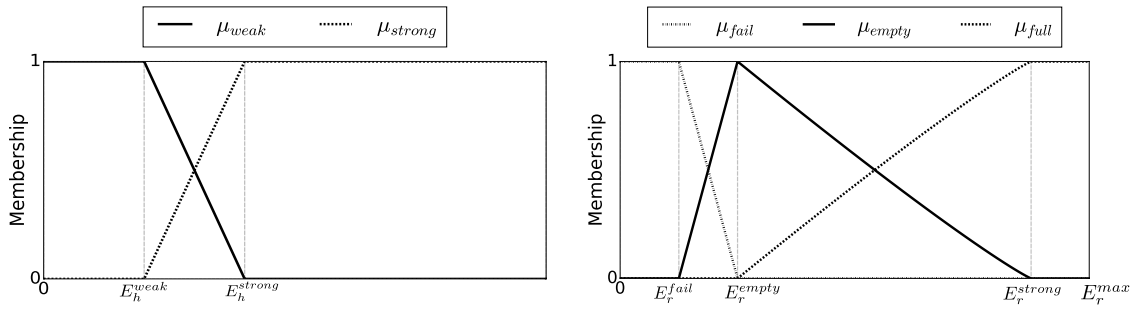
Le premier module de Fuzzyman est l'unité de floutage qui convertit chaque entrée, i.e. $e_r[k-1]$ et $e_h[k-1]$, en ensemble flou.

Floutage de l'énergie récupérée : Pour décrire la valeur physique de l'énergie récupérée $e_h \geq 0$, deux ensembles flous appelés "WEAK" et "STRONG" sont définis. Ils sont associés respectivement aux fonctions d'appartenance μ_{weak} et μ_{strong} décrites dans la Figure 3.5a. Soit E_b^{min} la quantité d'énergie assurant une qualité de service minimale pour un intervalle de temps, E_h^{weak} tient compte des fuites de puissance P_L et du taux de conversion de puissance $\eta \in (0, 1]$ et est donc égale à :

$$E_h^{weak} = \frac{E_b^{min}}{\eta} + P_L T_s. \tag{3.5}$$

Ainsi, si la source est entièrement "WEAK", alors la quantité d'énergie récupérée est insuffisante pour fournir le budget énergétique minimale E_b^{min}. E_h^{strong} est le seuil pour lequel l'énergie récupérée est considérée comme entièrement "STRONG".

Floutage de l'énergie résiduelle : Trois ensemble flous sont définis pour décrire l'énergie résiduelle $e_r \in [0, E_r^{max}]$. Ces ensembles sont appelés "FAIL", "EMPTY" et "FULL" et sont associés respectivement aux fonctions d'appartenance μ_{fail}, μ_{empty} et μ_{full} décrites dans la Figure 3.5b. E_{res} est définie comme la quantité d'énergie qui doit être réservée afin d'assurer la qualité de service minimale lorsqu'aucune énergie n'est récupérée. E_r^{full} indique que le nœud a réservé suffisamment d'énergie E_{res}. E_r^{empty} indique que toute l'énergie récupérée E_{res} a été utilisée. Ainsi, $E_{res} = E_r^{full} - E_r^{empty}$. Choisir $E_r^{empty} > E_r^{fail}$ permet d'éviter les pannes lorsque toute l'énergie réservée a été utilisée. Enfin, choisir $E_r^{full} < E_r^{max}$ permet d'éviter les gaspillages d'énergie lors de la saturation de l'élément de stockage.



(a) Fonctions d'appartenance pour l'énergie récupérée.

(b) Fonctions d'appartenance pour l'énergie résiduelle.

FIGURE 3.5 – Fonctions d'appartenance utilisée par l'unité de floutage.

e_h \ e_r	FAIL	EMPTY	FULL
STRONG	E_b^{eds} (R_1)	(3.6) (R_2)	(3.6) (R_3)
WEAK	E_b^{eds} (R_4)	E_b^{min} (R_5)	$e_b[k-1]$ (R_6)

TABLE 3.2 – Ensemble de règles utilisé par le moteur d'inférence.

Moteur d'inférence

Le rôle du moteur d'inférence est de créer des règles de contrôle en fonctions des entrées floutées par l'unité de floutage. La stratégie du moteur d'inférence est décrite par un ensemble de six règles IF-THEN R_i avec $i \in \{1 \dots 6\}$ données dans la Table 3.2. Pour chaque intervalle k, la sortie de la règle i est notée $e_b^i[k]$. Lorsque le nœud a un risque élevé de panne électrique, il passe dans un *état de détresse énergétique* et la quantité d'énergie nécessaire pour assurer un fonctionnement minimal pour un intervalle de temps dans cet état est noté E_b^{eds} tel que $E_b^{eds} < E_b^{min}$. Les sorties des règles R_2 et R_3 sont données par (3.6), et celle de la règle R_6 est le budget énergétique utilisé pour l'intervalle précédent.

$$e_b^i[k] = E_b^{min} + \mu_{full}\left(e_r[k-1]\right)\left(e_h[k-1] - \frac{E_b^{min}}{\eta} - P_L T_s\right)\eta. \tag{3.6}$$

Dans (3.6), on suppose que les quantités d'énergie récupérée pour deux intervalles consécutifs sont similaires. Cette hypothèse est raisonnable pour une durée d'intervalle suffisamment courte.

Toutes les règles contiennent l'opération logique AND et appliquent l'implication suivante :

$$\text{IF } (e_r[k-1] \text{ is } X_r^i) \text{ AND } (e_h[k-1] \text{ is } X_h^i) \text{ THEN } e_b[k] \text{ is } e_b^i[k], \tag{3.7}$$

où la phrase "x is X" signifie "x appartient à l'ensemble flou X avec une valeur d'appartenance $\mu_X(x)$", X_r^i peut être FAIL, EMPTY ou FULL et X_h^i STRONG ou WEAK. Il est important de noter que quatre règles au plus peuvent être activées simultanément par une exécution de Fuzzyman. La stratégie de gestion d'énergie décrite par la Table 3.2 correspond aux cinq scénarios suivants :

1. R_1 et R_4 : Si l'énergie résiduelle est FAIL, alors le nœud est en détresse énergétique. Dans ce cas, le budget énergétique est fixé à E_b^{eds}.
2. R_5 : Si la quantité d'énergie récupérée est WEAK et que le dispositif de stockage d'énergie est EMPTY, alors le budget énergétique est fixé à la quantité d'énergie requise pour assurer une qualité de service minimale, i.e. E_b^{min}.
3. R_6 : Si la quantité d'énergie récupérée est WEAK et que le dispositif de stockage d'énergie est FULL, le budget énergétique est inchangé.
4. R_2 : Si la quantité d'énergie récupérée est STRONG et que le dispositif de stockage d'énergie n'est pas encore complètement chargé, i.e. $\mu_{full}(e_r[k-1]) < 1$, alors une partie de l'énergie récupérée est utilisée pour alimenter le nœud, tandis que le reste est stocké. La fraction de l'énergie récupérée utilisée pour alimenter le nœud sera au moins E_b^{min}, et dépend de μ_{full} comme indiquée par (3.6).
5. R_3 : Si la quantité d'énergie récupérée est STRONG et que le dispositif de stockage d'énergie est complètement chargé, i.e. $\mu_{full}(e_S[k-1]) = 1$, alors seule la quantité d'énergie compensant les fuites est stockée, tandis que le reste est utilisé pour alimenter le nœud, minimisant ainsi le risque de gaspillage d'énergie.

L'opérateur flou AND est implémenté par la fonction `min` [31]. La *valeur d'activation* $\alpha_{R_i}[k] \geq 0$ pour chaque règle R_i est défini par :

$$\alpha_{R_i}[k] = \texttt{min}\left\{\mu_{X_r^i}\left(e_r[k-1]\right), \mu_{X_h^i}\left(e_h[k-1]\right)\right\}. \tag{3.8}$$

La règle R_i est *activée* si sa valeur d'activation strictement positive. Si la règle n'est pas activée, sa valeur de sortie $e_b^i[k]$ est alors simplement mise à 0. A chaque exécution, au moins une règle est

activée, et donc :

$$\sum_{i=1}^{6} \alpha_{R_i}[k] > 0. \tag{3.9}$$

La valeur d'activation de chaque règle est interprétée comme la valeur d'appartenance du budget énergétique à la sortie de la règle.

Défloutage du budget énergétique

L'unité de défloutage calcule une valeur physique du bilan énergétique à partir des sorties du moteur d'inférence. La formule du "centre de gravité" est la plus courante [31] pour effectuer le défloutage. La valeur physique du budget énergétique est ainsi calculée par :

$$e_b[k] = \frac{\sum_{i=1}^{6} \alpha_{R_i}[k] e_b^i[k]}{\sum_{i=1}^{6} \alpha_{R_i}[k]}. \tag{3.10}$$

Finalement, la complexité de l'algorithme Fuzzyman est $O(1)$. Il requiert peu de calcul et d'espace mémoire. Fuzzyman est donc parfaitement adapté aux nœuds de réseaux de capteurs.

3.2.2 Évaluation de Fuzzyman

Environnement de simulation

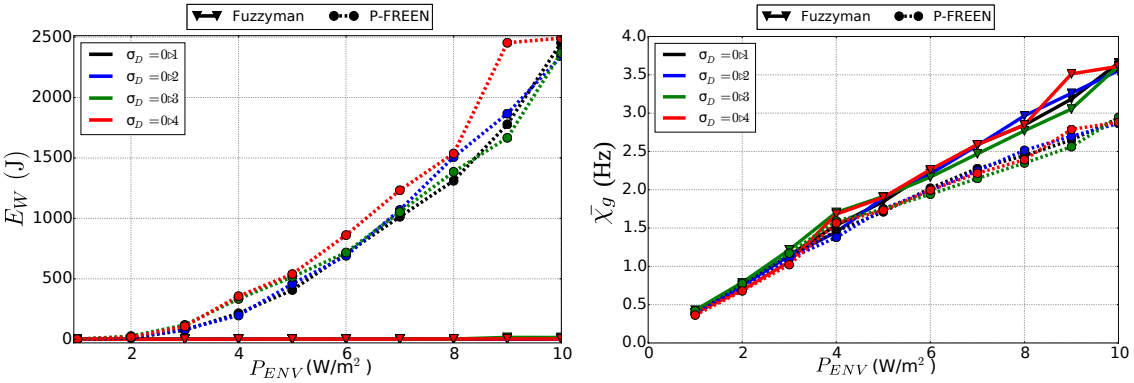
Nous utilisons les caractéristiques de la plate-forme PowWow [32] dont le stockage de l'énergie est une super-capacité de 0.9 F. Les nœuds sont alimentés par une lumière ambiante d'intérieure. Cette source d'énergie est diurne, et son niveau typique d'éclairement varie de 1 W/m^2 [33] à 10 W/m^2 [34]. Les simulations ont été réalisées en utilisant GreenCastalia [35], un environnement open-source de simulation de la récupération d'énergie pour le simulateur de réseaux Castalia/OMNeT++ [36]. Le réseau simulé est composé d'un nœud puits (alimenté par batterie) et de quatre nœuds alimentés par des panneaux solaires. La surface du panneau solaire est fixée à 28 cm^2, et son efficacité du panneau à 15% (valeur réaliste par rapport aux technologies photovoltaïques actuelles [37]). La plate-forme PowWow simulée utilise une radio TI CC1000 qui consomme 22.2 mW en réception, 80.1 mW en transmission et 0.0006 mW en veille. Afin d'évaluer les performances des EM, seul le gaspillage d'énergie dû à la saturation de l'élément de stockage d'énergie est considéré i.e. l'énergie récupérée qui ne peut pas être stockée car l'élément de stockage est plein. L'efficacité de conversion de puissance η vaut 1 et la puissance de fuite P_L vaut 0 W. Enfin, T_s est fixé à 300 s, et chaque simulation dure 31 jours (temps simulé).

La consommation d'énergie du nœud est contrôlée par duty-cycling [38]. A chaque réveil, le nœud effectue une mesure et envoie la valeur ainsi obtenue au puits. Les sources d'énergie fournissant une énergie qui varie avec le temps et l'espace, le cycle de fonctionnement de chaque nœud est individuel et découplé des autres nœuds. Cela rend les protocoles MAC synchrones inadaptés aux WSN à récupération d'énergie car ils nécessitent un cycle de fonctionnement synchronisé [39]. Dans [aitaoudia15pimrc], nous avons montré qu'un simple protocole MAC CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance) conduit à une consommation d'énergie plus faible que le protocole TMAC proposé par Castalia. Par conséquent, ce protocole est utilisé dans ce travail. E_B^{min} est fixé à 0.0221 J ce qui correspond à une fréquence de réveil minimale de 1/60 Hz.

Fuzzyman est comparé à P-FREEN [22], un EM sans prédiction. Trois métriques sont évaluées : le taux de panne D_R, correspondant au temps passé en panne de courant, l'énergie gaspillée E_W, i.e. l'énergie récupérée qui n'a pas pu être stockée parce que l'élément de stockage d'énergie était plein, et le taux moyen de paquets générés $\overline{\chi_g}$.

Traces énergétiques

GreenCastalia a besoin de traces d'énergie pour simuler l'énergie récupérée. Dans [aitaoudia15pimrc], un algorithme de génération de traces diurnes a été proposé. Le but de cette algorithme est d'évaluer les EM en fonction des caractéristiques de la source d'énergie qui influencent de manière significative leurs comportements, et qui peuvent être définies par l'utilisateur en tant



(a) Energie gaspillée E_W.

(b) Taux moyen de paquets générés $\overline{\chi_g}$.

FIGURE 3.6 – Performances de Fyzzyman et P-FREEN en fonction de P_{AVG} pour différentes valeurs de σ_D.

que paramètres d'entrée du modèle proposé. En particulier, le paramètre P_{AVG} permet de définir la puissance moyenne récupérée pendant une journée et σ_D permet de contrôler l'intensité des fluctuations de la puissance récupérée. Pour une source lumineuse d'intérieure avec des densités de puissance typiques allant de 1 W/m^2 à 10 W/m^2, des simulations ont été effectuées pour des valeurs de P_{AVG} allant de 1 W/m^2 à 10 W/m^2, et des valeurs de σ_D comprises entre 0.1 (faibles fluctuations d'un jour à l'autre) et 0.4 (fluctuations importantes).

Résultats de simulations

Fuzzyman et P-FREEN ont tous les deux un ratio d'inactivité inférieur à 0.2 % dans tous les scénarios de simulation. Par conséquent, seul les résultats en terme d'énergie gaspillée et de taux moyen de paquets générés sont donnés dans la Figure 3.6.

La Figure 3.6a montre l'énergie gaspillée E_W quand P_{AVG} et σ_D varient. Si E_W est similaire pour les deux EM lorsque $P_{AVG} < 2$ W/m^2, Fuzzyman gaspille moins d'énergie que P-FREEN pour des valeurs plus élevées de P_{AVG}. La valeur de σ_D n'a pas d'impact sur les performances des deux EM. Afin d'évaluer l'impact de l'efficacité de la gestion d'énergie obtenue par Fuzzyman, on considère le taux moyen de paquets générés $\overline{P_R}$ qui est directement lié au budget énergétique alloué par les EM. La Figure 3.6b montre l'impact de P_{AVG} et de σ_D sur $\overline{P_R}$. Si le budget énergétique moyen est similaire pour les deux EM pour de faibles valeurs de P_{AVG}, Fuzzyman surpasse P-FREEN pour les valeurs élevées de P_{AVG}. De plus, l'avantage de Fuzzyman sur P-FREEN augmente quand P_{AVG} augmente. Ainsi, Fuzzyman permet d'obtenir un gain maximum de 25 % par rapport à P-FREEN (pour $P_{AVG} = 10$ mW et $\sigma_D = 0.4$).

3.2.3 Conclusion

Cette partie présente la conception et l'évaluation de Fuzzyman, un nouvel EM pour les nœuds à récupération d'énergie basé sur la théorie du contrôle flou. Fuzzyman est capable de fournir une efficacité élevée d'utilisation de l'énergie récupérée, tout en évitant les pannes de courant. Ce travail présente l'évaluation de Fuzzyman par rapport à P-FREEN. En utilisant des simulations réseaux alimentés par des traces d'énergie, nous avons montré que Fuzzyman surpasse P-FREEN en termes d'efficacité d'utilisation de l'énergie récupérée.

Cependant, Fuzzyman détermine le budget énergétique en fonction de mesures de l'énergie récupérée. Cette valeur étant difficile à mesurer précisément en pratique, une adaptation de Fuzzyman est proposée dans [aitaoudia18sensors] et utilisé expérimentalement dans la partie 4.4.2 avec la plateforme PowWow. De plus, cette considération a motivé la conception de RLMan, présenté dans la prochaine section, un EM basé sur la théorie de l'apprentissage par renforcement.

3.3 RLman : Utilisation de l'apprentissage par renforcement

3.3.1 Formulation du problème de gestion de l'énergie récupérée

Dans cette étude, l'EM est toujours exécuté périodiquement tous les T_s. La quantité d'énergie résiduelle est notée e_r et l'élément de stockage d'énergie est supposé avoir une capacité finie notée E_r^{max}. Le seuil de panne matérielle, i.e. la quantité minimale d'énergie résiduelle requise pour le fonctionnement du nœud, est notée E_r^{fail}. Le rôle du nœud est d'envoyer périodiquement des paquets avec un débit $\chi_g \in [X^{min}, X^{max}]$, et le but de l'EM est d'ajuster dynamiquement la performance du nœud en choisissant χ_g. L'objectif est de maximiser le débit χ_g tout en maintenant le nœud en vie, i.e. en évitant des pannes de courant.

En RL, il est supposé que tous les objectifs peuvent être décrits par la maximisation de la récompense cumulative moyenne, où une récompense R est une scalaire décrivant comment le nœud fonctionne. Le problème est formulé par un processus de décision markovien (MDP - Markov Decision Process) $\langle \mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R}, \gamma \rangle$ [40], avec :
L'ensemble des états $\mathcal{S}$: L'état du nœud à l'instant k est défini par l'énergie résiduelle courante $e_r[k]$. Ainsi, $\mathcal{S} = [E_r^{min}, E_r^{max}]$.
L'ensemble des actions $\mathcal{A}$ A chaque intervalle de temps, une action, notée $A[k]$, est réalisée, correspondant au choix du débit χ_g. Ainsi, $\mathcal{A} = [X^{min}, X^{max}]$.
La fonction de transition $\mathcal{T}$ La fonction de transition donne la probabilité d'une transition vers $e_r[k+1]$ lorsque l'action χ_g est effectuée dans l'état $e_r[k]$. L'espace des états étant continu, la fonction de transition est une densité de probabilité telle que :

$$\Pr\left(e_r[k+1] \in \mathcal{S}[k+1] \mid S[k] = e_r, A[k] = \chi_g\right) = \int_{\mathcal{S}[k+1]} \mathcal{T}\left(S[k] = e_r, A[k] = \chi_g, S\right) dS. \quad (3.11)$$

La fonction de récompense $\mathcal{R}$ Dans cette étude, la récompense est calculée en fonction de χ_g et e_r :

$$R = \phi \chi_g, \quad (3.12)$$

avec ϕ l'énergie résiduelle normalisée :

$$\phi = \frac{e_r - E_r^{fail}}{E_r^{max} - E_r^{fail}}. \quad (3.13)$$

Ainsi, maximiser la récompense revient à maximiser à la fois le débit et l'état de charge de l'élément de stockage d'énergie. Cependant, l'énergie résiduelle dépend de l'énergie consommée et de l'énergie récupérée, et ces variables sont stochastiques. La fonction de récompense est définie par :

$$\mathcal{R}(e_r, \chi_g) = \mathbb{E}\left[R[k+1] \mid S[k] = e_r,\ A[k] = \chi_g\right]. \quad (3.14)$$

Le facteur de dépréciation γ Le facteur de dépréciation prend des valeur entre 0 et 1. Sa fonctionnalité sera détaillée ensuite.

La fonction de transition $\mathcal{T}$ décrit la dynamique de l'environnement et est supposée inconnue. De la même façon, en calculant la récompense avec (3.12), sa valeur a un comportement stochastique car elle dépend de l'énergie résiduelle. Le comportement d'un nœud est défini par une *politique* π, qui est une distribution conditionnelle d'actions :

$$\pi(\chi_g | e_r) = \Pr\left[A[k] = \chi_g \mid S[k] = e_r\right]. \quad (3.15)$$

L'objectif est d'apprendre une politique qui maximise une fonction objectif. La fonction objectif J est définie comme la moyenne de la somme des récompenses pondérées par rapport à une politique π et une distribution de l'état initial ρ_0 :

$$J(\pi) = \mathbb{E}\left[\sum_{k=1}^{\infty} \gamma^{k-1} R[k] \,\middle|\, \rho_0,\ \pi\right] = \int_{\mathcal{S}} \rho_\pi(e_r) \int_{\mathcal{A}} \pi(\chi_g | e_r) \mathcal{R}(e_r, \chi_g) d\chi_g de_r, \quad (3.16)$$

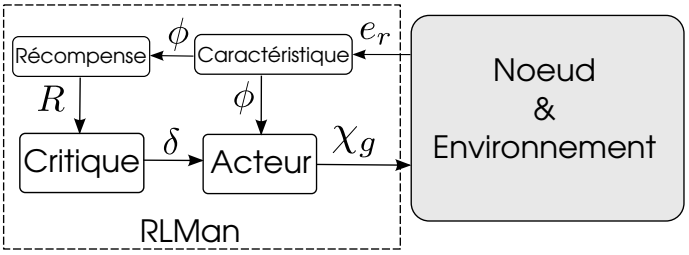


FIGURE 3.7 – Architecture globale de RLMan.

avec :

$$\rho_\pi(e_r) = \int_{\mathcal{S}} \rho_0(e'_r) \sum_{k=1}^{\infty} \gamma^{k-1} \Pr\left[S[k] = e_r \mid S_0 = e'_r,\ \pi\right] de'_r, \tag{3.17}$$

la distribution de l'état e_r sous une politique π. Dans (3.16), on peut voir que choisir une valeur de γ proche de 0 conduit à une évaluation "myope" car les récompenses immédiates sont privilégiées, alors que choisir une valeur de γ proche de 1 mène à une évaluation "à long terme".

L'utilisation de politiques stochastiques permet d'explorer l'environnement, ce qui est fondamental en RL. En effet, le RL est similaire à un apprentissage de type "essai et erreur", et le but de l'algorithme est de découvrir une bonne politique à partir de son expérience avec l'environnement, tout en minimisant la quantité de récompense "perdue" pendant l'apprentissage, ce qui conduit à un dilemme entre l'exploration (apprendre d'avantage sur l'environnement) et l'exploitation (maximiser la récompense en exploitant des informations connues).

3.3.2 Algorithme de RLMan

Une politique π est évaluée en estimant deux fonctions de valeur associées à cette politique. La fonction de valeur de l'état, notée v_π, prédit la future récompense si la politique π est utilisée à partir d'un état donné. Elle évalue ainsi la "qualité" des états [40] :

$$v_\pi(e_r) = \mathbb{E}\left[\sum_{i=1}^{\infty} \gamma^{i-1} R_{k+i} \;\middle|\; S[k] = e_r,\ \pi\right]. \tag{3.18}$$

De la même façon, la fonction de valeur action-état associée à la réalisation de l'action $A[k]$ dans l'état $S[k]$ est notée Q_π et évalue "qualité" du couple état-action quand la politique π est utilisée [40] :

$$Q_\pi(e_r, \chi_g) = \mathbb{E}\left[\sum_{i=1}^{\infty} \gamma^{i-1} R_{k+i} \;\middle|\; S[k] = e_r,\ A[k] = \chi_g,\ \pi\right]. \tag{3.19}$$

L'EM proposé est un algorithme de type acteur-critique, une famille de techniques RL connue pour sa capacité à rechercher des politiques optimales à l'aide d'estimations de gradient à faible variance [41]. La Figure 3.7 montre la relation entre l'acteur et le critique. L'acteur met à jour la politique π_ψ de paramètre ψ par une montée de gradient de la fonction objectif J (3.16). Un résultat fondamental pour calculer le gradient de J est donné par le théorème du gradient de la politique [42] :

$$\nabla_\psi J(\pi_\psi) = \int_{\mathcal{S}} \rho_{\pi_\psi}(e_r) \int_{\mathcal{A}} Q_{\pi_\psi}(e_r, \chi_g) \nabla_\psi \pi_\psi(\chi_g \mid e_r) d\chi_g de_r = \mathbb{E}\left[Q_{\pi_\psi}(e_r, \chi_g)\, \nabla_\psi \log \pi_\psi(\chi_g \mid e_r) \;\middle|\; \rho_{\pi_\psi},\ \pi_\psi\right]. \tag{3.20}$$

Ce résultat réduit le calcul à une moyenne et permet ainsi le développement d'algorithmes simples reposant sur l'échantillonnage de cette moyenne. Dans ce travail, une politique Gaussienne est utilisée pour générer f :

$$\pi_\psi(\chi_g \mid e_r) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\chi_g - \mu)^2}{2\sigma^2}\right). \tag{3.21}$$

σ est un paramètre constant qui contrôle l'exploration, et μ est linéaire avec la caractéristique ϕ :

$$\mu = \psi\phi. \tag{3.22}$$

Définir μ comme une fonction linéaire de ϕ permet un coût mémoire réduit avec seulement un flottant ψ stocké. De plus, le surcoût en calcul est aussi réduit car $\bigtriangledown_\psi \mu = \phi$ conduit à :

$$\bigtriangledown_\psi \log \pi_\psi(\chi_g \mid e_r) = \frac{(\chi_g - \mu)}{\sigma^2}\phi. \tag{3.23}$$

D'autres manières de calculer μ peuvent être utilisées, notamment par des réseaux de neurones artificiels pour lesquels ψ serait un vecteur des poids des neurones. Ces approches induiraient un coût en mémoire et en calcul important.

En introduisant l'erreur de différence temporelle (TD - Temporal Difference) définie par :

$$\delta = R[k+1] + \gamma v_{\pi_\psi}(e_r[k+1]) - v_{\pi_\psi}(e_r[k]), \tag{3.24}$$

le gradient de la politique peut être calculé par [41] :

$$\bigtriangledown_\psi J(\pi_\psi) = \mathbb{E}\left[\delta \bigtriangledown_\psi \log \pi_\psi(\chi_g \mid e_r) \;\middle|\; \rho_{\pi_\psi},\ \pi_\psi\right]. \tag{3.25}$$

L'erreur TD peut être intuitivement interprétée comme une critique de l'action précédente. Une erreur positive suggère que cette action doit être choisie plus souvent, alors qu'une erreur négative suggère le contraire. Le critique calcule l'erreur TD (3.24) et doit donc connaitre la fonction de valeur v_{π_ψ}. Comme l'espace d'état est continu, il est impossible de stocker la fonction de valeur pour chaque état. Une approximation de fonction linéaire est donc également utilisée pour estimer la fonction de valeur :

$$\hat{v}_\theta(e_r) = \theta\phi, \tag{3.26}$$

avec ϕ la caractéristique (3.13) et θ un paramètre de l'approximation. Le critique estime la fonction de valeur en mettant à jour le paramètre θ, il est implémenté par l'algorithme TD(λ) [43] défini par :

$$\nu[k] = \gamma\lambda\nu[k-1] + \phi[k] \tag{3.27}$$

$$\theta[k] = \theta[k-1] + \alpha\delta[k]\nu[k] \tag{3.28}$$

avec $\alpha \in [0,1]$ un paramètre de convergence des itérations, $\nu[k]$ un scalaire utilisé pour attribuer un crédit aux états visités plusieurs itérations auparavant et $\lambda \in [0,1)$ un facteur du taux de pondération.

L'Algorithme 1 montre l'EM proposé. A chaque exécution, l'algorithme utilise l'énergie résiduelle actuelle $e_r[k]$ et la récompense $R[k]$ calculée à l'aide de (3.12). La caractéristique $\phi[k]$ et l'erreur TD sont dans un premier temps calculés (lignes 1 et 2), puis la critique est mise à jour à l'aide de l'algorithme TD(λ) (lignes 4 et 5). Ensuite, l'acteur est mis à jour en utilisant le théorème du gradient de la politique à la ligne 7, où $\beta \in [0,1]$ est un paramètre de taille de pas. La moyenne de la politique gaussienne est liée à la plage des valeurs autorisées à la ligne 8. Enfin, une fréquence est générée par la politique gaussienne de la ligne 10, qui sera utilisée dans l'intervalle de temps actuel. La distribution gaussienne n'étant pas bornée, il est nécessaire de fixer la valeur générée à la plage autorisée (ligne 11).

L'algorithme développé a un faible encombrement en mémoire et une complexité réduite, il est donc adapté à une implémentation sur des nœuds WSN.

3.3.3 Évaluation de RLMan

RLman a été évalué et comparé à trois EM de l'état de l'art en utilisant des simulations. La plateforme simulée est PowWow [44]. La plate-forme PowWow utilise une super-capacité comme élément de stockage d'énergie, avec une tension maximale de 5.2 V et une tension de panne de 2.8 V. L'énergie récupérée a été simulée par deux traces d'énergie issues de mesures réelles : une de 270 jours correspondant à de la lumière intérieure [45] et l'autre de 180 jours correspondant à du

Algorithm 1 Gestionnaire d'énergie basé sur l'apprentissage par renforcement.

Input: $e_r[k]$, $R[k]$

1: $\phi[k] = \frac{e_r[k]-E_r^{fail}}{E_r^{max}-E_r^{fail}}$ ▷ Caractéristique (3.13)

2: $\delta[k] = R[k] + \gamma\theta[k-1]\phi[k] - \theta[k-1]\phi[k-1]$ ▷ Erreur TD (3.24), (3.26)

3: ▷ Critique : mise à jour de l'estimation de la fonction de valeur (3.27), (3.28) :

4: $\nu[k] = \gamma\lambda\nu[k-1] + \phi[k]$

5: $\theta[k] = \theta[k-1] + \alpha\delta[k]\nu[k]$

6: ▷ Acteur : mise à jour de la politique (3.22), (3.23), (3.25) :

7: $\psi[k] = \psi[k-1] + \beta\delta[k]\frac{(f[k-1]-\psi[k-1]\phi[k-1])}{\sigma^2}\phi[k-1]$

8: Fixer μ_t dans $[X^{min}, X^{max}]$

9: ▷ Génère une nouvelle action :

10: $\chi_g[k] \sim \mathcal{N}(\psi[k]\phi[k], \sigma)$

11: Fixe $\chi_g[k]$ dans $[X^{min}, X^{max}]$

12: **return** $\chi_g[k]$

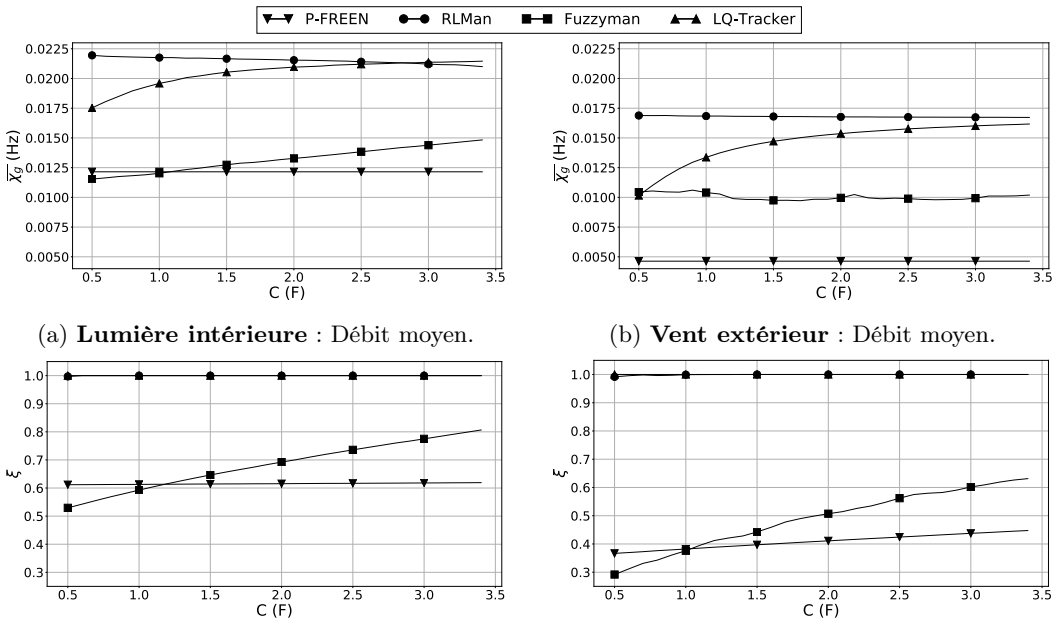
Paramètres communs		RLMan		P-FREEN [22]		Fuzzyman [aitaoudia16icc]		LQ-Tracker [18]	
E_c^{typ}	8.672 mJ	α	0.1	B_{OFL}	$0.95E_r^{max}$	η	1.0	μ	0.001
E_c^{max}	36.0 mJ	β	0.01	η	1.0	E_b^{eds}	$X^{min}T_sE_c^{typ}$	B^*	$0.70E_r^{max}$
X^{min}	$\frac{1}{300}$ Hz	σ	0.1			E_b^{min}	$X^{min}T_sE_c^{typ}$	α	0.5
X^{max}	5 Hz	γ	0.9			E_h^{strong}	$X^{max}T_sE_c^{typ}$	β	1.0
T_s	60 s	λ	0.9			E_h^{weak}	$X^{min}T_sE_c^{typ}$		

Table 3.3 – Valeurs des paramètres de simulations. Le détail des paramètres de P-FREEN, Fuzzyman et LQ-Tracker est donné dans les publications respectives.

vent extérieur [46], permettant l'évaluation des gestionnaires pour deux types de sources d'énergie différents. Le rôle du nœud consiste à mesurer des données, puis à les envoyer à un récepteur. Cependant, en pratique, la quantité d'énergie consommée par une exécution de ces tâches varie, e.g. en raison de multiples retransmissions. La quantité d'énergie requise pour exécuter une exécution a donc été simulée par une variable aléatoire qui suit une distribution beta, avec le mode de la distribution égal à l'énergie consommée si une seule tentative de transmission est requise, noté E_c^{typ}, et le maximum fixé à l'énergie consommée si cinq tentatives de transmission sont nécessaires, noté E_c^{max}. La Table 3.3 donne les valeurs de E_c^{typ} et E_c^{max}, ainsi que les valeurs des paramètres utilisés lors de l'implémentation des EM.

RLMan a été comparé à P-FREEN, Fuzzyman, et LQ-Tracker, trois EM qui visent à maximiser le débit. P-FREEN et Fuzzyman utilisent le niveau d'énergie récupérée en plus de l'énergie résiduelle, et ont donc été exécutés avec une connaissance parfaite de cette valeur. RLMan et LQ-Tracker utilisent seulement la valeur de l'énergie résiduelle. Les EM ont été évalués pour différentes valeurs de la capacité, allant de 0.5 F à 3.5 F, elle a en effet un impact important sur le comportement des EM, mais aussi sur le coût et la taille des nœuds. Le débit moyen $\overline{\chi_g}$ et l'efficacité énergétique ζ définie par (3.4) ont été évalués.

Tous les EM réussissent à éviter les pannes de courant lorsqu'ils sont alimentés par une lumière intérieure ou un vent extérieur. La Figure 3.8 donne les résultats de simulation. D'après les Figures 3.8c et 3.8d, RLMan et LQ-Tracker atteignent plus de 99.9 % d'efficacité, pour la lumière et le vent et pour toutes les valeurs de capacité, et en dépit du fait qu'ils n'utilisent que l'énergie résiduelle comme entrée. De plus, la Figure 3.8b montre que, lorsque le nœud est alimenté par le vent, RLMan surpasse les autres EM en termes de débit moyen pour toutes les valeurs de capacité. Lorsque le nœud est alimenté par une lumière intérieure, RLMan surpasse également les autres EM, sauf LQ-Tracker pour des valeurs de capacité supérieures à 2.8 F. En particulier, pour une capacité de 0.5 F, le débit moyen de RLman est supérieur de plus de 20 % par rapport à LQ-Tracker dans le cas de la lumière intérieure, et de près de 70 % pour le vent extérieur. Ce résultat est intéressant,



(a) **Lumière intérieure** : Débit moyen.

(b) **Vent extérieur** : Débit moyen.

(c) **Lumière intérieure** : Eff. énergétique.

(d) **Vent extérieur** : Efficacité énergétique.

FIGURE 3.8 – Débit moyen et efficacité énergétique en fonction de la valeur de la capacité pour de la lumière intérieure et du vent extérieur.

car l'utilisation d'une petite capacité permet de réduire les coûts et la taille des nœuds.

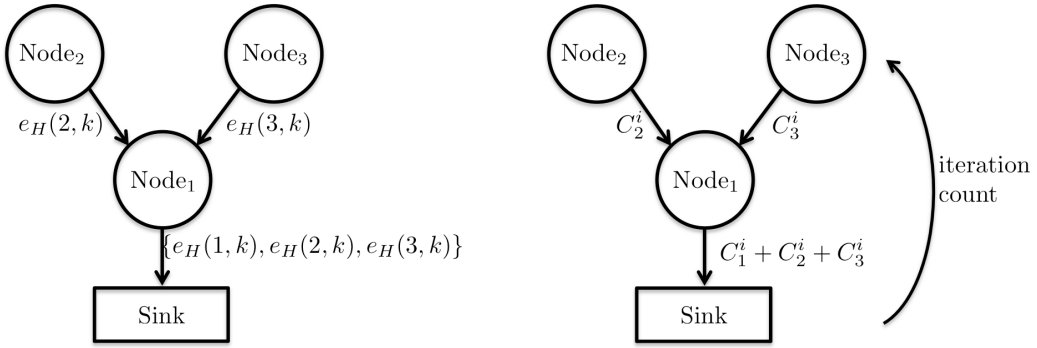
3.3.4 Conclusion

Dans cette partie, le problème de la maximisation de la qualité de service dans un nœud à récupération d'énergie est formulé à l'aide de la théorie de l'apprentissage par renforcement, et un nouveau système de gestion de l'énergie, RLMan, est présenté. RLMan nécessite uniquement l'état de charge de l'élément de stockage d'énergie en entrée, et utilise l'approximation de fonction pour minimiser l'empreinte mémoire et le temps de calcul, ce qui le rend pratique à mettre en œuvre et adapté aux nœuds de capteurs sans fil. Des simulations ont montré les avantages de RLMan en termes de débit et d'efficacité énergétique par rapport à trois gestionnaires d'énergie de l'état de l'art, dans le cas de la récupération d'énergie lumineuse intérieure et d'énergie éolienne extérieure. L'avantage de RLMan est d'autant plus important que l'élément de stockage d'énergie est petit. De plus, ces performances sont obtenues sans avoir besoin de modifier l'algorithme en aucune façon. Nous avons identifié des domaines d'amélioration potentielle, comme l'exploration d'autres récompenses et d'autres approximateurs de fonctions, qui pourraient entraîner une convergence plus rapide.

3.4 Autres contributions

3.4.1 Gestion d'énergie distribuée pour les réseaux multi-sauts

Le problème du calcul distribué des taux de génération de paquets des nœuds formant un réseau multi-saut et organisés dans une topologie en arbre a été formulé sous la forme d'un problème d'optimisation convexe. Dans le cas de réseaux multi-sauts, chaque nœud effectue des mesures pour produire des données et relaie des paquets de données provenant d'autres nœuds. Dans [aitaoudia17wcl], nous avons proposé un algorithme distribué pour le calcul des taux de générations de paquets équitables. Le problème du calcul des taux de génération est formulé comme un problème d'optimisation convexe, et en utilisant l'algorithme des directions alternées rapide (Fast ADMM - Alternating Direction Method of Multipliers), le problème original est décomposé en sous-problèmes plus petits pouvant être résolus en parallèle. Le Figure 3.9 montre les deux approches : centralisée,



(a) Approche centralisée : le puits résout le problème d'optimization convexe.

(b) Approche distribuée : chaque nœud calcule son résidu et transmet le cumul des résidus.

FIGURE 3.9 – Résolution centralisée vs distribuée. du calcul équitable des taux de génération de paquets.

où toutes les données sont transmises au puits pour une résolution centralisée du problème, et distribuée, où seuls des résidus sont transmis mais convergent itérativement. Des simulations utilisant des traces réelles d'énergie lumineuse montrent que l'algorithme atteint des précisions importantes pour un nombre d'itérations relativement faible.

3.4.2 Architecture de récupération d'énergie multi-source

La plupart des solutions existantes de récupération d'énergie reposent sur une seule source d'énergie, ce qui réduit potentiellement la fiabilité du capteur. Dans [gleonec16ogc], nous présentons un circuit qui commute entre plusieurs sources d'énergie hétérogènes et utilise un seul bloc de conditionnement de puissance (PIMC - Power Management Integrated Component). La Figure 3.10 présente l'architecture de ce circuit. Un prototype a été développé et validé avec une plateforme de capteurs sans fil existante, développée par la l'entreprise Wi6labs, utilisant le protocole de radio longue portée LoRa. Les mesures montrent que la commutation entre les sources d'énergie peut combiner efficacement deux sources d'énergie afin d'augmenter l'autonomie du dispositif et/ou la qualité du service.

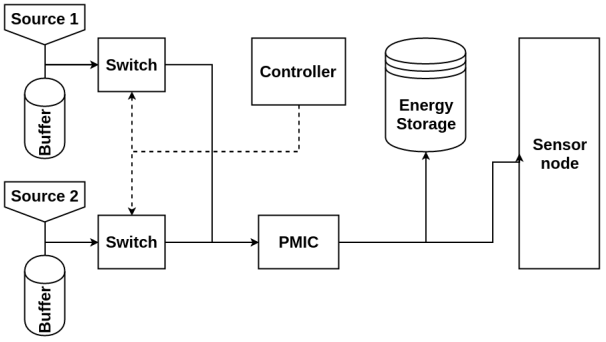


FIGURE 3.10 – Architecture de commutation de sources d'énergie.

3.4.3 Application de la gestion d'énergie à des nœuds LoRa

Les algorithmes de gestion d'énergie ont été principalement été implémentés et évalués par simulation. Nous avons évalué la mise en œuvre de ces algorithmes dans des conditions réelles de déploiements de nœuds de capteurs pour IoT. Ainsi, nous avons évalué les performances des différentes méthodes d'estimation du budget énergétique sur la plateforme LoRa. Présenté dans [gleonec18ict], les résultats montrent que dans ce cas d'utilisation, le choix de l'algorithme affecte la qualité de service du système de moins de 15%, permettant ainsi l'utilisation de méthodes à faible complexité d'implémentation.

La Figure 3.11 montre l'évolution du budget énergétique e_b en fonction de l'état de charge e_r. Ce graphique donne un aperçu rapide du comportement des différents algorithmes pour des états de charge faible et élevé, indiquant quel algorithme est plus ou moins conservateur.

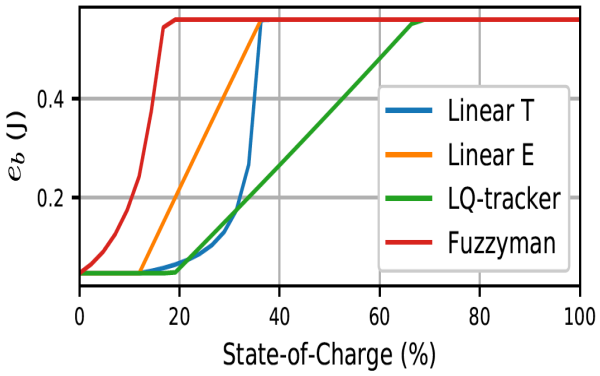


FIGURE 3.11 – Budget énergétique en fonction de l'état de charge.

3.5 Conclusion

L'énergie est le principal goulot d'étranglement au déploiement à long terme d'applications de réseaux de capteurs sans fil, car les nœuds sont typiquement alimentés par batteries qui ne peuvent stocker qu'une quantité finie d'énergie. Une solution prometteuse est de permettre à chaque nœud de récupérer de l'énergie directement de son environnement. Les sources d'énergie étant dynamiques et incontrôlées, il est nécessaire d'adapter continuellement les performances des nœuds afin d'optimiser la qualité de service des applications tout en évitant les pannes de courant.

Dans ce chapitre, deux nouveaux gestionnaires d'énergie sont proposés. Le premier, Fuzzyman, utilise la théorie du contrôle flou pour calculer un budget énergétique en fonction de l'énergie récupérée et de l'énergie résiduelle. Cet algorithme nécessite cependant un réglage fin de ses paramètres en fonction de la source et du matériel. De plus, il utilise l'énergie résiduelle qui est une quantité difficile à mesurer en pratique. Pour s'abstraire de ces contraintes, nous avons proposé un deuxième algorithme, RLman, reposant sur la théorie de l'apprentissage par renforcement. Enfin, nous avons appliqué la gestion d'énergie au cas pratique du standard LoRa, montrant ainsi la faisabilité de l'alimentation par récupération d'énergie pour développer des applications autonomes en énergie.

Les perspectives directes de ces travaux consistent en l'amélioration des algorithmes proposés. Pour adapter Fuzzyman aux variations de l'environnement, un contrôleur adaptatif doit permettre de modifier dynamiquement les paramètres du contrôleur flou (e.g. ceux de l'ensemble de règles ou ceux des fonctions d'appartenance) en fonction des entrées et des sorties du système. Une amélioration de RLman passe par l'ajout d'autres entrées au système permettant d'augmenter l'espace d'exploration en y associant les fonctions d'approximations adéquates. L'utilisation d'autres récompenses peut aussi être envisagée. Enfin, l'extension de ces travaux à des gestionnaires d'énergie distribués dans le cas de réseaux ad-hoc doit permettre une gestion globale de l'énergie et une équité entre les nœuds.

Un problème rencontré lors de l'évaluation d'algorithmes de gestion d'énergie utilisant des simulations est le manque de traces d'énergie récupérée. Une perspective importante serait donc de fournir à la communauté un ensemble de traces d'énergie pour différentes sources et différents environnements. La valorisation d'un générateur de traces permettrait également d'atteindre cet objectif.

Chapitre 4

Les wake-up radios

Une approche prometteuse pour pallier au problème de la consommation d'énergie dans les réseaux de capteurs sans fil est l'utilisation des wake-up radios [15]. Cette approche permet une communication parfaitement asynchrone, minimisant ainsi l'écoute du canal qui est une des principales sources de gaspillage d'énergie. Dans ce chapitre, nous évaluons analytiquement et expérimentalement les avantages de l'utilisation de la wake-up radio dans les réseaux de capteurs bas débit. De plus, pour exploiter pleinement cette technologie, de nouveaux protocoles de communication doivent être conçus et optimisés [aitaoudia16js].

Après une introduction du principe des wake-up radios, une approche analytique est présentée dans la Section 4.2. Cette évaluation repose sur une modélisation générique des protocoles MAC avec des chaînes de Markov absorbantes. Nos travaux reposent sur une approche hybride entre des modèles analytiques et des mesures précises de consommation énergétique qui sont intégrées dans les modèles. Les Sections 4.3 et 4.4 présentent deux protocoles MAC originaux utilisant la wake-up radio. Le premier, OPWUM, permet la sélection opportuniste du prochain relai dans un réseau multi-saut. Ce protocole a été implémenté et évalué par des simulations réseaux dans l'environnement Green-Castalia/OMNet++. Le second protocole, SNW-MAC, permet la collecte de données dans un réseau en étoile. Ce protocole a été implémenté sur une plateforme matérielle et testé dans des conditions réelles. Finalement, la dernière section présente nos autres contributions liées à la wake-up radio avec, entre autres, son association non-triviale mais complémentaire avec la radio longue portée LoRa.

Sommaire

4.1 Principe des wake-up radios 52
4.1.1 Architecture des wake-up radios 52
4.1.2 Caractéristiques des wake-up radios 53
4.2 Evaluation analytique des wake-up radios 54
4.2.1 Modélisation des protocoles 54
4.2.2 Performances analytiques et comparaisons expérimentales 56
4.2.3 Conclusion 56
4.3 OPWUM : protocole MAC opportuniste 57
4.3.1 Protocole MAC opportuniste utilisant la wake-up-radio 57
4.3.2 Évaluation de OPWUM 58
4.3.3 Application à la récupération d'énergie 59
4.3.4 Conclusion 60
4.4 SNW-MAC : collecte des données dans un réseau en étoile 60
4.4.1 Le protocole Star Network WuRx - MAC 60
4.4.2 Protocole expérimental : combinaison à la récupération d'énergie 61
4.4.3 Résultats expérimentaux 64
4.4.4 Conclusion 65
4.5 Autres contributions 65
4.5.1 Communications hétérogènes courtes et longues portées associant LoRa et WuRx 65
4.6 Conclusion 66

4.1 Principe des wake-up radios

4.1.1 Architecture des wake-up radios

Chaque nœud est équipé d'une radio secondaire très faible consommation permettant l'écoute en continu du canal et ainsi la mise en veille des autres éléments du nœud. Les Wake-up Receivers (WuRx) sont ensuite en mesure de réveiller le système complet en envoyant des interruptions quand elles reçoivent une balise de réveil, appelée Wake-up Beacon (WuB). Cette approche permet une communication parfaitement asynchrone, minimisant ainsi l'écoute du canal qui est une des principales sources de gaspillage d'énergie. De plus, les nouvelles WuRx intègrent des capacités de calcul permettant de traiter les données contenues dans le WuB. Ces capacités de calcul permettent notamment un décodage d'adresses afin que seul un destinataire soit réveillé par le WuB.

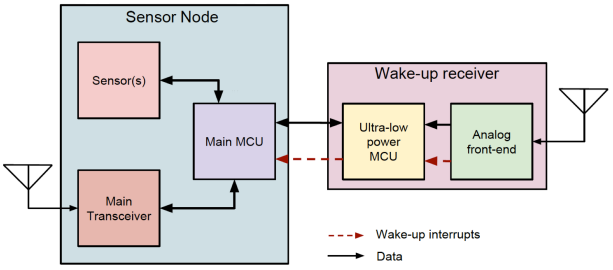


FIGURE 4.1 – Architecture d'un nœud équipé d'une wake-up radio.

La Figure 4.1 montre une architecture classique de la wake-up radio (celle utilisée dans nos travaux) qui est composée de deux blocs principaux, un front-end analogique et un micro-contrôleur (MCU) faible consommation (ULP - Ultra-Low Power). Seul le front-end radio est toujours éveillé, tandis que le micro-contrôleur est réveillé par le front-end uniquement lorsqu'une balise de réveil (WuB) est détectée.

La Figure 4.2 illustre les différentes étapes de réception d'un WuB pour le protocole TI-WuR (Transmitter Initiated WuRx protocol), initié par l'émetteur. L'émetteur envoie dans un premier temps un WuB contenant l'adresse du destinataire. Le front-end analogique de toutes les WuRx des nœuds voisins détecte l'activité sur le canal, et déclenche une interruption pour réveiller le MCU faible consommation embarqué dans la WuRx. Le MCU lit ensuite les données entrantes et effectue un décodage d'adresse. Si l'adresse contenue dans le WuB correspond à l'adresse du nœud, le micro-contrôleur de la WuRx déclenche une interruption pour réveiller le micro-contrôleur principal qui allume ensuite l'émetteur-récepteur principal pour recevoir la trame de données de l'émetteur. La trame de données est ensuite envoyée en utilisant des structures de trame de la norme IEEE 802.15.4.

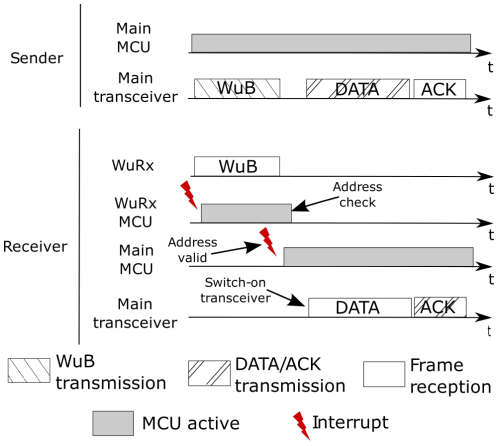


FIGURE 4.2 – Transmission d'un packet avec TI-WuR.

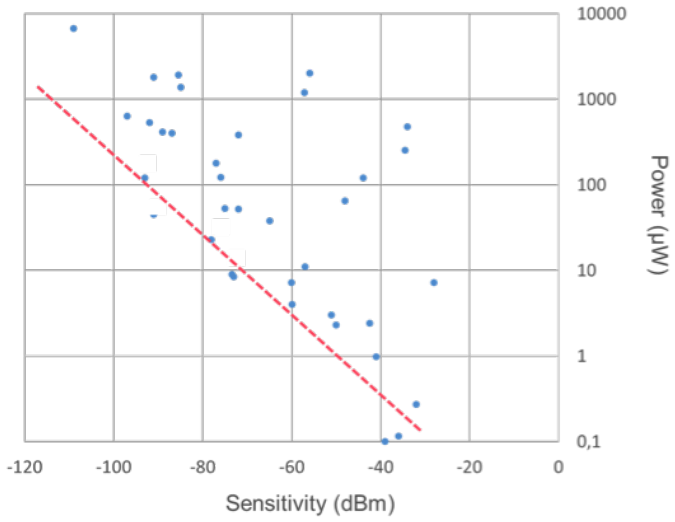


FIGURE 4.3 – Etat de l'art des WuRx.

4.1.2 Caractéristiques des wake-up radios

De nombreuses études ont répertoriées les avancées technologiques associées aux WuRx [47–50], aussi bien au niveau des architectures matérielles qu'au niveau des protocoles de communication. La Figure 4.3 montre la variété des architectures de WuRx proposées dans la littérature. Cette figure est issue des premiers résultats du projet ANR Wake-up, non publiée actuellement. Elle compare différentes solutions au travers de critères important des WuRx : la sensibilité et la consommation. Cependant, il est difficile de s'abstraire du débit (et donc de la bande passante et de la latence) de la WuRx pour comparer différentes architectures. Les résultats de cette figure sont donc à interpréter avec précautions. Cependant, un compromis existe clairement entre la sensibilité et la consommation avec des sensibilités allant de -40 dBm à -100 dBm pour des puissances consommées de l'ordre du μW au mW.

Dans nos travaux, la WuRx de [51] a été utilisée pour des expérimentations et des simulations. Elle reçoit des données modulées en OOK (On-Off Keying) à 1 kbps dans la bande 868 MHz. Dans ces conditions, sa sensibilité est mesurée à -55 dBm ; et est donc nettement inférieure à celle d'un émetteur-récepteur WSN, qui peut atteindre environ -115 dBm (e.g. celle d'un CC1120 de Texas Instruments). Par conséquent, les WuB sont transmis à une puissance d'émission plus élevée que les autres trames. De plus, les débits binaires des WuRx sont typiquement de quelques kbps (1 kbps pour celle utilisée dans ce travail), conduisant à des latences plus élevées, notamment lorsqu'une adresse est transmise dans le WuB. Les WuB sont donc transmis à une puissance de transmission plus élevée et pour une durée plus longue, ce qui rend leur envoi énergétiquement coûteux. En conséquence, les avantages potentiels apportés par la WuRx doivent être étudiés, et les protocoles de communication exploitant ces dispositifs émergents doivent être soigneusement conçus. Le micro-contrôleur ULP intégré dans la wake-up radio est un PIC12LF1552 de Microchip. La consommation mesurée est de 1.83 μW en mode d'écoute continue et de 284 μW quand le micro-contrôleur ULP est éveillé pour traiter les données reçues. La consommation d'énergie de la WuRx devient donc significative lorsque le PIC est actif. A chaque réveil, le PIC est actif pendant 19 ms pour effectuer une détection d'adresse. L'énergie consommée par la WuRx à chaque réveil du PIC est donc de 5.40 μJ. En considérant un nœud typique, n'utilisant pas une WuRx, mais utilisant l'approche duty-cycle avec un duty-cycle d'une valeur typique de 0.05% et consommant 100 mW lorsque l'émetteur-récepteur est actif, alors l'énergie totale consommée par ce nœud sur une période de 24 h est de 4.32 J. Cette quantité d'énergie correspond à plus de 8×10^5 réveils du PIC. Le nombre de faux réveil du PIC, i.e. les réveils non causés par des WuB mais par le bruit du canal sans fil, a été mesuré sur une période de 24 h dans un environnement intérieur. La Figure 4.4 donne le nombre total cumulé de faux réveils en fonction du temps. On peut observer que la plupart des faux réveils se produisent pendant la journée. Au total, 3110 faux réveils ont été comptabilisés sur une période 24 h, ce qui est deux ordres de grandeur en dessous du scénario précédemment

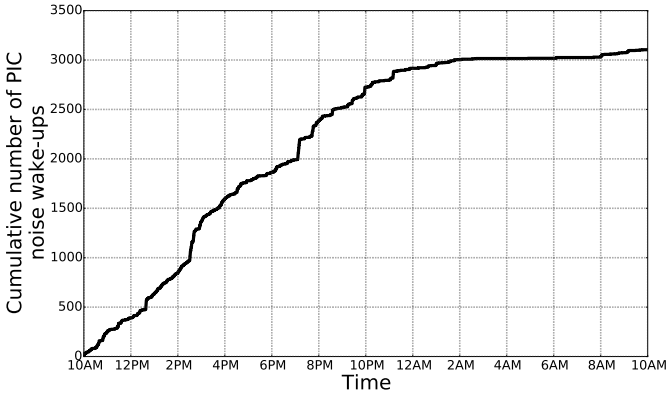


FIGURE 4.4 – Nombre de faux réveils du PIC pendant une durée de 24 h dans un environnement intérieur.

considéré avec un duty cycle de 0.05%. De plus, aucun faux réveil du MCU principal n'est arrivé. Ces résultats montrent l'importance du microcontrôleur embarqué dans la WuRx.

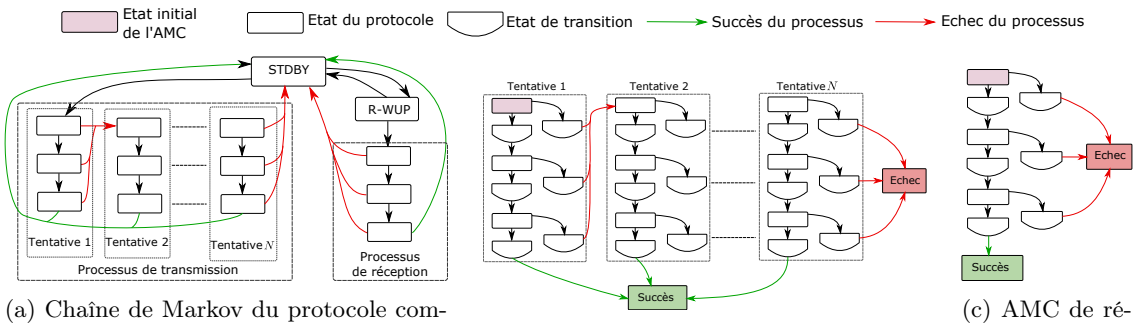
4.2 Evaluation analytique des wake-up radios

Un point clef dans la conception des réseaux de capteurs sans fil est le choix du protocole MAC, qui joue un rôle essentiel pour la fiabilité, la latence, le débit et l'efficacité énergétique (la communication étant généralement une des tâches les plus consommatrices d'énergie). Pour identifier les scénarios d'application pour lesquels les WuRx sont avantageuses par rapport aux protocoles traditionnels, un modèle analytique est requis. Dans cet partie, nous proposons une modélisation générique des protocoles MAC combinant des résultats analytiques et des mesures expérimentales. Le modèle repose sur les Chaînes de Markov Absorbantes (AMC - Absorbing Markov Chain) [52] et permet d'exprimer analytiquement la consommation d'énergie, la latence et la fiabilité. En utilisant des mesures expérimentales, nous avons prouvé que le modèle fournit des estimations précises. Finalement, des modélisations de protocoles MAC utilisant des WuRx ou utilisant l'approche cyclique traditionnelle ont permis de quantifier l'intérêt de l'utilisation des WuRx. Les contributions de cette étude sont donc :

- Une modélisation générique des protocoles MAC pour les réseaux de capteurs bas débit [aitaoudia17ton],
- La prise en compte des surcoûts matériels et logiciels d'une implémentation des protocoles [aitaoudia17gretsi],
- L'évaluation quantitative des protocoles utilisant les WuRx [aitaoudia16globecom].

4.2.1 Modélisation des protocoles

Cette section présente le modèle analytique proposé pour calculer la consommation d'énergie, la latence et la fiabilité des protocoles MAC. Un protocole MAC donné est modélisé par une Chaîne de Markov (MC) décrivant son fonctionnement. Une modélisation typique est illustrée par la Figure 4.5a. L'état STDBY (Stand-by) est l'état du protocole lorsqu'il ne reçoit ou ne transmet pas de paquet. La réception d'un paquet est classiquement précédée d'une étape d'écoute périodique du canal pour les protocoles à échantillonnage de préambule [53] ou de réception d'une balise. Cette étape correspond à l'état R-WUP (Receive Wake-UP) de la Figure 4.5a et peut conduire à la réception d'un paquet. Nous appelons respectivement *processus de transmission* et *processus de réception* les procédures définies par le protocole pour transmettre et recevoir un paquet. Chacun de ces procédés consiste en une ou plusieurs étapes élémentaires, e.g. la transmission d'une balise, la transmission d'une trame de données, la réception d'un accusé de réception (ACK - Acknowledgement). Certaines étapes peuvent éventuellement réussir ou échouer, e.g. la réception d'un ACK peut échouer en raison d'une interférence. Lors de la transmission d'un paquet, un nombre $N \in \mathbb{N}^*$ de tentatives est généralement autorisé. L'échec d'une tentative conduit soit à une nouvelle tentative, soit à l'échec du processus de transmission pour la dernière tentative.



(a) Chaîne de Markov du protocole complet.

(b) AMC de transmission.

(c) AMC de réception.

FIGURE 4.5 – Modélisation générique d'un protocole avec des chaînes de Markov.

Dans l'approche proposée, les processus d'émission et de réception sont modélisés individuellement par deux AMC illustrées par les Figures 4.5b et 4.5c. Ces AMC sont extraites de la MC modélisant l'ensemble du protocole et construites en définissant deux états d'absorption "Succès" et "Echec" notés "s" et "f" respectivement. Trois sortes d'états sont définis pour construire une AMC : les *états du protocole*, les *états de transition* et les *états finaux*. Les états du protocole sont définis par le protocole lui-même. Cependant, pour évaluer correctement la consommation d'énergie et la latence du protocole et du matériel, des états de transition, n'influant pas sur le comportement du protocole, sont nécessaires pour modéliser le coût d'énergie et de latence des transitions entre les états. Les états finaux sont les états "s" et "f".

En appliquant des résultats classiques de la théorie des AMC [52], chaque AMC représentant un processus de transmission/réception (noté t/r) peut être modélisée par sa matrice de transition $\mathbf{P_{t/r}}$ définie à partir des probabilités de transitions entre les différents états de l'AMC. De $\mathbf{P_{t/r}}$ se déduit son *vecteur fondamental*, notée $\mathbf{n_{t/r}}$, dont les éléments $n_{t/r,i}$ correspondent au nombre moyen de fois que la chaîne passe dans l'état transitoire i avant d'être absorbée. Enfin, le *vecteur de probabilité d'absorption*, notée $\mathbf{b_{t/r}}$, peut être calculé dont l'élément $b_{t/r,j}$ est la probabilité que la chaîne soit absorbée par l'état absorbant $j \in$ {s, f}.

Ces résultats sur les AMC permettent de déduire les performances analytiques des protocoles :

- La probabilité de succès de transmission/réception d'un paquet est tout simplement l'élément $b_{t/r,s}$ du vecteur de probabilité d'absorption de l'AMC de transmission/réception.
- Le coût énergétique de transmission ou de réception d'un paquet $\overline{e_{t/r}}$ est donné par :

$$\overline{e_{t/r}} = \mathbf{n_{t/r}} \cdot \mathbf{e_{t/r}}, \tag{4.1}$$

  avec $\mathbf{e_{t/r}}$ le *vecteur des coûts énergétiques*, i.e. le vecteur dont l'élément i est le coût énergétique induit au protocole quand il passe dans l'état transitoire i.
- La latence du processus de transmission d'un paquet est :

$$\overline{l_t} = \mathbf{n_{|s,t}} \cdot \mathbf{l_t}, \tag{4.2}$$

  avec $\mathbf{l_t}$ le vecteur des coûts de latence du processus de transmission et le $\mathbf{n_{|s,t}}$ le *vecteur fondamental conditionnel* dont l'élément $n_{|j,i}$ est le nombre moyen de fois que la chaîne est passée dans l'état transitoire i et qu'elle a été absorbée par l'état $j \in \{s, f\}$.
- La puissance moyenne consommée $\overline{P_C}$ due au protocole de communication peut être calculée par :

$$\begin{aligned}\overline{P_C} = \overline{a}\lambda_r\overline{e_r} + (\lambda_g + b_{t,s}\lambda_r)\,\overline{e_t} + \lambda_w e_w + \Big(1 - \overline{a}\lambda_r\,(\mathbf{n_r} \cdot \mathbf{l_r}) \\ - (\lambda_g + b_{t,s}\lambda_r)\,(\mathbf{n_t} \cdot \mathbf{l_t}) - \lambda_w l_w\Big) P_{SBY},\end{aligned} \tag{4.3}$$

  sous l'hypothèse qu'un nœud retransmet tous les paquets qu'il reçoit avec succès. λ_r et λ_g sont respectivement le taux moyen de réception et de génération de paquets en local. λ_w est le taux de réveil en réception. e_w et l_w sont respectivement le coût énergétique et la durée d'un unique réveil en réception. $\overline{a}$ est le nombre moyen de tentatives de transmission. Enfin, P_{SBY} est la puissance consommée par le nœud quand le protocole est dans l'état STDBY.

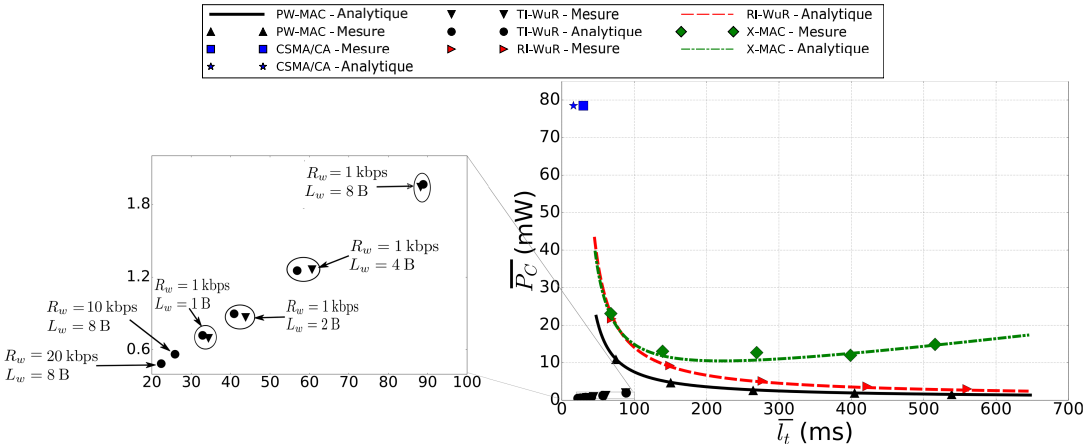


FIGURE 4.6 – Puissance consommée et latence des protocoles évalués : résultats analytiques et expérimentaux.

Le premier terme correspond à la puissance consommée lors de la réception de paquets, le second celle consommée lors de la transmission. Le troisième terme permet de tenir compte de la puissance consommée lors de réveils en réception (écoute périodique, transmission d'une balise, ...). Enfin, le dernier terme est la puissance consommée quand le nœud est endormi.

4.2.2 Performances analytiques et comparaisons expérimentales

Cinq protocoles ont été évalués : deux utilisant les WuRx (TI-WuR présenté dans la Figure 4.2 et RI-WuR, initié par le récepteur), PW-MAC [54] et X-MAC [55] sont deux protocoles dédiés à l'efficacité énergétique des réseaux de capteurs, et CSMA-CA [10] qui est le protocole utilisé dans la norme IEEE 802.15.4. Chaque protocole a été modélisé par le cadre proposé et implémenté sur les nœuds de capteur PowWow [44]. La Figure 4.6 donne les performances analytiques et les mesures expérimentales des protocoles en terme de latence et de consommation d'énergie. La première conclusion est que les modèles analytiques correspondent bien aux mesures expérimentales.

Ensuite, ces résultats permettent de comparer les protocoles entre eux. Les protocoles RI-WuR, PW-MAC et X-MAC dépendent de la fréquence de réveil du récepteur λ_w : $\lambda_w \in [1\,;\,20]$ Hz pour les résultats analytiques alors que les mesures expérimentales ont été réalisées pour des λ_w de 1, 1.3, 2, 4 et 10 Hz. Ces résultats montrent qu'un compromis entre la latence et la consommation d'énergie doit être fait en utilisant ces protocoles, et ce compromis est fixé par λ_w. CSMA/CA minimise la latence car les nœuds sont toujours en écoute : aucune synchronisation ni envoi de WuB ne sont requis. Cependant, cela se fait au prix d'une forte consommation d'énergie.

Pour le protocole TI-WuR, les résultats sont donnés pour différents débits binaires (R_w) et différentes tailles (L_w) des WuB. TI-WuR permet une latence et une consommation d'énergie nettement inférieures à celles des autres protocoles. Comme les WuB sont envoyés avec une puissance d'émission plus élevée et un débit binaire inférieur aux autres trames, réduire le temps de transmission WuB permet de réduire la latence et la consommation d'énergie.

4.2.3 Conclusion

Cette étude présente une modélisation générique des protocoles pour les réseaux de capteurs sans fil. Cette modélisation utilise les chaînes de Markov absorbantes et permet d'exprimer analytiquement la consommation d'énergie, la latence et la fiabilité tout en prenant en compte les surcoûts matériels engendrés par l'implémentation du protocole. Cet outil permet de comparer différents protocoles et notamment de montrer l'intérêt de l'utilisation des wake-up radios.

4.3 OPWUM : protocole MAC opportuniste

Dans le contexte d'un WSN multi-saut avec des nœuds équipés de récupérateurs d'énergie, un nœud peut décider de relayer ou non les paquets de ses prédécesseurs en fonction de son état énergétique. Afin de limiter les échanges d'informations entre les nœuds du réseau, la solution développée dans cette étude repose sur les communications opportunistes. Ces communications permettent de traiter le problème des liens non-fiables typiques des WSN à récupération d'énergie et d'améliorer l'efficacité énergétique en exploitant la diversité multi-nœuds [56]. Les solutions basées sur des fenêtres de contention temporelles sont prometteuses afin de permettre la sélection opportuniste du prochain relais. Cependant, l'écoute inutile peut être significative réduisant ainsi la durée de vie du réseau. Pour résoudre ce problème, nous proposons d'exploiter les WuRx.

Dans cette partie, nous présentons le protocole OPWUM (OPportunistic Wake Up MAC) [aitaoudia16js], un protocole multi-saut basé sur une contention temporelle. Il permet à un nœud la sélection opportuniste du meilleur relai parmi ses nœuds voisins en fonction d'une métrique donnée (par exemple, l'énergie restante), sans connaissances à-priori sur ses voisins. OPWUM exploite les wake-up radios afin de réduire considérablement la consommation d'énergie des nœuds. Des simulations réseaux exhaustives (avec le framework Green-Castalia d'OMNET++) montrent l'efficacité de OPWUM rapport au protocole 1-hop MAC, un protocole de l'état de l'art utilisant une contention temporelle [56].

4.3.1 Protocole MAC opportuniste utilisant la wake-up-radio

Le principe de la contention temporelle a été principalement abordé au niveau de la couche réseau. Watteyne et al. ont proposé de l'utiliser dans la couche MAC en concevant le protocole 1-hopMAC [56]. Un protocole basé sur une contention temporelle permet à un nœud de sélectionner le prochain relai en fonction d'une métrique donnée en n'ayant aucune connaissance de ses voisins. 1-hopMAC est un protocole semi-asynchrone avec une communication initiée par l'émetteur. Dans ce protocole, tous les nœuds se réveillent périodiquement pour écouter le canal et vérifier si un nœud voisin veut envoyer des paquets. La période de réveil des nœuds est notée T_{WI}. Lorsqu'un nœud veut envoyer un paquet, il envoie une balise RTS (Request To Send) suffisamment longue (supérieure à T_{WI}) afin que les nœuds voisins puissent la recevoir. La contention temporelle peut alors commencer. Quand un nœud veut transmettre un paquet, il diffuse une balise RTS. Ensuite, chaque nœud v qui reçoit la balise RTS et qui est un relais potentiel du nœud émetteur fixera un délai d'attente d'une durée $B(v) \in [0, D_c]$ avec D_c la durée de la fenêtre de contention. $B(v)$ est une durée qui dépend d'une métrique $M(v)$: $B(v) = f(M(v))$. Le mieux le nœud v se comporte par rapport à $M(v)$, le plus court sera $B(v)$. Dans le cas de la récupération d'énergie, si $M(v)$ est l'énergie restante du nœud v, un $M(v)$ faible impliquera un $B(v)$ important. Lorsque l'attente $B(v)$ expire, le nœud v envoie un trame CTS (Clear To Send). L'émetteur sélectionne le nœud qui a répondu le premier, et donc celui qui a le meilleur état en fonction de la métrique $M(v)$. Ainsi, le prochain relai est choisi de manière opportuniste.

Pour le protocole OPWUM, chaque nœud est équipé d'une wake-up radio et d'un émetteur-récepteur principal avec une détection de canal (CCA - Clear Channel Assessment). La Figure 4.7 présente un exemple de transmission de paquets utilisant OPWUM. La principale différence avec 1-hopMAC est l'utilisation des WuRx pour l'envoi des balises de contrôle. Après avoir reçu une balise RTS de l'émetteur sur sa WuRx, chaque récepteur potentiel calcule un délai d'attente en fonction de sa métrique d'état $M(v)$. Le récepteur qui a le plus court délai d'attente est donc le premier à répondre en envoyant une balise CTS sur la wake-up radio de l'émetteur. Ainsi, ce récepteur est choisi par l'émetteur pour devenir le prochain relai. Les autres récepteurs potentiels interrompent leur attente de plusieurs manières. Un récepteur peut recevoir également la balise CTS envoyée par le récepteur concurrent le plus rapide à répondre, provoquant ainsi l'interruption de son attente. Si des récepteurs ne reçoivent pas le CTS (e.g. parce que le relais choisi est hors de portée de leur WuRx), ils doivent être informés qu'ils ont perdu la contention. Ainsi, lorsque l'émetteur reçoit le CTS du premier récepteur, il envoie une trame About To Send (ATS) afin d'informer les nœuds restants encore en course que le relai suivant a déjà été choisi.

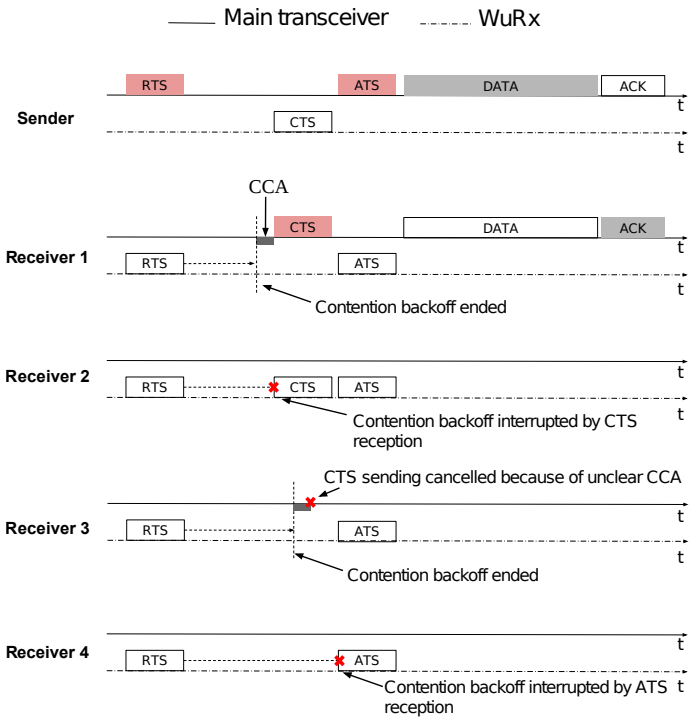


FIGURE 4.7 – Communication opportuniste utilisant OPWUM.

4.3.2 Évaluation de OPWUM

Environnement de simulation : OPWUM et 1-hopMAC ont été implémentés sous Green-Castalia/OMNeT++ [35]. Nous avons implémenté un module Castalia pour modéliser la WuRx utilisée dans [57]. Chaque simulation dure 3600 s (temps simulé). La puissance consommée par la radio principale est de 0.6 μW en veille, 26.7 mW pour l'envoi de trame hors WuB, de 80.1 mW pour l'envoi d'un WuB (RTS, CTS ou ATS) et de 22 mW en écoute. La puissance consommée par la WuRx est de 1.83 μW quand seul le front-end analogique est actif et de 284 μW quand le front-end et le microcontrôleur ULP sont actifs. La durée d'un WuB est de 5.2 ms, celle des données est de 12.5 ms et celle de l'accusé de réception ACK est de 3.33 ms.

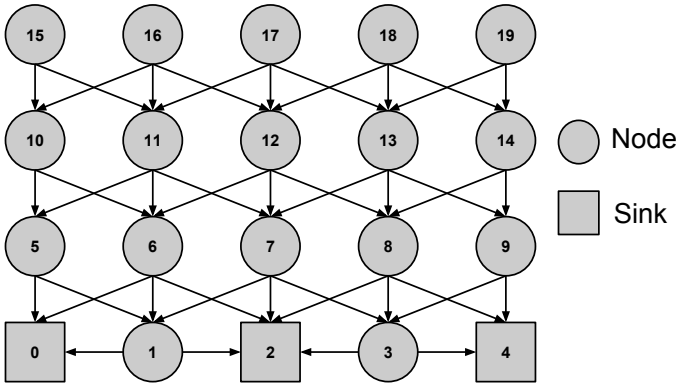
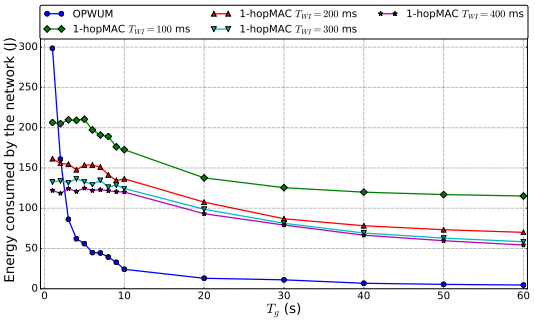
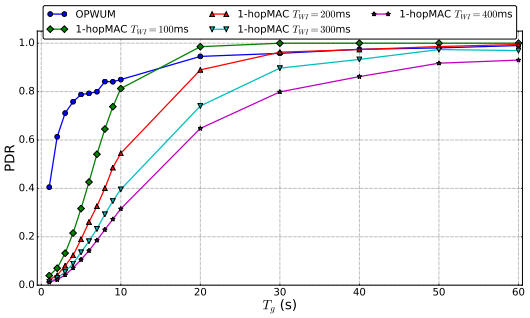


FIGURE 4.8 – Réseau utilisé pour comparer OPWUM et 1-hopMAC. Les flèches indiquent les récepteurs potentiels de chaque nœud.

Le scénario courant de la littérature décrit par la Figure 4.8 est implémenté. C'est un réseau multi-sauts statique, avec trois puits de réception. Les nœuds intermédiaires sont en charge de la transmission des paquets de leurs prédécesseurs immédiats. A l'exception des puits, chaque nœud génère périodiquement des paquets. La période de génération de paquet est notée T_g et est



(a) Énergie dépensée par le réseau en fonction de la période de génération des paquets T_g. La fenêtre de contention est fixée à 50 ms.



(b) Taux de bonne réception PDR en fonction de la période de génération des paquets T_g. La fenêtre de contention est fixée à 50 ms.

FIGURE 4.9 – Energie dépensée et Taux de bonne réception des paquets des protocoles OPWUM et 1-hopMAC.

identique pour tous les nœuds. Pour évaluer le routage opportuniste, chaque nœud possède plusieurs relais potentiels, indiqués par les flèches sur la Figure 4.8. En utilisant des fenêtres de contention temporelles, OPWUM et 1-hopMAC choisissent le prochain relai parmi les relais potentiels à chaque tentative de transmission de paquets. Pour 1-hopMAC, différents intervalles de réveil (T_{WI}), allant de 100 à 400 ms sont considérés. De plus, nous ne faisons dans un premier temps aucune hypothèse sur la métrique d'état utilisée pour calculer le délai d'attente de la contention, elle est donc choisie de manière uniforme dans l'intervalle $[0, D_c]$.

Résultats de simulations La Figure 4.9a représente l'énergie totale consommée par le réseau en fonction de la période de génération de paquet T_g, pour une fenêtre de contention D_c de 50 ms. On constate que la consommation d'énergie de 1-hopMAC dépend de la valeur de T_{WI} et qu'elle a tendance à réduire pour des valeurs élevées de T_{WI}. OPWUM permet une amélioration significative de la consommation d'énergie avec une réduction allant jusqu'à un facteur 5 par rapport à 1-hopMAC. La seule exception est pour un trafic élevé, lorsque T_g est inférieure à 5 s. Mais, comme nous le verrons ensuite, le taux de bonne réception des paquets de 1-hopMAC s'effondre à presque 0% quand le trafic est élevé.

La Figure 4.9b donne le taux de bonne réception des paquets (PDR - Packet Delivery Rate) d'OPWUM et de 1-hopMAC en fonction de T_g, pour une fenêtre de contention D_c de 50ms. Pour des trafics relativement élevés (i.e. T_g inférieure à 10 s), OPWUM est plus performant que 1-hopMAC. De plus, le taux n'est jamais inférieur à 40% avec OPWUM, alors qu'il peut être proche de 0% pour 1-hopMAC. Pour des débits faibles (i.e. T_g supérieure à 10 s), les résultats d'OPWUM sont comparables à ceux de 1-hopMAC avec un T_{WI} de 100 ms. Mais comme le montrent les résultats de consommation d'énergie, cette configuration de 1-hopMAC est plus gourmande en énergie.

4.3.3 Application à la récupération d'énergie

Le protocole OPWUM est maintenant utilisé pour un réseau de capteurs à récupération d'énergie. Le réseau de capteur de la Figure 4.10 a été implémenté dans Green-Castalia. Dans ce cas d'étude, les nœuds de la rangée supérieure (les nœuds « source ») génèrent des paquets à un rythme moyen d'un paquet toutes les 10 s, alors qu'un seul nœud (le "puits") reçoit les paquets générés. Tous les autres nœuds servent seulement de relais en utilisant le protocole OPWUM. Chaque nœud, sauf le puits, est équipé d'un dispositif de récupération d'énergie (e.g. un panneau solaire). Les nœuds situés sous les nuages (a) et (b) récupèrent moins d'énergie que les autres nœuds : le nuage (a) est "plus mince" que le nuage (b), ce qui signifie que le taux de récupération d'énergie des nœuds situés sous le nuage (a) est plus élevé que celui des nœuds situés sous le nuage (b). Dans ce contexte, la métrique M utilisée par les relais potentiels pour calculer leur temps de réponse à la contention est liée à l'énergie récupérée. Ainsi le temps de réponse sera inversement proportionnel à l'énergie récupérée : plus un nœud aura d'énergie, plus il cherchera à relayer des paquets de ses prédécesseurs.

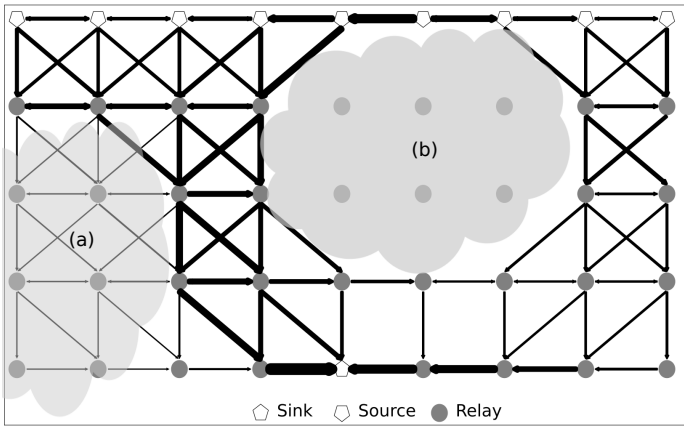


FIGURE 4.10 – Application de OPWUM aux réseaux de capteurs à récupération d'énergie.

La simulation dure 10000 s (temps simulé). L'ensemble de récepteurs potentiels de chaque nœud est défini par une table de routage, il consiste en les nœuds voisins du réseau situés soit sur la même ligne soit sur la ligne inférieure. OPWUM est ensuite utilisé pour choisir de manière opportuniste un relai parmi cet ensemble chaque fois qu'un paquet doit être envoyé. La Figure 4.10 montre les liens de communications utilisés. L'épaisseur des flèches est proportionnelle à leur utilisation. Comme nous pouvons le voir, seuls quelques chemins de routage contiennent des nœuds de la région (a), alors qu'aucun chemin de routage ne contient des nœuds de la région (b). OPWUM permet aux nœuds de choisir les récepteurs potentiels qui récupèrent le plus d'énergie. Les chemins de routage contournent les zones nuageuses afin d'atteindre le puits, permettant ainsi aux nœuds de survivre pendant les périodes de pénurie d'énergie.

4.3.4 Conclusion

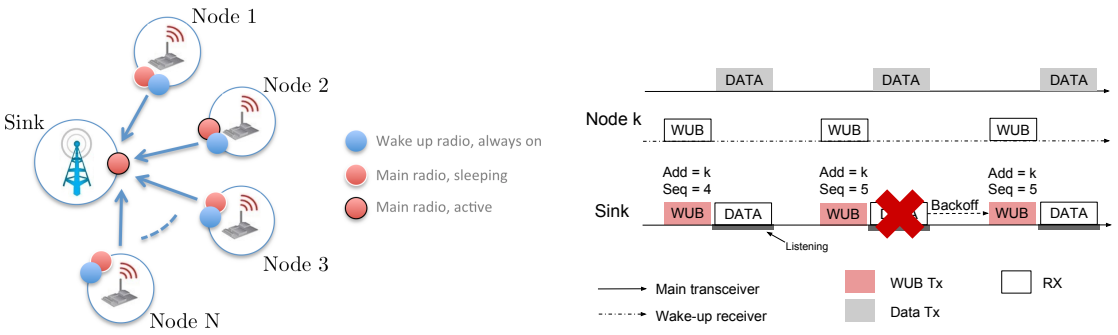
OPWUM est un des rares protocoles MAC qui utilise pleinement la wake-up radio afin de réduire l'énergie globale du réseau. Dans un réseau multi-saut, OPWUM permet aux nœuds une sélection opportuniste du prochain relai parmi un sous-ensemble de leurs nœuds voisins, sans connaissance a priori sur ce voisinage. Ces travaux ont été étendus à un protocole tenant compte de l'énergie pour utiliser ou non les relais, protocole qui a été implémenté sur une plate-forme matérielle [djidi18icecs].

4.4 SNW-MAC : collecte des données dans un réseau en étoile

Cette partie présente un nouveau protocole MAC asynchrone pour les réseaux en étoile, où tous les nœuds sont connectés à un puits central qui collecte les données. SNW-MAC (Star Network WuRx - MAC) utilise les WuRx pour réduire l'énergie nécessaire à la transmission d'un paquet et pour totalement annuler les collisions dans le cas d'un réseau en étoile. La contribution majeure de cette étude est l'évaluation expérimentale. SNW-MAC a été implémenté et testé en conditions réelles, en utilisant la plate-forme PowWow [44] qui intègre des capacités de récupération d'énergie et une ULP WuRx [58]. Une comparaison a été réalisée avec deux protocoles MAC faible consommation traditionnels implémentés sur le même matériel. Il s'agit surement de la première expérimentation combinant récupération d'énergie et WuRx [aitaoudia18sensors].

4.4.1 Le protocole Star Network WuRx - MAC

La topologie de réseau ciblée est la topologie en étoile décrite sur la Figure 4.11a. Le réseau est composé de deux types de nœuds. Le nœud puits "sink" initie les communications et récupère les données des autres nœuds. Il utilise une seule radio (IEEE 802.15.4 dans nos expérimentations), qui est tout le temps en écoute. Les autres nœuds sont équipés d'une radio principale et d'une



(a) Topologie en étoile. (b) Transmission des paquets avec SNW-MAC.

FIGURE 4.11 – SNW-MAC : topologie utilisée et description du protocole.

Sync bits	Sensor address	Sequence number
3 bits	8 bits	8 bits

FIGURE 4.12 – Structure du Wake-up Beacon utilisé par SNW-MAC.

WuRx. La WuRx est toujours en écoute, alors que la radio principale est allumée seulement sur demande du nœud puits.

Le protocole SNW-MAC est un protocole initié par le récepteur. La Figure 4.11b illustre la transmission d'un paquet en utilisant SNW-MAC. Le puits initie la communication en envoyant un WuB contenant l'adresse d'un nœud spécifique et écoute ensuite le canal afin de recevoir le paquet. Le nœud ciblé est réveillé par sa ULP WuRx, et peut commencer à envoyer le paquet de données. Par rapport aux protocoles initiés par le récepteur traditionnel, cette approche réduit l'énergie consommée par le puits et les nœuds puisqu'aucun processus de rendez-vous n'est nécessaire. Ne pas envoyer périodiquement des WuB réduit la consommation du puits alors que celle des nœuds est réduite par l'absence d'écoute inutile du canal.

Contrôle des erreurs et retransmission En coordonnant entièrement l'envoie des paquets au niveau du puits, le protocole SNW-MAC annule le risque de collisions puisque chaque nœud est interrogé spécifiquement et indépendamment des autres. Cependant, les interférences du canal sans fil peuvent conduire à des pertes de paquets. Un contrôle d'erreur et de retransmission des paquets est donc nécessaire. Comme le récepteur est entièrement responsable de la coordination de la transmission des paquets, il est responsable de la détection de l'erreur de transmission et de la planification d'une autre tentative d'envoi. Ce contrôle utilise un WuB spécifique décrit dans la Figure 4.12. Sa taille est de 19 bits : 3 bits de synchronisation, les 8 bits d'adresse du nœud interrogé et 8 bits du numéro de la sequence du paquet de données souhaité.

Le puits tient à jour une table qui associe à chaque nœud le numéro du prochain paquet à envoyer. Quand la WuRx d'un nœud reçoit le WuB, elle lit à la fois l'adresse et le numéro de séquence. Si l'adresse correspond à celle du nœud, la WuRx réveille le MCU principal et lui transmet le numéro de séquence par un port série. Le paquet correspondant au numéro de séquence demandé par le puits est ensuite envoyé. Si le puits ne reçoit pas le paquet demandé, il programme une nouvelle tentative pour le même numéro de séquence après un temps d'attente aléatoire. Ce mécanisme est illustré sur la Figure 4.11b. Comparé aux schémas traditionnels de contrôle d'erreur reposant sur l'envoi de trame ACK, le surcoût énergétique est réduit à la fois pour le puits qui ne doit pas envoyer d'ACK et pour les nœuds qui ne doivent pas attendre de le recevoir. Néanmoins, ce gain en énergie est contre-balancé par l'utilisation de WuB plus long (contenant le numéro de séquence).

4.4.2 Protocole expérimental : combinaison à la récupération d'énergie

Le protocole SNW-MAC a été implémenté sur la plateforme PowWow intégrant des mécanismes de récupération d'énergie. Cette section présente les configurations logicielles et matérielles utilisées pour les expérimentations.

Architecture logicielle

L'architecture logicielle du nœud décrite sur la Figure 4.13 et est composé de deux composants : le gestionnaire d'énergie (EM) et le protocole MAC de communication.

Gestion d'énergie L'objectif de l'EM est d'adapter dynamiquement les performances du nœud en fonction des variations de l'énergie que le nœud peut récupérer dans son environnement. On suppose le temps divisé en intervalles de temps identiques de durée T_s, le taux de génération de paquets $\chi_g[k]$ (en paquets par seconde) du nœud pour l'intervalle courant k est fixé par l'EM au début de chaque intervalle en fonction de l'énergie récupérée $e_h[k]$.

L'EM est composé de deux sous-parties comme décrit par la Figure 4.13. L'estimateur de budget énergétique évalue le budget énergétique $e_b[k]$ que le nœud peut consommer pendant l'intervalle de temps tout en restant en vie. L'estimateur utilisé dans cette étude est une version simplifiée de Fuzzyman [aitaoudia16icc] qui utilise seulement l'énergie résiduelle associée à une base de règle pour estimer le budget énergétique.

Le second module calcule le débit $\chi_g[k]$ en fonction du budget énergétique. Le taux de génération de paquets $\lambda_g[k]$ d'un nœud dépend fortement du protocole MAC utilisé. En notant e_T le coût énergétique et τ_T le temps total d'une transmission d'un paquet, l'énergie consommée par le nœud pendant l'intervalle de temps k est donné par :

$$e_c[k] = \chi_g[k] T_s e_T + (T_s - \tau_T \chi_g[k] T_s) P_S, \tag{4.4}$$

avec P_S la puissance consommée par le nœud quand le MCU et la radio sont en vieille. Ainsi, afin que l'énergie consommé $e_c[k]$ soit égale au budget énergétique $e_b[k]$, le taux de génération de paquets doit être égal à :

$$\chi_g[k] = \frac{\frac{e_b[k]}{T_s} - P_S}{e_T - \tau_T P_S}. \tag{4.5}$$

Pour SNW-MAC, chaque nœud inclut son taux de génération de paquets dans la trame de données. Le puits peut ainsi tenir à jour une table qui associe à chaque nœud son taux de génération de paquets et donc interroger chaque nœud aux bons instants. Le taux de génération de paquets est codé sur un mot entier de 16 bits, engendrant un faible surcoût de l'insertion de cette information dans la trame de données.

Comparaison à des protocoles de l'état de l'art SNW-MAC est comparé aux protocoles PW-MAC [54] et X-MAC [59], deux protocoles de l'état de l'art couramment utilisés. X-MAC est un protocole initié par l'émetteur, la Figure 4.14a illustre la transmission d'un paquet avec X-MAC. Quand un paquet de données doit être envoyé, le nœud le transmet continuellement, chaque transmission étant suivie par une période d'écoute. Le processus continu tant qu'une trame ACK venant du puits n'est pas correctement reçue. De son coté, le puits se réveille périodiquement et écoute le canal. S'il détecte une activité, il reste éveillé jusqu'à la réception complète du paquet de données et accuse ensuite la bonne réception du paquet.

PW-MAC est un protocole initié par le récepteur et permet d'obtenir une efficacité énergétique à la fois au niveau de l'émetteur et du récepteur. Avec PW-MAC, les nœuds prédisent précisément les instants de réveil du puits en utilisant un mécanisme de prédiction et en considérant la dérive de l'horloge et les latences logicielles et matérielles. La transmission d'un paquet avec PW-MAC est décrite dans la Figure 4.14b. Quand un nœud doit envoyer un paquet, il calcule une prédiction du prochain réveil du puits et se réveille à cet instant. Il attend ensuite l'envoi d'une balise par le

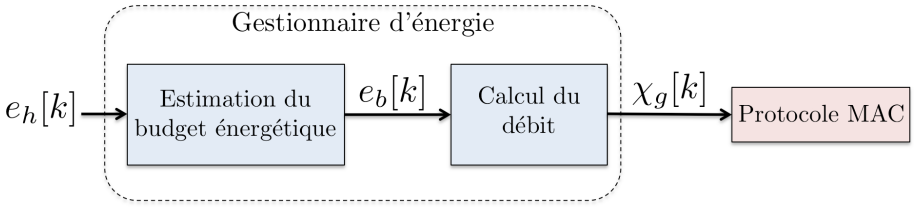
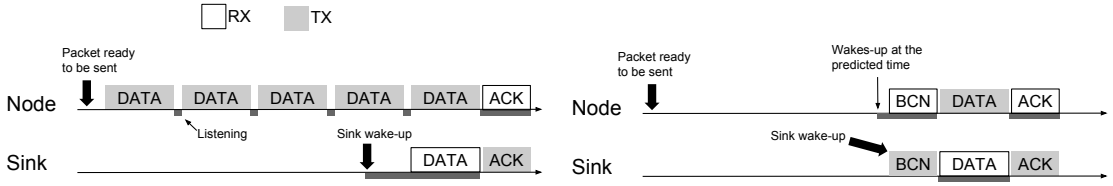
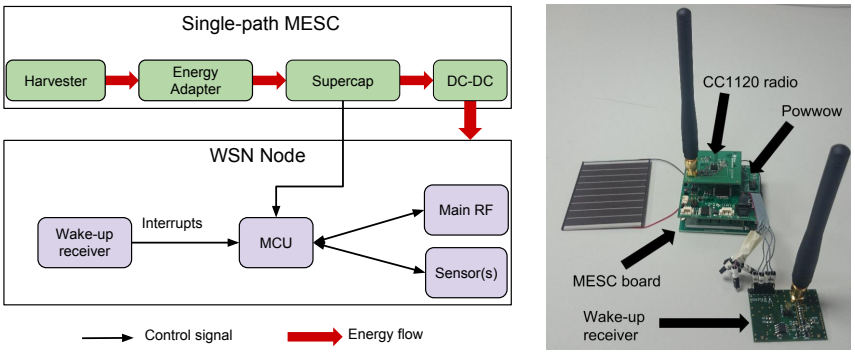


Figure 4.13 – Architecture logicielle.



(a) Transmission d'un paquet avec XMAC. (b) Transmission d'un paquet avec PW-MAC.

FIGURE 4.14 – Les protocoles X-MAC et PW-MAC.



(a) (b)

FIGURE 4.15 – Architecture matérielle du nœud utilisant la récupération d'énergie et une WuRx : (a) description fonctionnelle (b) plateforme PowWow utilisée pour les expérimentations.

puits (représenté par la trame BCN dans la Figure 4.14b). Une fois la balise reçue, le nœud envoi le paquet de données et attend une trame ACK. A chaque transmission de paquet, une erreur de prédiction est calculée afin de corriger la prochaine prédiction.

Architecture matérielle

La Figure 4.15 décrit l'architecture matérielle du nœud ainsi qu'une photographie de la plateforme PowWow utilisée pour les expérimentations [44]. La WuRx introduite dans la Section 4.1.2 est associée à la plateforme PowWow développée par le laboratoire IRISA qui intègre une architecture de récupération d'énergie.
Les protocoles SNW-MAC, X-MAC et PW-MAC ont été implémentés sur cette plateforme ainsi que l'EM présenté précédemment. Les paramètres utilisés pour les expérimentations sont donnés dans la Table 4.1.

Architecture de la plateforme PowWow utilisant la récupération d'énergie De nombreuses plateformes de réseaux de capteurs à récupération d'énergie ont été proposées dans le

Parametres		Valeurs
MAC	Intervalle de réveil du puits (X-MAC and PW-MAC)	250 ms
	Nombre maximum de retransmissions	2
PHY	Débit des paquets WuB	1 kbps
	Débit des paquets Data/ACK/beacon	20 kbps
	Puissance de transmission des paquets WuB	12.5 dBm
	Puissance de transmission des paquets Data/ACK/beacon	−6 dBm

TABLE 4.1 – Paramètres utilisés pour les expérimentations.

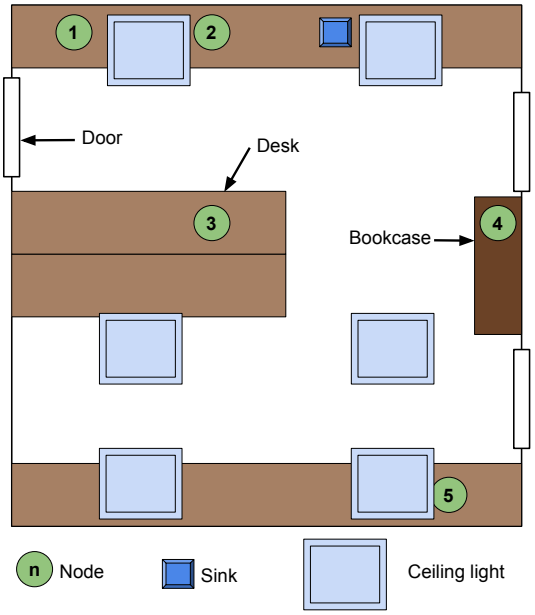


FIGURE 4.16 – Déploiement du réseau en étoile.

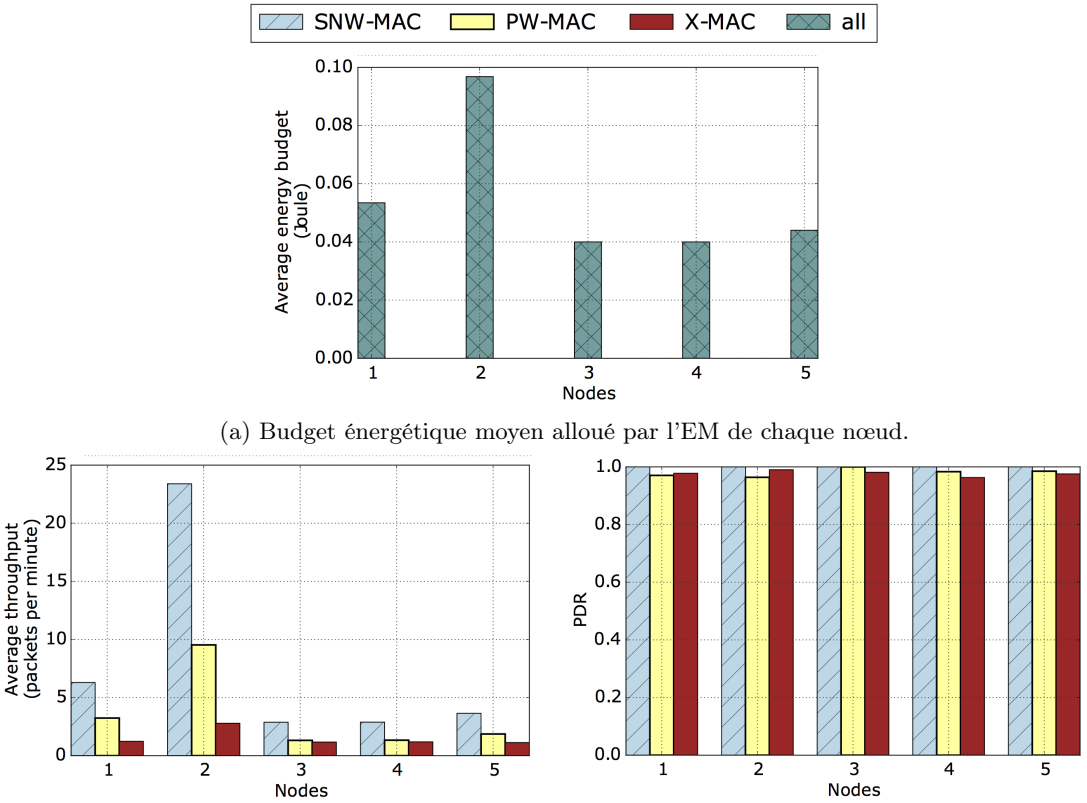
milieu universitaire et dans l'industrie. Dans ce travail, nous considérons l'architecture Multiple Energy Source Converter (MESC) proposé dans [60]. Dans la version dite "single-path" de cette architecture, il n'y a qu'un seul dispositif de stockage d'énergie et toute l'énergie récupérée est utilisée pour charger le dispositif de stockage qui alimente le nœud directement par un convertisseur DC-DC. La Figure 4.15 montre l'architecture du bloc MESC, qui peut être utilisée par différents récupérateurs d'énergie (cellules photovoltaïques, générateurs thermoélectriques, éoliennes) en utilisant un adaptateur d'énergie approprié afin de normaliser l'énergie en sortie. Une super-capacité est utilisée comme dispositif de stockage car les super-capacités ont une durée de vie importante et offrent une densité de puissance supérieure à celle des batteries [61].

La plateforme PowWow [44] est utilisée dans cette étude, elle utilise l'architecture MESC de récupération d'énergie et est équipée d'un micro-contrôleur MSP430 et d'une puce radio Texas Instruments CC1120. Une super-capacité de 0.9 F est utilisée avec une tension maximale de 5.0 V et une tension minimale de 2.8 V requise pour alimenter le nœud.

4.4.3 Résultats expérimentaux

Les protocoles MAC ont été implémentés sur un réseau avec une topologie en étoile et composé de 6 nœuds PowWow (incluant le puits). Les nœuds étaient exclusivement alimentés par une lumière fluorescente d'intérieure, permettant une reproductibilité des expérimentations. De plus, les nœuds étaient positionnés avec différentes conditions d'éclairage, comme montré dans la Figure 4.16. Les nœuds 1, 2 and 5 étaient placés sur des bureaux, directement sous les lumières des plafonds alors que le nœud 3 était placé dans une zone plus sombre et le nœud 4 sur une armoire, proche du plafond, recevant ainsi moins d'énergie que les autres. Chaque expérimentation durait 3 heures, et les nœuds PowWow était équipés des ULP WuRx quand le protocole SNW-MAC était évalué.

Les métriques utilisées pour comparer les protocoles sont le taux moyen de génération des paquets (en paquets par minutes) et le taux de paquets correctement reçus (PDR - Packet Delivery Rate). La Figure 4.17 montre les résultats obtenus. Tout d'abord, la Figure 4.17a donne le budget énergétique moyen alloué par l'EM pour chaque nœud. Ce budget diffère d'un nœud à l'autre en fonction de la quantité d'énergie récupérée. Les performances de chaque nœud sont de manière évidente liées à au budget énergétique. La Figure 4.17b donne le taux moyen de génération des paquets atteint par les différents protocoles MAC. SNW-MAC surpasse de manière significative les deux autres protocoles, permettant d'atteindre un taux moyen jusqu'à deux fois supérieur à celui de PW-MAC pour le nœud 2. Ce gain est du au faible coût énergétique d'envoi des paquets. Enfin,



(a) Budget énergétique moyen alloué par l'EM de chaque nœud.

(b) Taux moyen de paquets généré par chaque nœud.

(c) PDR de chaque nœud.

FIGURE 4.17 – Résultats expérimentaux.

la Figure 4.17c donne le PDR obtenu par les trois protocoles. SNW-MAC est le seul protocole à atteindre un PDR de 100 % pour tous les nœuds. Ces résultats démontrent la grande fiabilité que permettent les protocoles utilisant les WuRx.

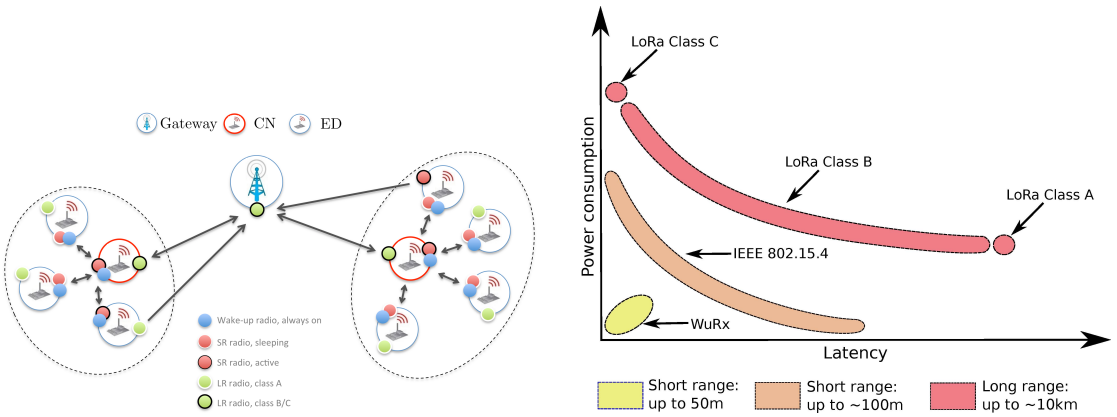
4.4.4 Conclusion

Cette partie a présenté un protocole MAC asynchrone dans le contexte de la collecte de données dans une topologie en étoile. La solution proposée utilise des WuRx pour augmenter l'efficacité énergétique du réseau de capteurs sans fil. Ce nouveau schéma est conçu pour être implémenté sur du matériel réel. Les résultats expérimentaux ont montré que l'approche proposée permet d'obtenir un débit 2 fois supérieur à celui des protocoles MAC de l'état de l'art, PW-MAC et X-MAC.

4.5 Autres contributions

4.5.1 Communications hétérogènes courtes et longues portées associant LoRa et WuRx

Apparues ces dernières années, des technologies sans fil permettent de transmettre à des portées de plusieurs kilomètres, au prix de débits généralement inférieurs à 10 kbps et avec une consommation d'énergie similaire aux transmetteurs habituels des WSN. Un exemple d'une telle technologie est LoRa [62]. LoRa fonctionne dans les bandes ISM 868/915 MHz et permet une portée théorique allant jusqu'à 22 km, et un débit compris entre 0.37 et 46.9 kbps [63]. Une communication LoRa en voie montante, i.e. des nœuds vers la passerelle, se fait avec une faible latence, car la passerelle écoute en permanence le canal. La communication en voie descendante nécessite un compromis entre la latence et la consommation d'énergie. Cependant, certaines applications, telles que la surveillance industrielle de machines, nécessitent à la fois une faible latence et une faible consommation



(a) Architecture de réseau hétérogène.

(b) Compromis latence-consommation pour différentes radios.

FIGURE 4.18 – Utilisation de plusieurs radios pour les communications hétérogènes : courte et longue portée.

d'énergie.

Nous avons donc proposé une architecture de réseaux [aitaoudia16dsd] [aitaoudia17micpro] qui repose sur des nœuds à plusieurs radios [magno17date]. Ces nœuds utilisent notamment des WuRx qui permettent des communications asynchrones, les nœuds ne se réveillant que pour communiquer. Ces radios réduisent donc la consommation d'énergie et la latence des communications au prix d'une portée réduite. Ce désavantage justifie l'architecture de réseaux proposée qui combine à la fois communications LoRa et WuRx. Dans l'architecture de réseaux proposée, les nœuds sont organisés dans une topologie en étoile, quelques dizaines de mètres autour d'un nœud central, comme le montre la Figure 4.18a. Une passerelle, située à une grande distance (quelques kilomètres) du réseau, recueille les données détectées et envoie des commandes aux nœuds (e.g. pour activer les actionneurs ou pour régler des paramètres de détection). Comme chaque nœud est équipé d'une WuRx en plus d'une radio LoRa, le nœud central peut réveiller un ou plusieurs nœuds en envoyant des balises de réveil [aitaoudia16dsd]. Dans cette étude, nous avons évalué le compromis entre la consommation d'énergie et la latence pour différents schémas LoRa avec l'architecture proposée. Ce compromis est illustrée par la Figure 4.18b qui montre la variété des configurations permise par la plateforme proposée.

4.6 Conclusion

L'autonomie d'un nœud de capteur ne peut être pas obtenue uniquement par une alimentation par récupération d'énergie. Il faut aussi réduire la consommation du nœud et donc le coût énergétique des communications. Les protocoles MAC ont donc un rôle important, étant responsable des communications point-à-point. Ce chapitre traite de l'apport de l'utilisation des wake-up radios dans la conception des protocoles efficaces en énergie.

Nous avons tout d'abord proposé une modélisation analytique des protocoles MAC utilisant les chaînes de Markov absorbantes. Cette modélisation a notamment permis de mettre en avant l'intérêt des wake-up radios. Ensuite, deux protocoles tirant pleinement profit des wake-up radios sont présentés : le premier pour le choix opportuniste d'un relai, le second pour la collecte des données dans un réseau en étoile.

Dans l'objectif de combiner expérimentalement la récupération d'énergie et les wake-up radios, nous avons conçu et implémenté un gestionnaire d'énergie basé sur Fuzzyman et un protocole MAC exploitant les wake-up radios. L'expérimentation était constituée de cinq nœuds PowWow organisés en étoile. Les résultats expérimentaux ont montré que, dans des conditions similaires, l'utilisation de wake-up radios permet l'amélioration du débit de paquets transmis.

Une perspective de ces travaux est la validation expérimentale du protocole OPWUM pour un réseau à récupération d'énergie et son utilisation pour résoudre efficacement des algorithmes dis-

tribués de gestion d'énergie. L'expérimentation d'un réseau hétérogène combinant wake-up radios et communications LoRa peut permettre également de mettre en avant l'utilisation des wake-up radios. La sélection dynamique du nœud central au sein du réseau courte portée peut également être étudiée dans le cas d'un scénario où l'énergie des nœuds varieraient au cours du temps.

Enfin, la modélisation des protocoles MAC peut être intégrée dans un outil d'évaluation et de comparaison de protocoles. Cet outil pourrait proposer des bibliothèques de composants logiciels et matériels et intégrer des modèles théoriques de couches physiques (e.g. de modulations, de canaux de propagation) et de couches réseaux (e.g. de collisions, de topologies).

Troisième partie

Les radios logicielles efficaces en énergie : prototypage rapide et exploration de l'espace de conception

Chapitre 5

Description haut-niveau des formes d'ondes et prototypage rapide

Le prototypage rapide d'applications pour les systèmes de communications est une approche importante pour réduire la phase de conception et donc le temps de mise sur le marché des systèmes. La radio logicielle est une solution permettant l'implémentation rapide sur des plateformes matérielles. Cependant, si la radio logicielle offre une nouvelle infrastructure de conception des communications sans fil et mobile, les solutions basées sur des processeurs présentent des limites en termes de capacité de calcul, de temps de reconfiguration ou de consommation d'énergie [64].

Une alternative à cette radio uniquement logicielle est le développement d'une radio logicielle sur des architectures reconfigurables de type FPGA. En effet, les FPGA apparaissent comme un support parfaitement adapté à la mise en œuvre d'une radio logicielle tant pour leur puissance de calcul que pour leurs capacités en matière de reconfiguration. Dans ce chapitre, nous étudions une radio logicielle sur une architecture reconfigurable permettant d'atteindre l'objectif du prototypage rapide d'une couche physique flexible et faible consommation.

Les deux clés fondamentales pour atteindre cet objectif sont une méthodologie de développement spécifique permettant la flexibilité d'une spécification avec une description de haut niveau et une architecture modulaire permettant le déploiement de diverses applications (e.g. couches physiques, formes d'onde, protocoles). On se propose à travers nos travaux d'outiller la communauté radio logicielle d'un flot de conception basé sur un langage dédié combiné à la synthèse de haut niveau. La Section 5.2 présente le langage et le flot de conception associé. Une validation expérimentale sur la forme d'onde IEEE 802.15.4 est présentée dans la Section 5.3 ainsi que l'exploration de l'espace de conception permise par le flot de conception. Enfin, la Section 5.4 résume les autres contributions associées à ce flot de conception. La partie 5.4.1 présente la génération de blocs de traitement reconfigurables dynamiquement en associant la reconfiguration dynamique partielle et la synthèse de haut niveaux. L'étude d'un récepteur OFDM (Orthogonal Frequency Division Multiplexing) faible consommation utilisant plusieurs dimensionnements virgule fixe est introduite dans la partie 5.4.2.

Sommaire

5.1 La radio logicielle et le prototypage rapide 72
5.1.1 La radio logicielle 72
5.1.2 Méthodes de conception 73
5.1.3 La synthèse de haut-niveau 74
5.2 Description des formes d'ondes 75
5.2.1 Langage de description spécifique (DSL) 75
5.2.2 Flot de conception 77
5.3 Prototypage rapide : expérimentation et validation 78
5.3.1 Description DSL de la norme IEEE 802.15.4 79
5.3.2 Implémentation de la norme IEEE 802.15.4 80
5.3.3 Exploration de l'espace de conception 81
5.4 Autres contributions 82
5.4.1 Reconfiguration dynamique des formes d'ondes 82
5.4.2 Adaptation dynamique d'un récepteur OFDM en virgule fixe 84
5.5 Conclusion 85

5.1 La radio logicielle et le prototypage rapide

5.1.1 La radio logicielle

La radio logicielle a été introduite au début des années 90 par Joe Mitola [65] qui à travers des articles pointait du doigt l'inéluctable évolution des plateformes radio vers des plateformes comprenant de plus en plus de composants logiciels. Ainsi, il prévoyait que les approches traditionnelles de conception radio, qui consistaient essentiellement en l'implémentation d'architectures dédiées, seraient peu à peu abandonnées au profit d'une conception centrée sur des composants logiciels. Elle a naturellement été proposée, dans un premier temps, comme support matériel à la radio cognitive [66] : une radio intelligente qui peut s'adapter à l'environnement et coopérer avec d'autres équipements de type radio logicielle. La radio logicielle permet la reconfiguration dynamique de l'architecture en bande de base par le biais d'un ou plusieurs processeurs hôtes. Elle offre ainsi la possibilité de changer le schéma de codage, de modulation, de bande passante et de technique d'accès au canal sans modifier directement le matériel [67].

Une plateforme radio logicielle idéale est illustrée à travers la Figure 5.1 dans laquelle les données reçues sont échantillonnées directement après les antennes. Cette architecture permettrait de réaliser tous les traitements Radio Fréquence (RF) et bande de base sur un processeur ou un microprocesseur généraliste. Une telle architecture est qualifiée d'idéale du simple fait que la technologie actuelle ne permet pas la réalisation d'un tel scénario où une très large bande de fréquence (typiquement plusieurs GHz) serait numérisée. Les solutions radios logicielles rencontrées à ce jour intercalent un *front-end* radio analogique entre les convertisseurs et les composants numériques afin de démoduler le signal en fréquence intermédiaire ou directement en bande base.

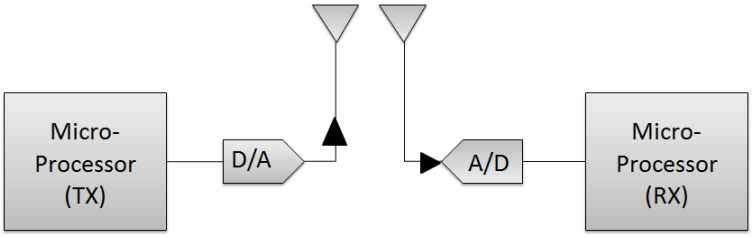


FIGURE 5.1 – Plateforme d'émission-réception radio logicielle idéale.

Une radio logicielle est caractérisée par trois caractéristiques majeures : (i) la programmabilité, pour pouvoir décrire l'application (e.g. une couche physique, une forme d'onde, un protocole) avec un langage de haut niveau, (ii) la portabilité, pour pouvoir adresser différents types de plateforme, et (iii) la reconfiguration, pour pouvoir modifier l'application. Ces éléments sont à la base de la définition d'une radio logicielle et de nombreux travaux de recherche mettent l'accent sur au moins une de ces trois caractéristiques [68–70]. Ces travaux ont conduit principalement à définir d'une part un ensemble de méthodologies de conception dédiées à la radio logicielle et d'autre part un certain nombre de plateformes radios logicielles. Ces plateformes peuvent être classées en trois familles en fonction de ressources matérielles utilisées [69] :

— Les plateformes se basant sur un processeur généraliste (GPP - General Purpose Processor) : ces plateformes réalisent le traitement en bande de base sur un ordinateur associé à un *front-end* radio pour la transmission du signal dans le canal. L'USRP [71] (Universal Software Radio Peripheral) est l'exemple le plus connu de ce type de plateforme.
— Les plateformes hybrides : elles associent au processeur généraliste des accélérateurs dédiés à certaines opérations de traitement du signal. Ces composants varient entre des FPGA, des microprocesseurs DSP, des Graphics Processing Units (GPU) ou différentes combinaisons. Des exemples de ce type d'architectures sont la plateforme KUAR de l'Université de Kansas [72], ADRES [73] développée par l'IMEC, NXP EVP16 [74], le circuit Tomawak [75] de l'université de Dresde et l'architecture Embb [76] proposée par Eurecom.
— Les plateformes se basant uniquement sur un FPGA : elles réalisent l'ensemble du traitement en bande de base sur des FPGA de taille importante et capables de supporter des

formes d'ondes relativement complexes. Le réseau sur puce Magali [77] du CEA-LETI et les plateformes WARP [78] de l'université de Rice University et Nutaq [79] sont des exemples de plateformes permettant l'implémentation complète de formes d'onde sur FPGA.

D'autre part, les recherches sur les méthodologies de conception dédiées à la radio logicielle ont conduit aux deux grandes catégories de méthodes suivantes :

— Les **standards** [80] [81] ont pour objectif d'unifier l'approche de conception. Ils peuvent être présentés comme un ensemble d'intergicielles faisant office d'interface entre le matériel et le logiciel. Le principal exemple est Software Communication Architecture (SCA) [80] développé par l'US army pour instancier les multiples formes d'ondes utilisées.

— Les **langages** qui ont pour objectif principal de faciliter la spécification et la programmation de couches physiques sur des radios logicielles. De nombreux exemples cherchent à abstraire le matériel comme Waveform Description Language (WDL) [82], SPEX [83], DiplodocusDF [84], Platform and Hardware Abstraction Layer (P-HAL) [85] et SDRPHY [86], ils ciblent principalement des platformes à base de GPP ou DSP. Cependant, le framework GNU Radio [87] est aujourd'hui le langage le plus utilisé dans la recherche académique et dans l'industrie.

Pour les plateformes basées sur des FPGA, les travaux se concentrent principalement sur l'amélioration des capacités de traitement et sur la reconfiguration matérielle des formes d'ondes [88, 89]. La programmabilité de ce type de plateforme est peu étudiée, la plupart des solutions utilisant des langages de description bas niveau.

5.1.2 Méthodes de conception

Les méthodologies d'implémentation de systèmes numériques, souvent appelées flot de conception, ont motivé de nombreux travaux de recherche. Le niveau d'abstraction a ainsi été augmenté du transistor (description bas-niveau) jusqu'à une approche système (description haut-niveau) en passant par au niveau des portes logiques et de la fonction. Chacune de ces avancées a été accueillie avec beaucoup de scepticisme quant à la performance réalisable, même si leur objectif principal était de permettre aux concepteurs de se concentrer davantage sur la fonctionnalité plutôt que sur sa mise en œuvre. Ainsi, plusieurs étapes du processus de conception ont été automatisées grâce à des algorithmes adaptés associés à des mécanismes de génération de code source. Ces processus d'automatisation sont généralement réalisés par un compilateur dédié.

Les méthodologies de conception peuvent être divisées en deux catégories. D'une part, certains outils permettent le développement d'applications numériques sur des technologies logicielles telles que les microprocesseurs ou les DSP. Leur point d'entrée consiste en une description haut-niveau de l'application prévue, écrite dans un langage tel que C/C++. Cette description est ensuite compilée dans un langage machine entièrement composé de nombres (bits) interprétable par le processeur cible. D'autre part, certains outils permettent le développement de systèmes numériques sur des technologies matérielles telles que les FPGA ou les ASIC. Ces technologies permettent au concepteur de décrire des architectures de circuits personnalisées adaptées aux besoins de l'application. Les FPGA offrent plus de flexibilité par rapport à leurs homologues ASIC. Les FPGA permettent en effet de décrire une architecture de circuit spécifique via un fichier de configuration et peuvent être reprogrammés pour implémenter une autre fonctionnalité alors que les ASIC sont des circuits entièrement câblés. Leurs outils de conception considèrent généralement comme point d'entrée une description de l'application sous la forme d'un HDL (Hardware Description Language) tel que le VHDL (VHIC Hardware Description Language) ou le Verilog.

Une approche émergente consiste à élever le niveau d'abstraction tout en ciblant des technologies matérielles FPGA ou ASIC. Appelée synthèse de haut niveau (HLS - High Level Synthesis), elle utilise des langages de haut niveau comme point d'entrée de flots de conception FPGA et ASIC. Ainsi, une application peut être spécifiée en C/C++ par exemple, puis directement compilée en un fichier de configuration pour les FPGA ou ASIC. L'approche permet aux concepteurs de se concentrer sur les fonctionnalités de l'application plutôt que sur le matériel sous-jacent. La HLS s'appuie sur un ensemble d'outils académiques et industriels tels que *Catapult* de Calypto et *Vivado HLS* de Xilinx. Les outils HLS peuvent donc être utilisés pour le prototypage rapide lorsque des cibles matérielles sont considérées dans le processus de conception de système numérique. Cependant, ils mettent davantage l'accent sur la conception du chemin de données plutôt que sur le chemin de

contrôle. Par conséquent, le contrôle des blocs de traitement doit souvent être traité séparément. En effet, la HLS a d'abord été pensé comme un générateur de processeurs utilisant des langages séquentiels natifs. Les compilateurs associés permettent d'extraire efficacement du parallélisme dans les applications spécifiées afin de créer un chemin de données performant. Le chemin de contrôle est géré au niveau de la fonction plutôt qu'au niveau du système. Par exemple, les outils de HLS ne traitent pas correctement la spécification d'une machine à états, qui constitue souvent la structure principale pour définir la logique de contrôle d'une application au niveau système.

Dans le contexte de la radio logicielle, les outils de conception doivent prendre en charge la spécification ainsi que l'implémentation de toutes les formes d'ondes, tout en étant indépendants du matériel, ce qui pourrait assurer la portabilité de la solution sur différentes plates-formes. De plus, le point d'entrée des outils doit être un langage de haut niveau afin d'assurer la programmabilité et, enfin, ces outils doivent traiter la reconfigurabilité de l'application à un niveau d'abstraction élevé.

5.1.3 La synthèse de haut-niveau

La synthèse de haut niveau peut être définie comme étant le processus permettant de générer automatiquement une description RTL (Register Transfer Level) d'une application à partir de sa spécification à haut niveau. Elle entend ainsi réduire le fossé existant entre les équipes chargées de spécifier les algorithmes et celles chargées de leur implémentation. Les spécifications de haut niveau sont réalisées dans des langages tel que le C/C++, le SystemC ou Matlab. La Figure 5.2 présente un flot de conception HLS qui commence par une spécification haut niveau du système qui est ensuite analysée afin de produire une représentation intermédiaire sous forme d'un graphe de contrôle et de flot de données (CDFG - Control Data Flow Graph). Ce graphe peut éventuellement subir des transformations de haut niveau avant d'être utilisé pour réaliser l'allocation des ressources matérielles selon la technologie choisie par l'utilisateur et en utilisant des algorithmes d'ordonnancement [90, 91] minimisant la latence ou les ressources. Les contraintes spécifiées par l'utilisateur sont prises en compte dans cette étape et il s'en suit la génération du code RTL décrivant le système.

La synthèse de haut niveau présente de nombreux avantages en ce sens qu'elle permet un prototypage rapide d'une application du fait de l'utilisation de langages de haut niveau. Ainsi, il a été montré dans [92] qu'un circuit d'un million de portes nécessite environ 300k lignes de code en RTL alors que seule 40k lignes de C/C++ sont nécessaires en HLS. D'autre part, la HLS permet d'explorer assez aisément l'espace des solutions possibles pour une application donnée. En effet, elle propose un ensemble d'optimisation permettant d'aboutir à différentes solutions sans avoir à

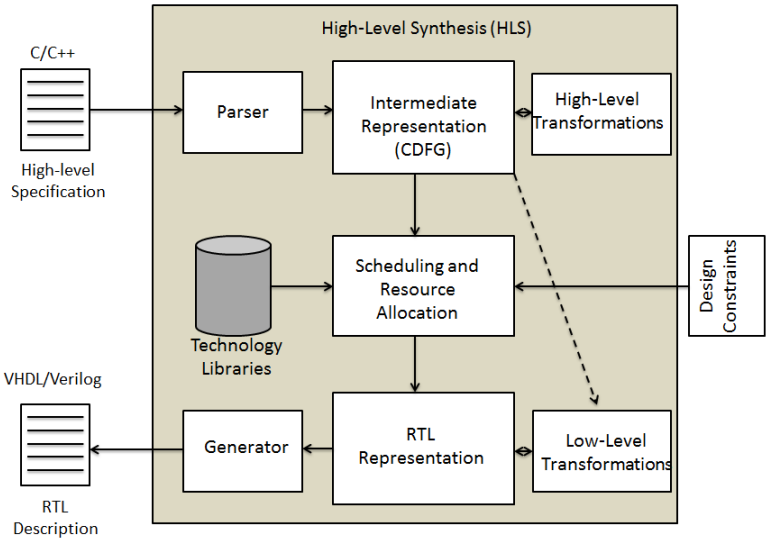


FIGURE 5.2 – Flot de conception HLS.

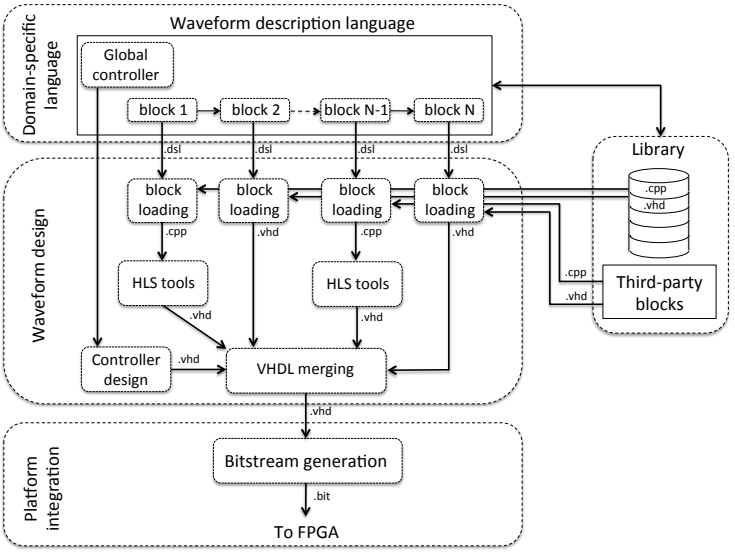


FIGURE 5.3 – Flot de conception de la radio logicielle sur FPGA.

rééditer le code source. Ces optimisations peuvent entre autres être faites sur des structures de boucles (e.g. le pipeline ou le déroulage de boucle) ou encore sur des accès mémoires.

5.2 Description des formes d'ondes

L'approche de conception proposée dans ces travaux est illustrée par la Figure 5.3. L'idée de base est de s'appuyer sur un langage de description spécifique (DSL - Domain Specific Language) afin de générer automatiquement (*Waveform Compiler*) des formes d'ondes variées et de les intégrer (*Platform integration*) sur une plateforme matérielle à base de FPGA. Le langage repose sur une bibliothèque d'IP (Intellectual Properties) de traitement du signal et permet de générer différentes architectures relatives à des standards de télécommunications. La partie 5.2.1 décrit le DSL proposé alors que le flot de conception est détaillé dans la partie 5.2.2.

5.2.1 Langage de description spécifique (DSL)

Un DSL est un langage qui à travers son expressivité ou sa syntaxe demeure spécifique à un domaine d'application [93]. C'est un outil qui simplifie la description d'application et qui est peu utilisé dans le domaine des télécommunications. Par ailleurs, certains travaux de recherches utilisent cette approche DSL afin de modéliser le domaine d'application considéré, par exemple pour modéliser des applications communicantes car-to-X dans [94]. On distingue 2 types de DSL [93] : interne ou externe. Les DSL internes réfèrent à des langages fortement dépendant d'un langage hôte. Ils sont en effet imbriqués dans ce dernier et permettent de spécialiser une partie du traitement à effectuer à partir d'instructions spécifiques. Ce type de DSL bénéficie en général de l'infrastructure de compilation du langage hôte mais ne s'avère pas être le meilleur candidat lorsqu'il s'agit de modéliser entièrement un domaine d'application. A l'opposé, les DSL externes se définissent comme des langages entièrement autonomes et dédiés à la modélisation d'un domaine d'application. Ils proposent une forme de spécialisation du modèle à travers une syntaxe et des outils de compilation adéquats. Cette approche permet aux concepteurs du langage de prendre différentes initiatives selon la finalité désirée sans se soucier d'éventuelles incompatibilités avec un langage hôte. Pour finir, un aspect notable d'une conception à base de DSL est de faciliter la communication entre experts du domaine et de ce fait de favoriser une prise en main assez rapide de l'environnement de programmation. Ces différents arguments nous ont conduit à proposer un DSL dédié au prototypage d'application radio logicielle sur cible FPGA.

Dans le contexte de la radio logicielle, la définition d'un DSL peut se montrer utile à la description d'une application pour une cible FPGA en ce sens qu'il permettra d'insérer à plus haut

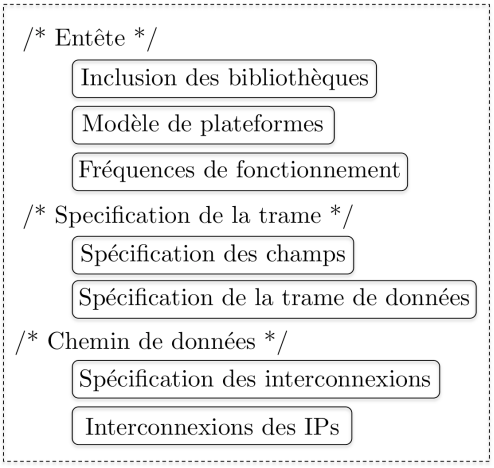


FIGURE 5.4 – Structure de la description d'une forme d'onde avec le DSL proposé.

niveau des informations relatives à la cible matérielle. Ces informations couvrent un panel assez large allant de l'ensemble des fréquences de fonctionnement dans le cas des systèmes multi-cadences aux informations propres à la plateforme d'exécution (ressources disponibles). Les différents scénarios de reconfiguration peuvent aussi être gérés à ce niveau de description, ce qui représente une indéniable plus value en terme de temps de prototypage. Une caractéristique supplémentaire du DSL proposé est la possible interaction avec les outils de synthèse de haut niveau qui permettent de générer l'architecture de certains blocs de traitement à partir de leur spécification. Cette association vise à conférer un caractère hétérogène (i.e. avec différents niveaux de description) au flot de conception tout en garantissant d'éventuelles extensions de ce dernier. Le langage consiste à définir un ensemble de mots clés qui seront ensuite interprétés par le compilateur.

La structure typique d'une description avec le DSL proposé est donnée dans la Figure 5.4. Elle est composée d'une description du modèle de plateforme utilisée, d'une description de la trame de données et d'une description de flot de données. Cette approche innovante utilise la description de la structure de flot de données ainsi que le modèle de la trame de données pour inférer automatiquement de l'unité de contrôle en VHDL. Comme discuté dans la partie 5.1.2, le contrôle des blocs de traitement doit souvent être traité séparément et le DSL permet de pallier ce problème. Ainsi, la spécification d'une forme d'onde à partir du langage que nous proposons se fait selon les étapes suivantes, une partie de ces étapes sera illustrée pour un récepteur IEEE 802.15.4 dans la Section 5.3.1 :

→ **Modèle de plateforme** : Une description de la plateforme FPGA ciblée doit être fournie avec le type de FPGA utilisé et les caractéristiques des convertisseurs de signaux de la plateforme (i.e. fréquence d'échantillonnage, nombre de bit). Ces informations sont ensuite utilisées par le compilateur afin de produire, avec les outils HLS, une description RTL appropriée.

→ **Bibliothèques et fréquences de fonctionnement** : Cette étape consiste d'une part en l'inclusion des bibliothèques d'IP utiles à la synthèse de la forme d'onde et d'autre part en la déclaration des fréquences de fonctionnement des différentes parties du système. Les mots clés prévus à cet effet sont *#include* pour l'inclusion d'une bibliothèque et *frequency* pour la spécification d'une fréquence. Deux types de fréquence peuvent être définies à savoir des fréquences de référence ou des fréquences dérivées.

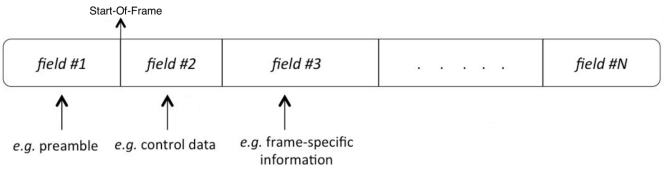


FIGURE 5.5 – Structure générique d'une trame de données.

→ **Champs et trame de données** : Chacun des champs d'une trame de donnée peut être spécifié avec différents arguments référant à sa taille ou encore à la présence éventuelle d'une redondance de donnée. La nature constante ou variable d'un champs est précisée en début de spécification à partir des mots-clés *#fieldC* et *#fieldV* respectivement. Cette spécification permet entre autres de réaliser des optimisations matérielles consistant par exemple à mapper les données des champs constants en mémoire puis à construire la trame finale par un mécanisme de multiplexage temporel des données. La trame de donnée est ensuite construite en renseignant de façon chronologique les différents champs de cette trame.

→ **Le Start-Of-Frame** : Le Start-Of-Frame désigne un champ spécifique de la trame et est utilisé comme événement de synchronisation en réception. En effet, l'un des objectifs du DSL proposé est de générer automatiquement la logique contrôle des architectures spécifiées. Pour ce faire, un point de synchronisation doit être introduit lors de la spécification de la trame. La Figure 5.5 donne un exemple générique de structure de trame avec l'indication du Start-Of-Frame.

→ **Déclaration d'IP** : Dans un but d'hétérogénéité de l'approche, la déclaration des blocs de traitement se fait précédée d'annotation renvoyant à un outil de synthèse de haut niveau. Les blocs décrits en RTL sont aussi pris en considération ce qui permet de décrire des systèmes mixtes en termes de langage. Ces annotations sont entre autres *#CatapultC* or *#RTL* qui désignent l'outil HLS de Calypto ou la description VHDL d'un bloc.

→ **Déclaration d'interconnexion** : L'objectif du DSL est de faciliter le prototypage de forme d'ondes radio logicielle sur FPGA. Ces formes d'ondes sont de types flots de données et requièrent une structuration particulière de l'interfaçage des différents blocs qui les composent. Un exemple du flot de données d'une telle forme d'ondes est donné dans la Figure 5.6. On y voit des interconnexions implémentées sous forme de FIFO (First In First Out), ce qui est une approche classique pour des applications flots de données. Ces FIFO permettent de réduire les contraintes sur les blocs de calculs en leur permettant de fonctionner à un rythme spécifique. La primitive *framedata* génère une interconnexion de type FIFO et, pour les données véhiculées dans l'architecture sans contrainte de débit, la primitive *data* permet de les implémenter sous formes de bus de données. Les outils de synthèse de haut niveau permettent de dimensionner des FIFO à la volée et nous tirons profit dans le DSL de cette caractéristique.

→ **Instanciation d'IP** : Après la déclaration des interconnexion et des IP, ces IP sont instanciés selon la forme d'ondes à définir et leurs ports sont connectés aux différentes interconnexions déclarées. Il est possible lors de cette instanciation de préciser quels champs de la trame seront traités par le bloc instancié. Cette astuce permet d'aiguiller la trame de façon optimisée à travers le chemin de donnée et est réalisée avec des signaux de contrôle de type *enable*.

Le langage ainsi décrit a donc pour tâche principale d'assembler différentes IP (ou *block*) tout en générant une logique de contrôle ainsi qu'une infrastructure de communication adéquate. Aussi, l'origine de ces blocs de base se veut diverse. Le développement récent ainsi que la mise sur le marché d'outils de synthèse de haut niveau (*HLS tools*) permet de considérer dans notre approche des IP décrites en *C/C++* ou *SystemC*, langages présentant une certaine abstraction en comparaison au *VHDL* ou au *Verilog* dédiés à l'architecture matérielle. Le caractère hétérogène ainsi conféré à la bibliothèque en cours de développement vise à élargir le spectre de couches physiques mais aussi à bénéficier de travaux antérieurs. L'hétérogénéité de la librairie ainsi souhaitée permet de considérer des codes sources de natures différentes. L'assemblage se faisant à plus bas niveau i.e. au niveau *VHDL*, il semble assez aisé d'intégrer une IP décrite dans un langage de plus haut niveau tel que *C/C++* sous réserve de l'existence d'un compilateur générant le code RTL fonctionnel associé. La démarche ainsi décrite offre des possibilités d'extension du flot de conception.

5.2.2 Flot de conception

Dans la partie précédente, nous avons détaillé les différentes étapes de description d'une forme d'ondes par le biais du DSL proposé. Cette spécification est ensuite interprétée selon le flot proposé en Figure 5.3 qui illustre le mécanisme sous-jacent au DSL. Ce flot se décline en trois étapes essentielles. La première est une étape de spécification de la forme d'ondes à partir d'une syntaxe

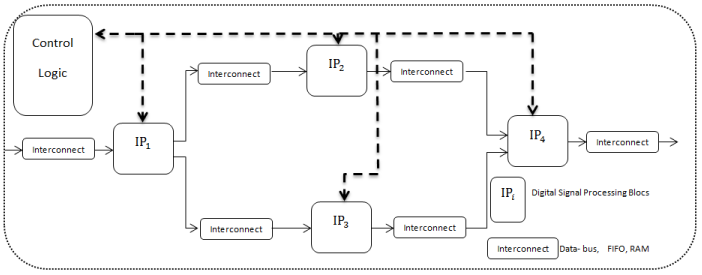


FIGURE 5.6 – Structure d'interconnexions et de contrôle des IP.

DSL. Ensuite, survient une étape de synthèse des blocs de traitements via les outils de synthèse de haut niveau et de leur assemblage. Au cours de cette étape, la description de la trame de donnée fournie dans la spécification DSL est utilisée afin de générer automatiquement la logique de contrôle du chemin de donnée. Pour finir, une étape d'intégration sur la plateforme cible est réalisée.

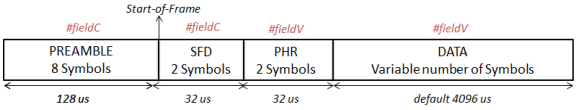
Ainsi, une des étapes essentielles de ce flot de conception est la phase d'assemblage des blocs de traitement qui consiste en l'interconnexion des différents blocs issus de la bibliothèque afin de produire la forme d'onde souhaitée. Cette étape est illustrée dans la Figure 5.3 comme *VHDL merging*. Il s'agit d'une part d'utiliser les outils HLS afin de générer les descriptions RTL des blocs instanciés ou de directement charger les blocs lorsqu'ils sont nativement décrits au niveau RTL. D'autre part, il faut aussi s'assurer de la cohérence des types de données à l'entrée et à la sortie de chaque bloc et de fournir entre autres la structure d'interconnexion (cf. Figure 5.6) selon que certaines contraintes de débits et/ou de latences soient prises en considération. On peut, comme mentionné plus haut, dégager différents profils d'interconnexion qui vont du simple bus de données à la mémoire RAM (Random Access Memory) en passant par des FIFO. Cette étape s'accompagne en plus de la génération d'une logique de contrôle bidirectionnelle dont la tâche est de fournir et recevoir un ensemble de signaux (*enable*, *clock*, ...) utile au séquencement de ces blocs. Il incombe au compilateur prévu à cet effet d'exécuter correctement cet ensemble de tâches.

Dans ce but, le compilateur analyse la description DSL afin d'en déduire une représentation intermédiaire sous la forme d'un arbre abstrait (AST - Abstract Syntax Tree). Le compilateur extrait de la représentation les sous-graphes modélisant la plateforme, la trame de données ainsi que la structure flot de données de la forme d'onde. Pour chacun des blocs instanciés, le compilateur génère un script de synthèse .tcl destiné à guider la synthèse de ce bloc par les outils de HLS. Aussi, le compilateur génère un fichier .vhd de signaux de contrôle afin que l'ensemble des blocs synthétisés soit assemblé de sorte à constituer la structure flot de données préalablement définie dans le DSL. Le détail du compilateur ainsi que l'algorithme de génération des signaux de contrôle sont donnés dans [ouedraogo14jasp].

Ce flot, en rappel, s'adresse à des plateformes radio logicielle à base de carte FPGA. Cette technologie relativement récente offre de nombreuses perspectives en termes de flexibilité et des capacités de calcul fortement intéressant pour des applications de type radio logicielle. Les interfaces de programmation de telles cartes sont dans la plupart des cas éditées par les concepteurs, ce qui contraint à l'utilisation d'outils spécifiques dans le processus d'intégration. A ce jour, le flot que nous proposons ne réalise pas une intégration automatique des solutions radios logicielles sur les cartes FPGA, nous utilisons donc comme point d'entrée les différents outils de génération de bitstream de chez Xilinx (EDK, XPS, ...).

5.3 Prototypage rapide : expérimentation et validation

Dans un souci de validation, le flot de conception a été expérimenté à travers la modélisation et l'implémentation de deux applications, à savoir les couches physiques IEEE 802.15.4 et IEEE 802.11a. Ces standards spécifient les couches physiques des technologies ZigBee et Wifi respectivement. Nous disposons à cet effet d'une plateforme matérielle Nutaq Perseus 6010. Elle est équipée d'un FPGA Virtex-6 de chez Xilinx, programmable à partir du logiciel XPS. De plus, la plateforme dispose de convertisseur avec une fréquence d'échantillonnage réglable pouvant atteindre les



(a) Structure de la trame IEEE 802.15.4 et type de champs associés.

```
/* Specification of the preamble field */
#fieldC PREAMBLE {          /* Preamble field declaration */
    constant zero;          /* Preamble constant data */
    redundancy 8;           /* Redundancy within the preamble */
    duration 128 us;        /* Preamble duration */
}
/* Specification of the SFD field */
#fieldC SFD {               /* SFD field declaration */
    constant sfdfield;      /* SFD constant data */
    duration 32 us;         /* SFD duration specification */
}
/* Specification of the PHR field */
#fieldV PHR {               /* PHR field declaration */
    data phrvalue;          /* PHR field content data */
    size 8 bits;            /* PHR size or length */
}
/* Specification of the DATA field */
#fieldV DATA {              /* DATA field declaration */
    data datasample;        /* DATA field content data */
    maxsize 128 bytes;      /* DATA field maximal length */
    minsize 4 bytes;        /* DATA field minimal length */
    duration 4096 us;       /* DATA field default duration */
}
/* Data frame specification */
 complex frame PPDU {       /* Complex frame declaration */
    PREAMBLE SFD PHR DATA   /* All fields composing the frame */
}  sof after PREAMBLE       /* Start-of-Frame designation */
```

(b) Description DSL de la trame IEEE 802.15.4 .

FIGURE 5.7 – Description de la trame IEEE 802.15.4 avec le DSL.

80 MHz. Ces convertisseurs ont une précision de donnée de 12 bits. L'interprétation des résultats est faite par le biais d'équipements d'analyse temps-réel des signaux tel que des oscilloscopes, des analyseurs de spectre ou encore des outils d'analyse de signaux en bande de base tel que Chipscope de chez Xilinx.

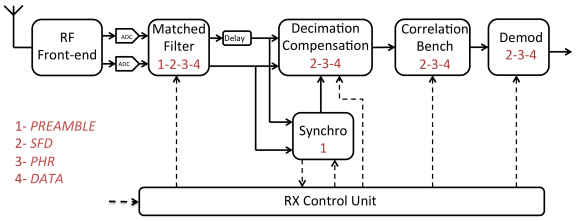
Dans cette partie, nous validons différentes étapes de notre étude en implémentant la couche physique IEEE 802.15.4 d'une forme d'ondes Zigbee avec le DSL. La première phase est la description de la structure de la trame et du graphe flot de données à partir du DSL, alors que la deuxième permettra de valider l'utilisation de la synthèse de haut niveau sur une plateforme matérielle. Enfin, une exploration architecturale sera présentée afin de valider la souplesse de conception de l'architecture.

5.3.1 Description DSL de la norme IEEE 802.15.4

A titre d'exemple, cette section présente une partie de description DSL associé à la norme IEEE 802.15.4, à savoir la description de la trame et la description du flot de données du récepteur. La description complète de cette norme est donnée dans [ouedraogo14crowncom, 95].

Le point d'entrée du flot de conception est la spécification de la structure d'une trame. La trame IEEE 802.15.4 est présentée sur la Figure 5.7a alors que le description DSL associée à cette trame est donnée dans la Figure 5.7b. Elle comprend deux champs de synchronisation SHR, un champ de données spécifique à la trame PHR et un champ de données utiles PSDU. Dans cette trame, le champ SHR composé du préambule et du SFD est constant (de type *#fieldC*). Le préambule possède une structure répétitive (composée de huit symboles 0) ce qui n'est pas le cas du SFD. Les champs PHR et PSDU sont des champs variables (de type *#fieldV*) dont les données seront issues de blocs de traitement du signal. Le DSL proposé exploite ces différentes caractéristiques afin d'optimiser la structuration de la trame en sauvegardant les données des champs constants dans des mémoires, données qui seront ensuite multiplexées en temps aux données des champs variables issues de blocs de traitement.

Le graphe flot de données associé à la norme IEEE 802.15.4 implémente une technique d'accès au canal du type étalement de spectre par séquence directe (DSSS - Direct Sequence Spread Spectrum). L'architecture du récepteur est illustrée par la Figure 5.8a. Elle est composée de cinq blocs de traitement. Après filtrage de l'ensemble de la trame, le bloc de synchronisation utilise uniquement le champ PREAMBLE. Après une étape de décimation, les symboles sont décodés par un banc de



(a) Architecture du récepteur IEEE 802.15.4 et champs associés à chaque bloc de traitement.

```
/* RxFilter block specification */
rxfir_i: ip rxfir processing all PPDU from ADC {
    read from_adcI on port ich2 at fe;
    read from_adcQ on port qch2 at fe;
    write ich_rxfilter on port ich3 at fe;
    write qch_rxfilter on port qch3 at fe;
    synthesis     catapult;
    constraint    throughput 2;
}
/* Synchronization block specification */
synchro_i: ip synchro processing PREAMBLE from rxfir_i {
    read ich_rxfilter on port ich4 at fe;
    read qch_rxfilter on port qch4 at fe;
    write topt_connect on port topt;
    write phiI_connect on port phiI;
    write phiQ_connect on port phiQ;
    write detect_connect on port pbdetect;
    synthesis     catapult;
    constraint    latency PBDETECTDL;
}
/* Compensation block specification */
decimadjust_i: ip decimadjust processing SFD PHR DATA from
rxfir_i {
    read ich_rxfilter on port ich5 at fe;
    read qch_rxfilter on port qch5 at fe;
    read topt_connect on port topt;
    read phiI_connect on port phiI;
    read phiQ_connect on port phiQ;
    read detect_connect on port pbdetect;
    write ich_decim on port ich6 at fc;
    write qch_decim on port qch6 at fc;
    constraint    throughput 2;
}
/* CorrBench block specification */
corrbench_i: ip corrbench processing SFD PHR DATA from
decimadjust_i {
    read ich_decim on port ich7 at fc;
    read qch_decim on port qch7 at fc;
    write sb_connect on port sb_rx at fs;
    synthesis     catapult;
}
```

(b) Description DSL du récepteur IEEE 802.15.4.

FIGURE 5.8 – Implémentation d'un récepteur IEEE 802.15.4 avec le DSL.

corrélateurs. Le dernier bloc converti les symboles en bits. La description DSL de ce graphe flot de données est donnée dans la Figure 5.8b.

5.3.2 Implémentation de la norme IEEE 802.15.4

L'intégration a été faite sur une plateforme Nutaq Perseus 6010 équipée d'un FPGA Virtex-6 et d'une tête d'émission radio large bande et agile en fréquence. Dans ce contexte radio logicielle, les travaux que nous avons effectués sur cette plateforme ont fait l'objet d'une publication [bhatnagar13vtc]. La forme d'onde IEEE 802.15.4 spécifiée à partir du DSL a été générée automatiquement, cette spécification s'est faite selon les étapes décrite dans la Figure 5.3.

Les résultats de synthèse sur FPGA sont présentés dans la Table 5.1 pour l'émetteur et le récepteur. On y compare deux approches d'implémentation : l'une se basant sur une description RTL de la forme d'onde (HC-VHDL pour Hand-Coded VHDL) et l'autre décrite à partir du DSL proposé (HLS-VHDL). Les 2 synthèses ont été contraintes à ne pas utiliser de DSP slices. Si la taille de l'émetteur est relativement petite, les ressources obtenues pour les récepteurs montrent les bonnes performances étonnamment obtenues, pour cet exemple, par la synthèse de haut niveaux.

La Figure 5.9 présente la validation expérimentale des implémentations de l'émetteur et du récepteur IEEE 802.14.5. Une photographie du banc de test expérimental est donnée en Figure 5.9a, on y voit les plateformes Nutaq (émetteur et récepteur) permettant de faire des mesures de taux d'erreur binaire. La Figure 5.9b montre les signaux IEEE 802.15.4 décodés par le biais de l'équipement Vector Signal Analysis (VSA) de chez Agilent. Elle donne le spectre reçu, les signaux bande

	Emetteur				Récepteur			
	Slices	FF	LUT	BRAM	Slices	FF	LUT	BRAM
HC-VHDL	35	134	86	0	3668	4292	10945	0
HLS-VHDL	76	188	202	0	543	1690	1058	0

TABLE 5.1 – Estimation des ressources de l'émetteur et du récepteur IEEE 802.15.4.

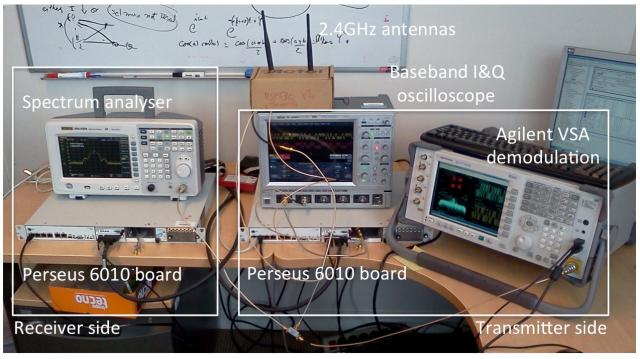
de base temporels ainsi que la constellation après décodage.

5.3.3 Exploration de l'espace de conception

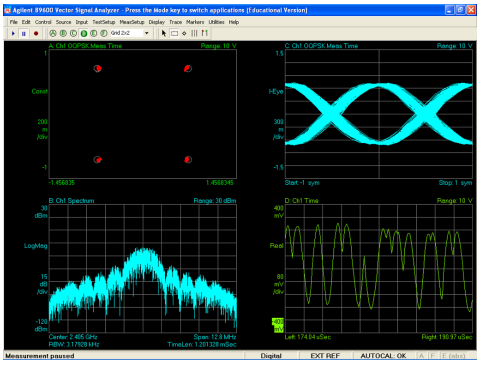
Le DSL proposé permet d'exploiter pleinement les fonctionnalités de DSE (Design Space Exploration) offertes par les outils HLS. L'exploration vise essentiellement à répondre à des spécifications de performance en terme de ressources utilisées ou de temps de calcul. Elle représente une valeur ajoutée importante des outils HLS en permettant d'explorer rapidement différentes solutions d'une même description. Une optimisation des ressources a été donc effectuée à partir de la description haut niveau des formes d'onde IEEE 802.15.4 et IEEE 802.11a.

Dans la Figure 5.10a, une exploration DSE est effectuée sur le bloc *CorrBench*, qui fait partie du récepteur IEEE 802.15.4 (cf Figure 5.8a), afin de faire un compromis entre la surface et la latence du bloc (correspondant au nombre de cycles pour réaliser le traitement). Ces résultats sont les estimations de surface données pour l'outil Catapult. Chaque courbe de cette figure correspond à un niveau de pipeline d'une boucle interne. Le pipeline d'une boucle permet de réduire sa latence globale en lançant l'itération n+1 avant la fin de l'itération n. Le nombre de cycles entre les itérations de la boucle s'appelle l'intervalle d'initiation (II - Initiation Interval). Ainsi, plus l'intervalle d'initiation (II) est faible, plus le pipeline sera profond. On peut voir à partir de ces courbes que les faibles latences sont obtenues pour un II faible au prix d'une augmentation des ressources utilisées. Le déroulage de boucle (U) impacte aussi considérablement la latence.

De la même façon, la Figure 5.10b montre une exploration DSE exécutée pour un bloc FFT (Fast Fourier Transform) de 256 points avec Catapult afin de comparer différentes solutions pour un même bloc de traitement FFT. L'objectif était de trouver un compromis entre le débit réalisable et la surface totale. Dans ce but, le compilateur DSL génère un script .tcl pour chaque valeur de pipeline II et extrait ensuite l'estimation des ressources à partir des rapports de synthèse publiés par les outils HLS. L'exploration du bloc FFT montre l'augmentation des ressources en fonction du débit réalisable. Une solution adaptée est automatiquement sélectionnée de manière à créer la forme d'onde souhaitée. Ce scénario d'exploration de la solution peut être exécuté à plusieurs reprises pour chaque bloc de traitement avant l'assemblage final du signal. Ainsi, le concepteur sélectionne les solutions appropriées qui répondent aux spécifications globales de l'application.

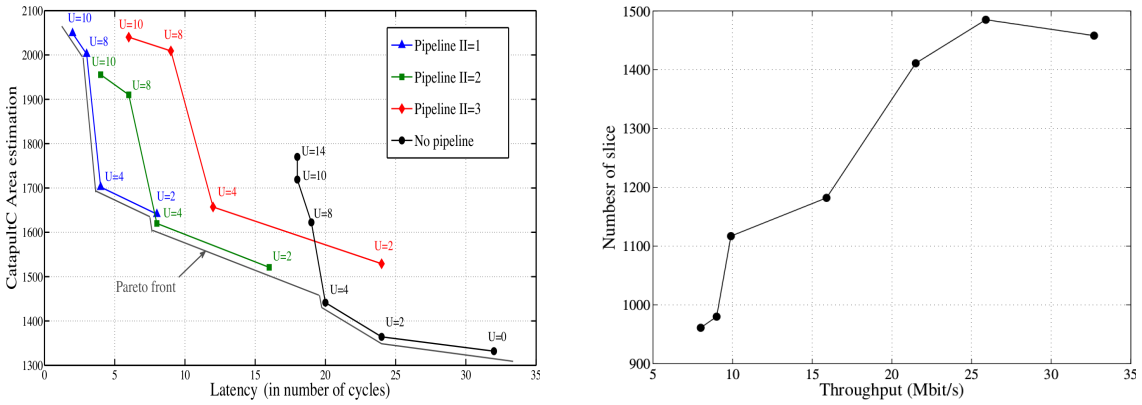


(a) Banc de test expérimental.



(b) Démodulation IEEE 802.15.4 par les équipements VSA.

FIGURE 5.9 – Validation expérimentale de l'implémentation IEEE 802.15.4 sur les plateformes Nutaq Perseus 6010.



(a) Exploration de l'espace de conception du bloc *CorrBench* du récepteur IEEE 802.15.4.

(b) Estimation des ressources en fonction du débit pour un récepteur IEEE 802.11a.

FIGURE 5.10 – Exploration de l'espace de conception.

5.4 Autres contributions

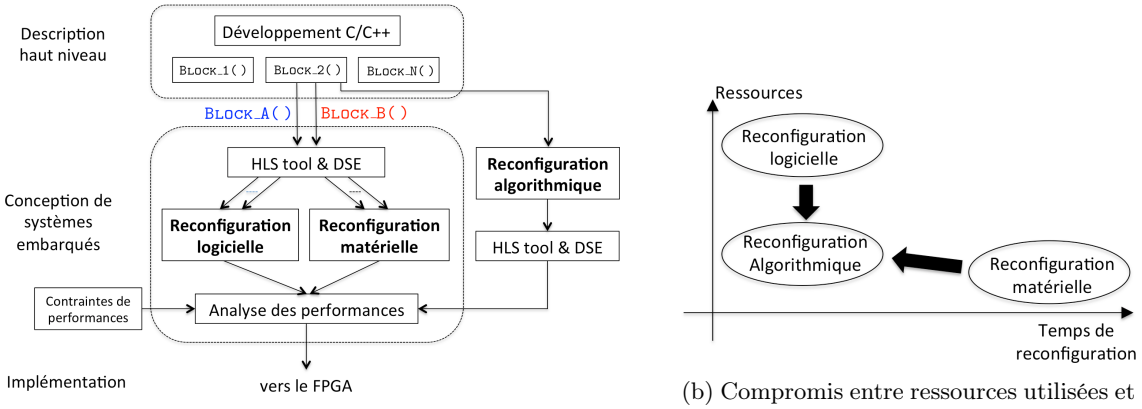
5.4.1 Reconfiguration dynamique des formes d'ondes

Dans ce chapitre, nous avons précédemment montré la capacité de prototypage rapide d'une radio logicielle sur une cible FPGA obtenue en exploitant les principes et les outils de HLS pour générer des descriptions RTL à partir de spécifications haut niveau [ouedraogo14crowncom]. Le prototypage rapide ainsi obtenu offre une flexibilité à la conception des architectures au travers des outils de DSE. Cependant, le problème de la flexibilité temps-réel est toujours ouvert. Présentée dans [tran16crowncom, tran16dasip], cette étude traite donc de l'implémentation sur FPGA de la reconfiguration dynamique d'une forme d'onde flexible à partir de sa description haut niveau.

Flot de conception de reconfiguration dynamique : Il existe deux types de travaux qui traitent de la flexibilité temps-réel d'une radio logicielle à base de FPGA. Les premiers proposent de concevoir des composants RTL multi-modes avec des signaux de contrôle pour basculer entre les différents modes [89, 96]. Dans notre approche, un bloc de traitement multi-mode peut être décrit en utilisant des modifications algorithmiques dédiées du bloc de traitement (*Reconfiguration algorithmique*) ou avec une génération automatique utilisant une encapsulation HLS (*Reconfiguration logicielle*). Les autres sont basés sur la reconfiguration dynamique partielle (DPR - Dynamic Partial Reconfiguration) [88]. La DPR, appelée ensuite *Reconfiguration matérielle*, permet de reconfigurer une partie du FPGA pendant que le reste du FPGA continu de fonctionner. C'est un sujet de recherche datant des années 90 [97] qui est maintenant couramment utilisé dans les FPGA, puisque Xilinx et Altera fournissent de tels circuits [98] [99]. Le principal avantage de la *Reconfiguration matérielle* est de permettre la réutilisation une zone matérielle, permettant ainsi de réduire la consommation d'énergie et les coûts de production.

L'objectif de notre flot de conception, décrit dans la Figure 5.11a, est de choisir ou de combiner ces reconfigurations tout en décrivant le bloc de traitement à un niveau de description élevé. Les différents modes d'un bloc de traitement peuvent être fournis soit par un codage manuellement ou soit en utilisant un outil HLS pour générer différentes versions d'un bloc de traitement en modifiant les contraintes de synthèse telles que le débit, la latence, la taille des données, ... Dans la Figure 5.11a, les *Contraintes de performances* sont des contraintes définies par l'utilisateur telles que les ressources/surfaces, le temps de reconfiguration, le débit ou la latence. L'*Analyse des performances* compare les performances des trois solutions aux contraintes définies par l'utilisateur.

La Figure 5.11b montre le compromis de conception entre les ressources utilisées et le temps de reconfiguration pour les différents types de reconfiguration. La *Reconfiguration algorithmique* permet de diminuer les ressources par rapport à la *Reconfiguration logicielle* et de réduire le temps de reconfiguration par rapport à la *Reconfiguration matérielle*. Elle offre donc un meilleur compromis ressources/temps de reconfiguration. Cependant, en fonction des blocs de traitement, le temps de



(a) Flot de conception de la reconfiguration dynamique.

(b) Compromis entre ressources utilisées et temps de reconfiguration.

FIGURE 5.11 – Génération d'un bloc de traitement reconfigurable dynamiquement.

description (i.e. de codage) de la *Reconfiguration algorithmique* peut être important par rapport à *Reconfiguration logicielle*.

Performances d'une FFT multi-mode pour le standard LTE : Un bloc de traitement FFT flexible est conçu en utilisant l'approche proposée. Pour la norme LTE, le composant FFT résultant doit avoir six modes correspondant à différentes tailles de FFT : {128 ; 256 ; 512 ; 1024 ; 1536 ; 2048}. L'architecture proposée et détaillée dans [tran16crowncom] repose sur deux descriptions haut-niveau de FFT. D'une part, une FFT pour les puissances de 2 `BLOCK_FFTPOW2()` est obtenue par *Reconfiguration algorithmique* afin de partager les ressources entre les différents modes. Ensuite, une fonction FFT pour la taille 1536 `BLOCK_FFT1536()` est décrite combinant trois FFT de taille 512 et une structure radix-3. Dans ce document, nous présentons l'étude du partage des ressources entre ces 2 fonctions pour la *Reconfiguration matérielle* et *Reconfiguration logicielle*.

La *Reconfiguration logicielle* est appliquée en premier pour concevoir une FFT réalisant les 6 modes LTE. Une fonction `MULTI_MODE_BLOCK_LTE()` est générée à partir des deux fonctions `BLOCK_FFTPOW2()` et `BLOCK_FFT1536()`. La *Reconfiguration matérielle* est ensuite appliquée aux deux fonctions `BLOCK_FFTPOW2()` et `BLOCK_FFT1536()`. Deux partitions séparées sont d'abord générées : une pour la FFT puissance de 2 et une pour la FFT 1536. Enfin, une partition commune est créée pour la DPR des 2 FFT.

La Table 5.2 donne les résultats de synthèse et la latence (en nombre de cycles d'horloge) pour les deux reconfigurations. Les ressources nécessaires à chaque fonction sont dans un premier temps données. Pour la *Reconfiguration logicielle*, les ressources utilisées par `MULTI_MODE_BLOCK_LTE()` sont presque la somme des ressources utilisées par `BLOCK_FFTPOW2()` et `BLOCK_FFT1536()` lorsqu'ils sont synthétisés séparément. L'outil HLS ne partage donc pas les ressources entre les deux fonctions même si elles ne sont pas exécutées en même temps. Pour la *Reconfiguration matérielle*, la partition de la FFT 1536 est plus petite que celle de la FFT puissance de 2. Ainsi, lors de la combinaison des 2 FFT en une seule partition, la partition résultante est basée sur la partition de la FFT puissance de 2.

Par rapport à la *Reconfiguration logicielle*, la FFT multimode basée sur la *Reconfiguration matérielle* utilise moins de ressources (les BRAM et DSP sont les composants logiques les plus coûteux). Lorsque la taille de la FFT doit être modifiée parmi les FFT puissance de deux, un seul cycle d'horloge est nécessaire pour la reconfiguration dans les deux cas. Toutefois, 32.9 ms sont nécessaires à la reconfiguration lors de la commutation d'une FFT 1536 vers une FFT puissance de deux (ou vice versa) avec une *Reconfiguration matérielle*, alors qu'un seul cycle est requis avec la *Reconfiguration logicielle*.

La FFT multimode utilisant la *Reconfiguration matérielle* a été intégrée dans une chaîne de transmission complète et implémentée sur une carte FPGA ML605. Cette implémentation a donné lieu à une démonstration en conférence [tran16dasip].

	Bloc de traitement : Ressources nécéssaires		**Reconfiguration matérielle** Partition : Ressources utilisées			**Reconfiguration logicielle**
	FFT puiss. 2	FFT 1536	FFT puiss. 2	FFT 1536	FFT LTE	FFT LTE
BRAM	12	14	17	14	17	26
DSP	65	40	68	56	68	103
LUT	2553	3054	4080	3360	4080	5256
FF	2497	2010	8160	6720	8160	4299
Taille bitstream	n/a	n/a	416016 Bytes	277344 Bytes	2x416016 Bytes	n/a
Temps de reconf.	n/a	n/a	32.9 ms	21.96 ms	32.9 ms	n/a
Latence	$L_{\texttt{FFTpow2}}$	52198	$L_{\texttt{FFTpow2}}$	52198	$L_{\texttt{FFTpow2}}$ & 52198	$L_{\texttt{FFTpow2}}$ & 52198

Table 5.2 – Performances de la FFT pour la norme LTE avec la *Reconfiguration matérielle* et la *Reconfiguration logicielle*. Les latences (en nombre de cycles) pour la FFT puissance de 2 sont $L_{\texttt{FFTpow2}}$={355 ;7604 ;16172 ;34284 ;72412} .

5.4.2 Adaptation dynamique d'un récepteur OFDM en virgule fixe

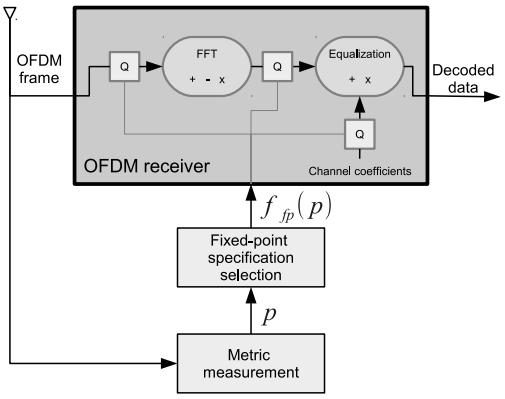
Une autre application de la génération rapide d'architecture matérielle est l'exploration de la dynamique de la représentation des nombres. Les algorithmes de traitement du signal sont souvent implémentés en virgule fixe. L'implémentation en virgule fixe (i.e. la taille des mots) est choisie pendant la phase de conception selon les conditions de fonctionnement du système. Plusieurs conditions de fonctionnement impliquent donc plusieurs implémentations en virgule fixe. Dans les systèmes statiques, le système est dimensionné pour la pire condition avec des tailles de mots importants afin d'assurer le bon fonctionnement dans toutes les circonstances. Cependant, utiliser des tailles de mot importantes n'est pas une solution optimisée du point de vue des ressources utilisées et de la consommation d'énergie. L'objectif principal de cette étude est l'évaluation de l'impact de la précision des données sur la consommation d'énergie afin de développer un récepteur OFDM en virgule fixe, pour lequel l'implémentation peut changer selon les conditions du canal de propagation. En utilisant un récepteur dynamique, une réduction de la consommation d'énergie est attendue.

Récepteur OFDM dynamique : Dans le cas d'une application spécifique utilisant un récepteur OFDM, on suppose qu'un taux d'erreur binaire (BER - Bit Error Rate) minimum est nécessaire pour garantir son bon fonctionnement. La Figure 5.12a présente le récepteur proposé : une métrique p est calculée à partir du signal reçu et une implémentation en virgule fixe est choisie en fonction de la métrique pour atteindre l'objectif en terme de BER. La réduction de l'arithmétique en virgule fixe augmentera le BER et économisera de l'énergie.

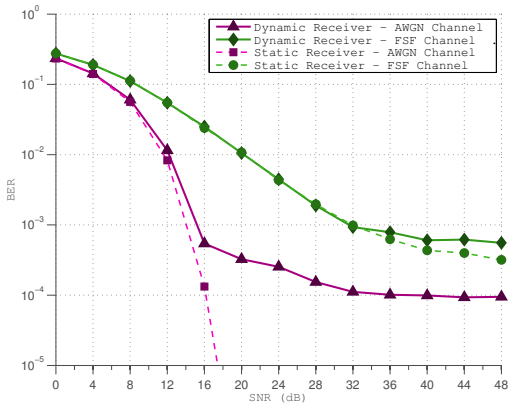
Dans cette étude, le récepteur ciblera un BER fixe de 10^{-3} et s'adaptera aux conditions du canal de propagation caractérisées par le rapport signal à bruit (SNR - Signal to Noise Ratio) en réception et le type de canal, bruit additif gaussien (AWGN - Additive White Gaussian Noise) ou sélectif en fréquence (FSF - Frequency Selective Fading). La consommation d'énergie du récepteur est calculée à l'aide de bibliothèques de consommation d'énergie d'opérateurs élémentaires pour différentes tailles de mot [cladera15isvlsi]. Ces bibliothèques sont obtenues à partir de nos mesures expérimentales exhaustives sur différentes cibles matérielles (FPGA Virtex 5 et ARM7TDMI). En connaissant le nombre d'opérateurs élémentaires du système et la taille des mots, la consommation énergétique peut être ainsi estimée.

Performance du récepteur dynamique : Le récepteur dynamique OFDM à virgule fixe est associé au sélectionneur d'implémentation présenté dans [cladera15eusipco] et des simulations ont été effectuées pour chaque condition de canal afin d'analyser son comportement. Les courbes de BER en fonction SNR sont présentées dans la Figure 5.12b. Les lignes pointillées correspondent à un récepteur statique utilisant l'approche du "pire cas", tandis que les lignes pleines correspondent au récepteur dynamique utilisant une précision adaptative. On peut observer que pour un BER $> 10^{-3}$ les deux courbes se confondent. Mais, lorsque le BER $< 10^{-3}$, notre système adaptatif essaie d'atteindre la valeur cible alors que le récepteur statique diminue continuellement son taux d'erreur.

Les estimations énergétiques montrent que cette technique permet d'économiser jusqu'à 63% de l'énergie lorsqu'un FPGA (Virtex 5) est utilisé. Avec un processeur ARM7, une économie d'énergie



(a) Récepteur OFDM en virgule fixe et algorithme d'adaptation dynamique de la précision.



(b) Courbes de taux d'erreur binaire (BER) en fonction du rapport signal à bruit (SNR).

FIGURE 5.12 – Algorithme et performance du récepteur OFDM adaptatif.

de 34% est obtenue. Une mise en œuvre ASIC permettrait probablement des économies d'énergie encore plus importantes.

5.5 Conclusion

Dans ce chapitre, nous avons présenté un flot de conception innovant dédié à l'implémentation de formes d'ondes sur des radios logicielles embarquant des cibles de calcul de type FPGA. Ce flot repose essentiellement sur un langage dédié à la radio logicielle combiné aux outils de synthèse de haut niveau. Les différentes étapes de la spécification à l'aide du langage ont été détaillées et un exemple de réalisation a été discuté. Quelques résultats d'intégration sur une plateforme matérielle sont présentés et commentés. Les perspectives directes consistent en un raffinement du DSL afin d'insérer des caractéristiques propres à la radio logicielle, à savoir les différents scénarios de reconfiguration dynamique une fois les formes d'ondes spécifiées. Une étape significative de ce projet consiste aussi en l'enrichissement de la bibliothèque de blocs de traitement.

Une extension de ces travaux vers des architectures hétérogènes, combinant ressources logicielles et matérielles, est aussi envisagée. Il convient alors de positionner le compromis entre niveau d'abstraction et performances par rapport au partitionnement des traitements entre les parties logicielles et matérielles et d'en évaluer les conséquences en termes de temps de conception. Une piste serait de combiner la synthèse de haut niveau au langage Julia [100] qui offre un niveau d'abstraction élevé avec d'excellentes performances.

Ces travaux ont aussi montré qu'une méthodologie associée au prototypage rapide ne se résout pas uniquement au choix du langage de programmation qui permet cette abstraction. Il est nécessaire de considérer le système en entier, composé de la couche applicative mais également des pilotes matériels, du système d'exploitation de la carte et de l'interface avec les parties analogiques.

Chapitre 6

Méthodes de conception pour les architectures multi-cœurs hétérogènes

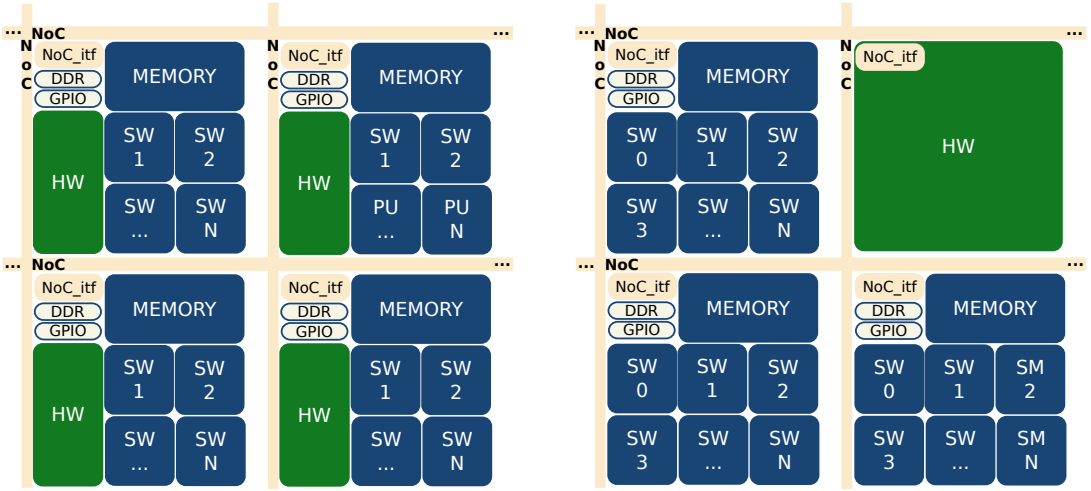
L'utilisation d'architectures reconfigurables de type FPGA permet d'augmenter significativement les capacités de traitement des radios logicielles. Cependant la question de la consommation d'énergie reste encore ouverte et ce chapitre étudie l'efficacité énergétique pouvant être atteinte par l'association d'architectures matérielles et logicielles. La flexibilité de l'architecture pouvant permettre également d'y associer des traitements adaptatifs permettant une réduction supplémentaire de la consommation.

Ainsi, le paradigme de la radio logicielle efficace en énergie est traité au travers de l'exploration des opportunités offertes par les architectures multi-cœurs hétérogènes (HMpSoC) intégrant plusieurs processeurs et des accélérateurs matériels. Outre le gain en performance attendu, ces architectures ont induit de nouvelles problématiques en termes de flot de conception. En effet, avec la multiplication du nombre et du type d'organes de calcul, l'espace d'exploration est devenue immense. Ainsi, lorsque la consommation devient un facteur clé, il est essentiel de pouvoir explorer cet espace au plus tôt dans le flot de conception pour obtenir des solutions efficaces en énergie.

Ce chapitre présente une méthode d'exploration de l'espace de conception d'accélérateurs matériels implémentés sur un processeur HMpSoC. Une vue d'ensemble du flot de conception est présentée dans la Section 6.1. Une modélisation analytique de l'énergie des HMpSoC ainsi que l'exploration sont ensuite présentées dans la Section 6.2. Enfin, la Section 6.3 donne les performances de la méthode pour deux applications.

Sommaire

6.1 Vue globale du flot de conception pour les architectures HMpSoC . 88
6.1.1 Structure des architectures et applications ciblées 88
6.1.2 Flot de conception orienté énergie 90
6.2 Exploration architecturale d'accélérateur 91
6.2.1 Applications, architectures et objectifs du DSE 92
6.2.2 Estimation de l'énergie et du temps d'exécution 94
6.2.3 Exploration de l'espace de conception 96
6.3 Résultats expérimentaux d'exploration 97
6.3.1 Environnement d'expérimentation 98
6.3.2 Extraction des paramètres 99
6.3.3 Performances de l'exploration MILP 100
6.4 Autres contributions 102
6.4.1 Plateforme d'émulation de HMpSoC 102
6.5 Conclusion 102



(a) HMpSoC distribué.

(b) HMpSoC partagé.

FIGURE 6.1 – Structure des architectures HMpSoC.

6.1 Vue globale du flot de conception pour les architectures HMpSoC

6.1.1 Structure des architectures et applications ciblées

Dans cette section, nous présentons le formalisme utilisé dans la suite du chapitre pour décrire les architectures et les applications utilisées.

Architectures cibles

Par le passé, la loi de Moore [101] prédisant le doublement du nombre de transistors sur puce tous les 18 mois, a été un moteur fondamental de l'évolution des micro-architectures. Associée à la loi de Dennard [102] prédisant la diminution de la tension et du courant d'alimentation des transistors en proportion de leur taille, une nouveau type de circuits, d'architectures et de compilateurs a émergé, conduisant ainsi à une augmentation exponentielle des performances. Ces derniers ont été obtenus en passant progressivement d'architectures simple-cœur à des architectures multi-cœurs pour tirer pleinement profit du nombre de transistors disponibles. Malheureusement, un nouveau phénomène appelé "dark silicon" lié à la limitation de la densité de puissance dissipée sur une puce [103] est apparu. Il empêche l'utilisation simultanée de l'ensemble des transistors présents sur la puce et a poussé les concepteurs à intégrer des accélérateurs matériels au sein des micro-architectures. C'est ainsi que sont nées les architectures multi-cœurs hétérogènes (HMpSoC) intégrant plusieurs processeurs et des accélérateurs matériels.

Les architectures multi-cœurs sont généralement composées d'un ensemble de mémoires, de processeurs, d'éléments d'interconnexion et d'entrées/sorties. Lorsqu'elles intègrent des accélérateurs matériels, elles deviennent hétérogènes dans le sens où elles associent des processeurs généralistes (SW) et des accélérateurs matériels (HW). Une représentation générique des architectures HMpSoC est proposée dans la Figure 6.1. Elle est construite à l'aide de clusters connectés par un réseau sur puce (NoC). Chaque cluster est composé de N cœurs de calcul logiciels associés à un accélérateur matériel de taille S. Au niveau du cluster, les communications sont effectuées grâce à une mémoire partagée. A partir de cette description, deux familles de HMpSoC peuvent être construites en fonction du placement des accélérateurs matériels dans l'architecture.

Dans le cas où ils sont placés au sein de chaque cluster ($S \neq 0$), on obtient des *HMpSoC distribués* (Figure 6.1a). Ils permettent d'obtenir des communications rapides entre la partie matérielle et la partie logicielle mais en conséquence, la taille maximale des accélérateurs est réduite. Dans le cas où les accélérateurs sont placés au niveau cluster, on obtient deux types de cluster : les clusters logiciels ($N \neq 0$ et $S = 0$) ; et les clusters matériels ($S \neq 0$ et $N = 0$). Les architectures

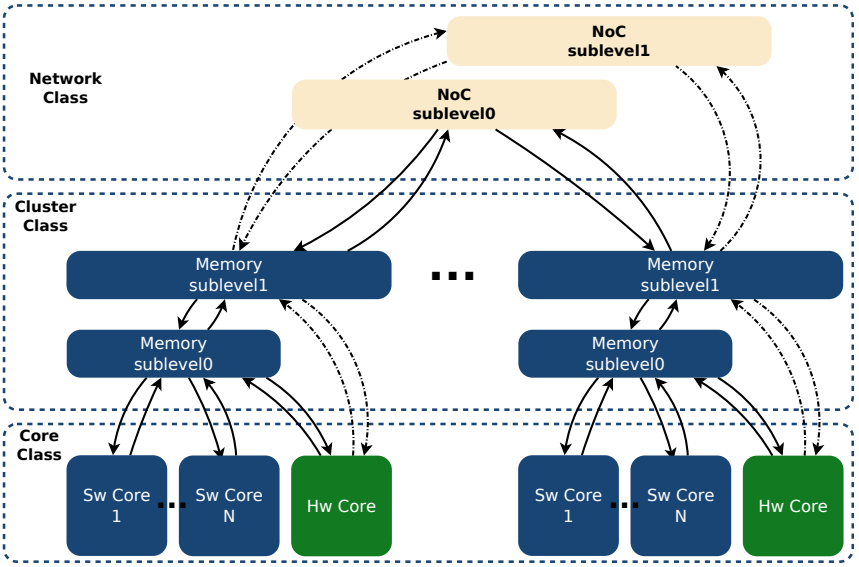


FIGURE 6.2 – Représentation d'un HMpSoC basée sur la hiérarchie mémoire.

ainsi obtenues sont appelées *HMpSoC partagés* (Figure 6.1b). Cette approche permet d'avoir des accélérateurs matériels de grande taille et de les partager entre les clusters. En contrepartie, ils induisent une augmentation des temps de communications entre les parties logicielles et matérielles avec potentiellement une congestion de ces communications.

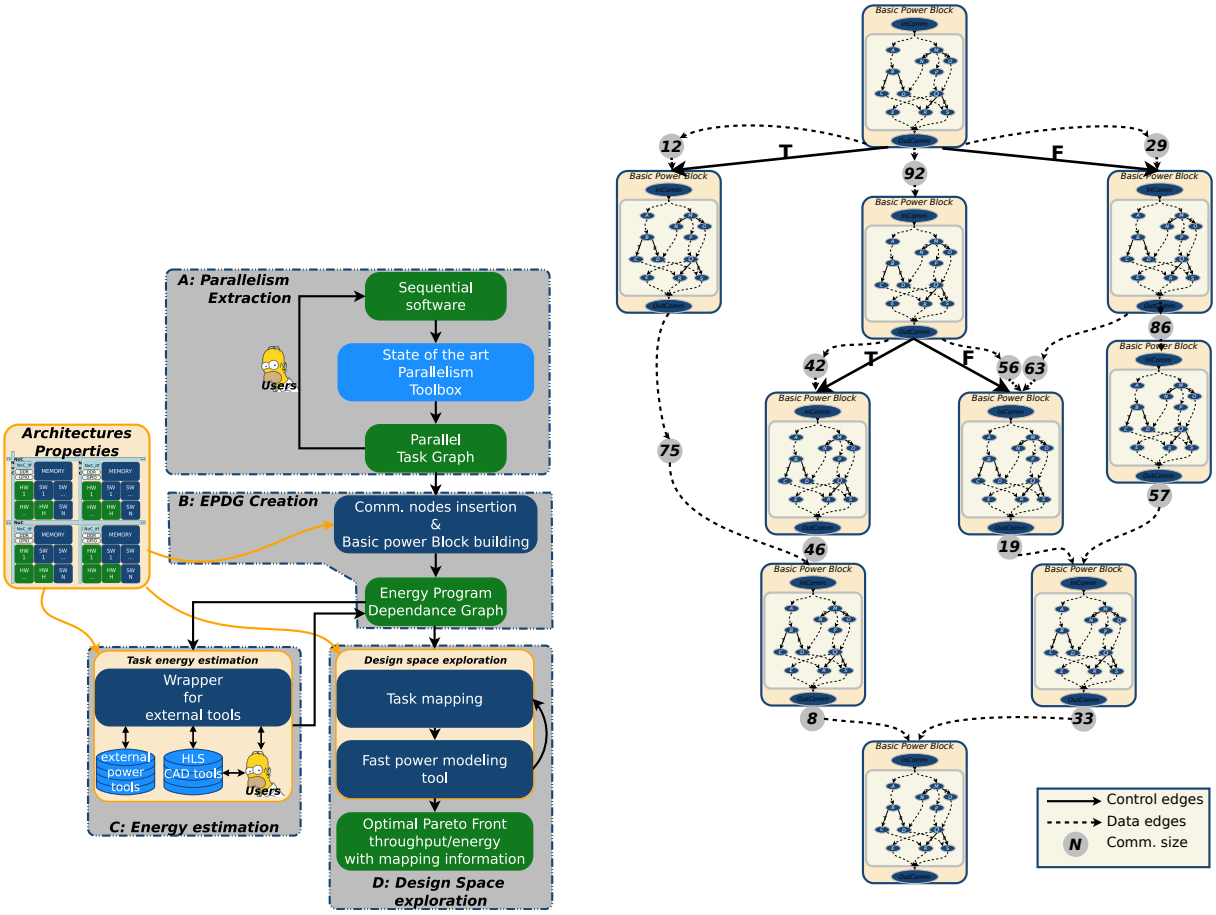
Les familles de HMpSoC introduites précédemment peuvent définir un grand nombre d'architectures. Pour cibler, avec précision, une architecture particulière au sein de ces familles, nous avons besoin d'une représentation générique. Pour cela, nous avons proposé une représentation basée sur la hiérarchie mémoire décrite à l'aide d'un arbre orienté (Figure 6.2). Cette représentation divise la hiérarchie mémoire d'un HMpSoC en trois niveaux principaux : réseau, cluster et cœur. Chaque niveau contient des sous-niveaux, avec les propriétés suivantes :

— Au niveau réseau, les sous-niveaux sont disjoints de façon à ce que chaque sous-niveau puisse être utilisé indépendamment.
— Au niveau cluster, les sous-niveaux sont mélangés, c'est-à-dire qu'ils sont liés entre eux mais restent accessibles à chaque sous-niveau.
— Au niveau cœur, les sous-niveaux sont chaînés. L'accès au sous-niveau de profondeur L passe par les $L-1$ sous-niveaux supérieurs.

Pour décrire avec précision une architecture cible, les caractéristiques de chaque classe mémoire doivent être définies, telles que la profondeur de chaque sous-niveau, le débit des canaux, ainsi que le coût des communications (énergie et temps).

Structure des applications

Après une étude des paradigmes de représentation des applications, nous formalisons la représentation qui sera utilisée dans le cadre de ce travail. Cette représentation est illustrée dans la Figure 6.3b et permet d'exposer les parallélismes ainsi que les coûts de communications. Elle est basée sur un graphe à deux niveaux appelé *Energy Program Dependence Graph* (EPDG). Le niveau supérieur expose le parallélisme gros grain de l'application et contient des annotations tels que le coût d'exécution des nœuds et la taille des communications. Le niveau inférieur, quant à lui, expose le parallélisme à grain fin. Au niveau supérieur, les nœuds ne sont pas de simples instructions, mais des blocs d'instructions appelés *Basic power Bloc* (BpB). Ces BpB représentent des macro-instructions avec un point d'entrée unique et un point de sortie unique et sont indépendants les uns des autres. Cela permet de les optimiser de manière indépendante. De plus, ces blocs sont enrichis d'informations comme la taille des communications et le coût énergétique de calcul.



(a) Vue d'ensemble du flot de conception. (b) Graphe EPDG de représentation des applications.

FIGURE 6.3 – Flot de conception orienté énergie pour les architectures HMpSoC.

6.1.2 Flot de conception orienté énergie

Dans cette section, une vue globale du flot de conception orienté énergie est présentée. Ce flot a pour objectif d'aider les concepteurs à aborder les nouvelles architectures multi-cœurs hétérogènes. L'idée principale est d'intégrer un modèle de consommation, calculable rapidement, au plus tôt dans le flot de conception. Cela permet de considérer les performances énergétiques de l'application tout en explorant rapidement l'espace des architectures possibles. L'approche proposée considère indépendamment l'impact des communications impliquées entre les organes de traitement et l'impact des calculs afin de proposer une formulation analytique qui diminue le temps d'estimation de la consommation énergétique.

La Figure 6.3a présente une vue d'ensemble du flot de conception. Ce flot est divisé en quatre parties distinctes. La première consiste à extraire le parallélisme d'une application décrite de manière séquentielle et fournit un graphe de tâches parallèles. La seconde enrichit le graphe obtenu avec la taille des communications et identifie des blocs indépendants pour former des BpB. Ces BpB sont ensuite utilisés pour construire un graphe parallèle hiérarchique avec le formalisme EPDG. La troisième étape du flot de conception utilise les caractéristiques de l'architecture cible pour enrichir l'EPDG avec le temps d'exécution et la consommation énergétique de chaque BpB. Enfin, l'espace de conception peut être exploré en prenant en compte l'énergie pour obtenir un front de Pareto des configurations en fonction de l'énergie et de la vitesse de calcul. Ces différentes étapes sont détaillées ci-après.

Extraction du parallélisme Cette première phase utilise une description séquentielle de l'application en entrée et repose sur les multiples outils proposés par la communauté pour extraire

le parallélisme de l'application [104–107] et fournir un graphe de tâches parallèles décrivant l'application. Cette étape se base sur une approche itérative qui permet à l'utilisateur de modifier la structure du programme d'entrée pour améliorer les résultats d'extraction.

Création de l'EPDG Cette deuxième étape utilise en entrée le graphe de tâches parallèles précédemment obtenu ainsi qu'une description de l'architecture cible contenant le degré de parallélisme utilisable. Ces informations permettent de calculer la taille des communications entre chaque tâche du graphe. Les petites tâches sont ensuite fusionnées entre elles ainsi que celles impliquant un gros flux de communication dans des macro-tâches. Ces macro-tâches sont ensuite enrichies de deux nœuds supplémentaires représentant les communications entrantes (*CommIn*) et sortantes (*CommOut*). Des BpB sont alors obtenus. Ils sont ensuite enrichis d'informations sur la taille des communications et reliés entre eux en fonction des dépendances pour obtenir un EPDG (voir Figure 6.3b). À ce moment du développement, l'EPDG ne contient aucune information sur le coût énergétique de calcul ou sur le temps d'exécution.

Estimation de l'énergie Un EPDG ne contient aucune information de coût et aucune idée des potentielles performances d'implémentation de chaque BpB. Cette étape repose également sur des outils extérieurs, tel que Aladdin [108], pour obtenir une estimation précise des performances atteignables par les implémentations des BpB. Il est aussi envisageable que l'utilisateur puisse utiliser des outils de synthèse de haut niveau HLS ou des mesures expérimentales, pour obtenir ces valeurs. Quelle que soit la méthode utilisée, les valeurs obtenues en termes de consommation et de temps de calcul sont insérées dans le flot de conception pour enrichir le EPDG.

Lorsque tous les BpB possèdent des informations de performance, un outil d'estimation de puissance est utilisé. Cet outil d'estimation permet d'obtenir pour chaque BpB sa consommation énergétique ainsi que son temps d'exécution. Ces valeurs sont nécessaires pour chaque implémentation matérielle et pour chaque type de processeur logiciel disponible. Ces informations de coût sont insérées dans l'EPDG afin de pouvoir réaliser l'étape d'exploration.

Exploration de l'espace de conception Les travaux de ce chapitre se concentrent principalement sur cette partie. Son point de départ est un EPDG contenant les informations relatives à la taille des communications ainsi qu'au coût/temps d'exécution des BpB pour chaque implémentation. Ces informations sont ensuite utilisées pour construire des configurations associant les implémentations utilisées pour les BpB, le placement et l'ordonnancement des BpB, etc. Pour chaque configuration obtenue dans l'espace de conception, le coût réel de consommation est estimé. Au vu du nombre de configurations à traiter, le temps de calcul est primordial. Pour répondre à cette problématique, un modèle de consommation analytique basé sur l'étude des communications est proposé dans la Section 6.2.2. Après avoir développé une infrastructure de mesure complète sur l'architecture Zynq de Xilinx, nous avons validé ce modèle de consommation sur une série d'applications générant différents schémas de communication [roux16mcsoc].

Malgré le temps d'estimation très rapide, il reste impossible d'explorer l'ensemble de l'espace de conception de manière exhaustive. Pour pallier cette problématique, nous proposons, dans la Section 6.2.3, une méthode basée sur une formulation de programmation linéaire en nombre entier mixte (MILP - Mixed integer linear programming). Cette dernière permet d'obtenir la configuration optimale pour des applications tuilées en moins d'une seconde. Pour des noyaux de calcul de type multiplication de matrice et filtre stencil, nous avons observé un gain de plus de 12 % sur la consommation par rapport à une approche traditionnelle.

L'objectif à plus long terme est de proposer des algorithmes d'exploration plus aboutis permettant d'obtenir un front de Pareto de configuration en fonction de la consommation énergétique et du temps de calcul. L'utilisateur pourra ainsi sélectionner la configuration idéale pour les contraintes de son système.

6.2 Exploration architecturale d'accélérateur

Basée sur une version simplifiée du flot de conception précédent, une méthode d'exploration de l'espace de conception pour les architectures HMpSoC est présentée dans cette section. Le

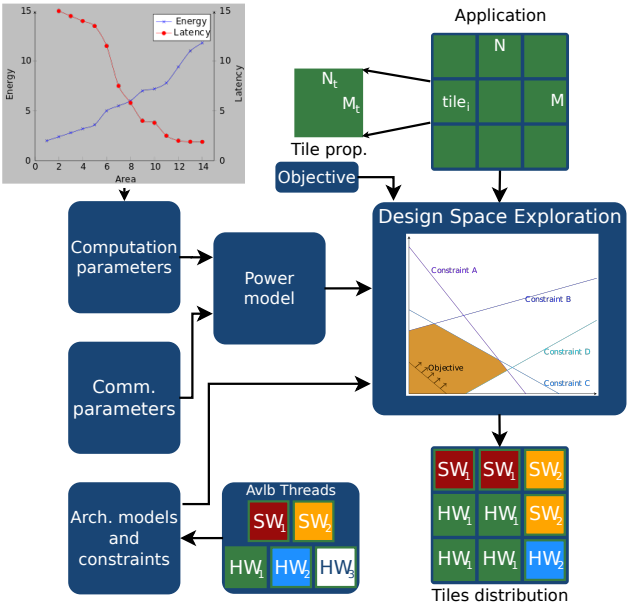


FIGURE 6.4 – Vue globale de l'exploration proposée.

DSE proposé, représenté dans la Figure 6.4, permet d'optimiser la partitionnement HW/SW pour différentes contraintes utilisateurs, en particulier la contrainte énergétique. Le flot cible les applications parallèles tuilées, afin de s'abstraire des étapes d'extraction du parallélisme et de création de l'EDPG du flot de la Figure 6.3a, et repose sur un modèle énergétique analytique, dont les paramètres énergétiques sont extraits par des micro-benchmarks.

6.2.1 Applications, architectures et objectifs du DSE

Applications parallèles tuilées

Le tuilage est une technique de parallélisation issu du domaine de la compilation [109]. Le tuilage fait référence au partitionnement de l'espace d'itération en blocs polyédriques. Cette transformation a été initialement introduite pour augmenter la localisation des données. Elle réduit le volume de données accédées entre plusieurs utilisations d'un élément, permettant à un élément de rester dans le cache jusqu'à sa prochaine utilisation. La transformation en tuile offre trois avantages intéressants pour la conception d'accélérateurs : l'augmentation de la réutilisation des données dans les mémoires tampon ; l'exposition de parallélisme gros grain qui pourrait être exploité dans le partitionnement HW/SW ; l'exposition d'un parallélisme grain fin dans chaque tuile qui pourrait être utilisé pour concevoir efficacement l'architecture HW. Les benchmarks PARSEC [110], considérés comme un lot représentatif d'applications, proposent 8 applications sur les 12 disponibles avec un parallélisme de données. Ainsi, avec certaines transformations sur l'espace d'itération, on peut considérer que 2/3 des applications sont de bons candidats pour le tuilage. Dans cette étude, nous considérons des applications 2D avec un espace d'itération N par M. Le tuilage est appliqué avec des tuiles rectangulaires composées de N_t par M_t éléments. Ainsi, un ensemble de $N_{tiles} = \frac{N}{N_t} \times \frac{M}{M_t}$ tuiles est réparti parmi les cœurs d'exécution de l'architecture cible. L'application tuilée est représentée en entrée du flot DSE de la Figure 6.4 alors que l'allocation des tuiles aux ressources HW et SW est la sortie du flot de conception.

Architectures hétérogènes

Dans notre modèle, des architectures hétérogènes sont considérées et sont composées de N_{SW} cœurs de calcul logiciels et de N_{HW} cœurs de calcul matériels. Tous les cœurs SW peuvent s'appuyer sur la même architecture de processeur ou sur un ensemble de cœurs différents. Les cœurs HW sont implémentés dans la partie PL (Programmable Logic) de l'architecture. La PL peut être partitionnée en un maximum de N_{HW} cœurs HW indépendants avec N_{HW} le nombre de ports

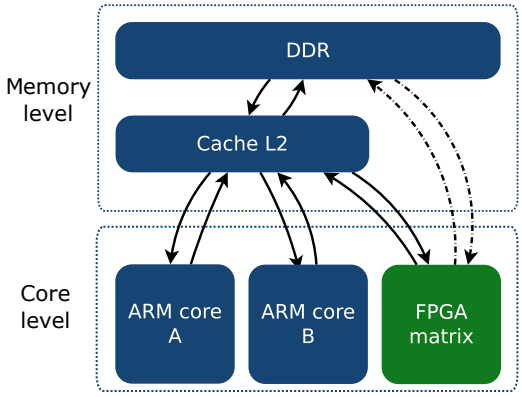


FIGURE 6.5 – Hiérarchie mémoire de l'architecture Zynq.

mémoire disponibles dans la PL. Cette restriction, basée sur le nombre maximum de transferts de données simultanés vers/depuis la mémoire principale, empêche notamment l'apparition d'une congestion mémoire entrainant un temps d'accès mémoire et un coût énergétique imprévisibles. De plus, toutes les ressources utilisées par les cœurs doivent être inférieures aux ressources globales disponibles dans l'architecture. Par exemple, l'architecture Zynq de Xilinx, qui sera utilisée dans nos expérimentations, est un cas représentatif d'architectures hétérogènes. Cette architecture est composée de deux cœurs SW identiques et d'une partie PL FPGA. La Figure 6.5 donne la représentation de la hiérarchie mémoire de l'architecture Zynq. Les cœurs HW communiquent avec les cœurs SW via quatre ports dédiés HP (High Performance) de la mémoire DDR (Double-Data-Rate). Quatre autres ports General Purpose (GP) sont disponibles pour la configuration et la synchronisation du FPGA et pour la communication avec les périphériques standard.

Objectifs du DSE

Le but de la méthode d'exploration proposée est de trouver la meilleure configuration qui minimise l'objectif de l'utilisateur (e.g. le temps d'exécution ou la consommation totale d'énergie de l'application). Une configuration est un vecteur $\vec{C}$ composé de la distribution des tuiles entre les cœurs SW et HW et du type d'implémentation utilisé pour chacun des N_{HW} cœurs . Un exemple d'une telle distribution est représenté sur la Figure 6.4 en sortie du DSE. Soit $tiles^i_{SW}$ et $tiles^j_{HW}$ le nombre de tuiles (tiles en anglais) allouées respectivement au cœur SW i et au cœur HW j. Les cœurs HW peuvent utiliser un bloc de calcul parmi un ensemble d'implémentations disponibles qui ont des performances différentes en surface, en énergie et en latence (comme illustré dans la Figure 6.4 en haut à gauche où chaque point correspond à une implémentation). Dans notre étude, l'exploration matérielle a été effectuée manuellement à l'aide des outils de synthèse de haut niveau. Le bloc de calcul est synthétisé sous différentes contraintes de latence pour obtenir différentes implémentations variant de 1 à N_{impl}, avec $impl_j$ représentant l'implémentation utilisée dans le cœur HW j. Une configuration $\vec{C}$ est donc définie par :

$$\vec{C} = [tiles^1_{SW}\,;\ldots;\,tiles^{N_{SW}}_{SW}\,;(tiles^1_{HW}, impl_1)\,;\ldots;\,(tiles^{N_{HW}}_{HW}, impl_{N_{HW}})], \tag{6.1}$$

avec le couple $(tiles^j_{HW}, impl_j)$ le nombre de tuiles allouées au cœur j et calculées par l'implémentation $impl_j$.

L'objectif du DSE est donc de trouver la meilleure configuration $\vec{C}_{opt}$ qui minimise le coût (latence, énergie) tel que :

$$\vec{C}_{opt} = \min_{\forall \vec{C}} Cost(\vec{C}), \tag{6.2}$$

avec $Cost(\bullet)$ la fonction de coût définie par des modèles analytiques définis dans la partie suivante.

6.2.2 Estimation de l'énergie et du temps d'exécution

Modèles énergétiques et du temps d'exécution

Les modèles du temps d'exécution et de l'énergie sont présentés dans cette partie. La consommation d'énergie d'une application exécutée sur une architecture multi-cœur hétérogène dépend de trois sources principales : l'énergie statique dissipée pendant le temps d'exécution, la consommation d'énergie dynamique utilisée pour les calculs et l'énergie utilisée pour les communications entre les cœurs de calcul. Les applications sont composées de N_{tiles} tuiles indépendantes qui pourraient être exécutées en parallèle. L'exécution d'une tuile est atomique, l'énergie et le temps nécessaires à son calcul ne dépendent que du cœur d'exécution ciblé. La quantité de données requises pour le calcul des tuiles est supposée connue au moment de la compilation. Les cœurs HW sont gérés avec des appels de bas niveau et des interruptions qui sont encapsulées dans un thread SW. La répartition des threads entre les cœurs de calcul est effectuée séquentiellement et commence par les cœurs HW afin de minimiser le nombre de changements de contexte.

Temps de calcul Pour une configuration $\vec{C}$, le temps total de calcul est :

$$T_t(\vec{C}) = \max\left[T_{HW}(\vec{C}), T_{SW}(\vec{C})\right], \tag{6.3}$$

avec T_{SW} et T_{HW} correspondant respectivement aux temps de calcul sur les cœurs SW et HW qui sont déterminés par :

$$T_{HW}(\vec{C}) = \max_{j\in(1\ldots N_{HW})}\left[T_{comp}^{impl_j} \times tiles_{HW}^{j} + j \times T_{spawn}\right] \tag{6.4}$$

$$T_{SW}(\vec{C}) = \max_{i\in(1\ldots N_{SW})}\left[T_{comp}^{SW_i} \times tiles_{SW}^{i} + (N_{HW}^{used}(\vec{C}) + i) \times T_{spawn}\right] \tag{6.5}$$

avec $T_{comp}^{impl_j}$ and $T_{comp}^{SW_i}$ le temps de calcul d'une tuile exécutée respectivement sur l'implémentation HW j et sur le cœur SW i. T_{spawn} représente le temps nécessaire à configurer le thread de calcul. $N_{HW}^{used}(\vec{C})$ est le nombre de cœurs HW utilisés pour la configuration $\vec{C}$:

$$N_{HW}^{used}(\vec{C}) = \sum_{j=1}^{N_{HW}} \begin{cases} 1, & \text{si } tiles_{HW}^{j} > 0 \\ 0, & \text{sinon} \end{cases}. \tag{6.6}$$

Consommation d'énergie L'énergie totale $E_t(\vec{C})$ consommée par l'exécution d'une application tuilée avec la configuration $\vec{C}$ est composée de l'énergie statique et de l'énergie dynamique due aux calculs et aux communications :

$$E_t(\vec{C}) = E_{stat} + \sum_{i=1}^{N_{SW}} (E_{comp}^{SW_i} + E_{com}^{SW}) \times tiles_{SW}^{i} + \sum_{j=1}^{N_{HW}} (E_{comp}^{impl_j} + E_{com}^{HW}) \times tiles_{HW}^{j}, \tag{6.7}$$

avec E_{stat} l'énergie statique consommée, E_{com}^{SW} et E_{com}^{HW} les énergies des communications requises pour l'exécution d'une tuile sur un cœur SW et HW respectivement. $E_{comp}^{SW_i}$ et $E_{comp}^{impl_j}$ sont les énergies nécessaires au calcul d'une tuile sur le cœur SW i et sur l'implémentation HW $impl_j$.

Les paragraphes suivants présentent une méthode pour estimer l'énergie des communications entre les cœurs et l'énergie de calcul d'une tuile.

Extraction des paramètres de communication

Dans (6.7), les paramètres de communication E_{com}^{SW} et E_{com}^{HW} représentent l'énergie requise pour lire et écrire les données pour le calcul d'une tuile sur un cœur SW et HW respectivement. L'extraction de ces paramètres est réalisée en utilisant la méthode décrite dans [roux16mcsoc].

La méthodologie proposée repose sur l'utilisation de micro-benchmarks qui stimulent un chemin de communication spécifique de l'architecture. Comme illustrée par la Figure 6.5, une architecture

HMpSoC de type Zynq est composée de trois types de canaux de communication : du cœur SW à la mémoire SW (cache L2) ; de la mémoire SW vers la mémoire DDR ; et du cœur HW à la mémoire DDR. Pour déterminer les paramètres du modèle de cette architecture, les micro-benchmarks suivants ont été implémentés :

- SwChannel : ce micro-benchmark permet d'évaluer les coûts de communication entre le cœur de calcul SW et la mémoire SW. Il génère des accès en lecture (read) ou en écriture (write) dans un tableau alloué dans le cache L2.
- HwChannel : ce micro-benchmark permet d'évaluer les coûts de communication entre les accélérateurs HW et la mémoire DDR. Il génère des accès en lecture (read) ou en écriture (write) dans un tableau alloué dans la mémoire DDR.
- MemoryChannel : ce micro-benchmark permet d'évaluer les coûts de communication entre les mémoires SW et DDR. Ces coûts ne pouvant être mesurés directement, des accès entre le cœur SW et un tableau alloué dans la mémoire DDR sont générés. Les paramètres MemoryChannel sont ensuite déduits en soustrayant les valeurs SwChannel.

L'exécution de ces micro-benchmarks est paramétrée par la taille N des données à communiquer.

Lorsque qu'un cœur de calcul exécute une tuile, une fonction peut être utilisé pour calculer le coût des communications résultant. Soit $\mathcal{C}$ l'ensemble des canaux de communications utilisés pour exécuter une tuile, $\mathcal{C} = \{\text{HwChannel_read, HwChannel_write}\}$ pour les cœurs HW et $\mathcal{C} = \{\text{SwChannel_read, SwChannel_write, MemoryChannel_read, MemoryChannel_write}\}$ pour les cœurs SW. L'énergie des communications peut être calculée par :

$$E_{com}^{HW/SW}(N) = \sum_{c \in \mathcal{C}} e_{0_c} N + e_{1_c}, \tag{6.8}$$

avec e_{0_c} et e_{1_c} les coûts énergétiques du canal traversé c qui sont estimés par les micro-benchmarks. La Section 6.3.2 présente les coûts énergétiques extraits expérimentalement dans le cas d'une architecture Zynq.

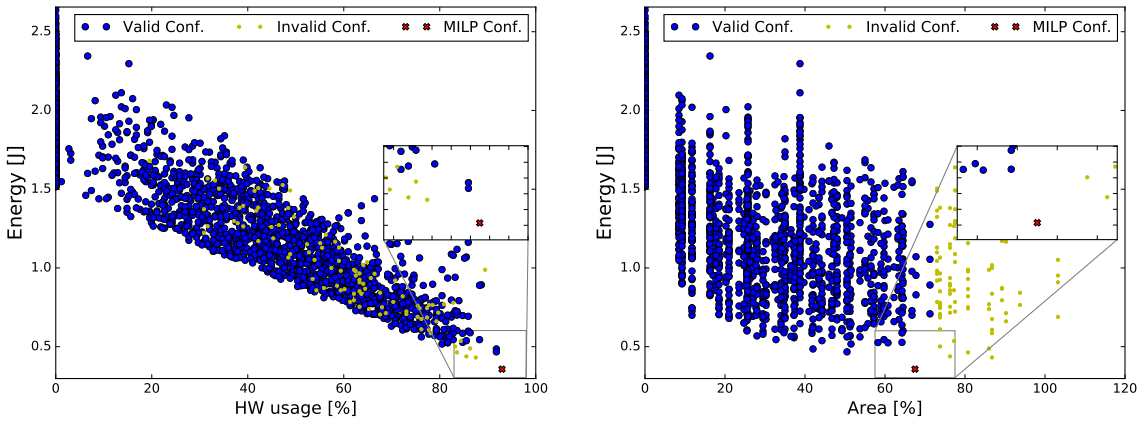
Extraction des paramètres de calcul

Les paramètres de calcul sont l'énergie et le temps nécessaires à l'exécution d'une tuile pour une configuration donnée ($E_{comp}^{SW_i}$ ou $E_{comp}^{impl_j}$). Pour extraire ces paramètres, des traces d'exécution sont nécessaires pour chaque configuration. Ces traces permettent d'obtenir les informations de temps d'exécution et d'énergie totale d'une configuration. Pour faciliter le processus d'extraction au regard du nombre de configuration, seul un sous-ensemble de l'espace de conception est évalué pour une architecture de base composée d'un seul cœur SW et d'un seul cœur HW. Les traces de mesures sont obtenues en faisant varier le partitionnement SW/HW des tuiles dans ce sous-ensemble d'architecture.

Dans la Sous-section 6.3.2, nous montrons que seules 9 configurations (i.e. 9 partitionnements SW/HW) par cœur SW et implémentation HW sont nécessaires pour obtenir une bonne précision de l'extraction des paramètres pour une application tuilée composée de $N_{tiles} = 256$. Les traces d'exécution sont ensuite traitées avec un algorithme de régression linéaire (LSR - Least Square Regression) pour extraire les paramètres. Cet algorithme calcule les valeurs des paramètres qui minimisent l'erreur quadratique entre les mesures et la fonction de coût.

Les paramètres du temps d'exécution sont extraits avec les fonctions de coût temporel définie dans (6.4) et (6.5). Pour chaque trace d'exécution, la configuration impliquant l'implémentation $impl_j$ et le type de logiciel SW_i est connu, ainsi que le temps d'exécution global. L'algorithme LSR extrait les valeurs des paramètres $T_{comp}^{impl_j}$, $T_{comp}^{SW_i}$ et T_{spawn}.

Les paramètres d'énergie sont extraits avec le même processus. L'extraction est basée sur la fonction de coût énergétique (6.7). Pour une trace d'exécution impliquant l'implémentation $impl_j$ et le type de logiciel SW_i, la consommation d'énergie globale est calculée, ainsi que l'énergie des communications E_{com}^{SW} et E_{com}^{HW} et l'énergie statique E_{stat}. Ensuite, l'algorithme LSR extrait les valeurs des paramètres $E_{comp}^{impl_j}$ et $E_{comp}^{SW_i}$.



(a) Matmult : Energie vs utilisation HW.

(b) Matmult : Energie vs surface.

FIGURE 6.6 – Exploration de l'espace de conception pour l'application Matmult.

6.2.3 Exploration de l'espace de conception

Cette parte adresse l'exploration de l'espace de conception par deux approches. Une recherche exhaustive permet de montrer que la taille importante de l'espace de conception peut entraîner une variation importante de la fonction de coût. Ensuite, une formulation MILP (Mixed Integer Linear Programming) est proposée pour résoudre le problème d'exploration et trouver la meilleure solution.

Recherche exhaustive

En considérant l'architecture cible similaire à l'architecture Zynq avec deux cœurs SW du même type et quatre cœurs HW, l'espace de conception est décrit comme un polyèdre et sa taille peut être calculée à l'aide de la bibliothèque `isl` [111] en fonction des paramètres N_{impl} et N_{tiles}. Par exemple, avec $N_{impl} = 3$ et $N_{tiles} = 256$, l'espace de conception est composé de 7 866$\times 10^{11}$ points. En supposant que la consommation d'énergie peut être obtenue en environ 1 ms, une recherche exhaustive nécessiterait plus de 9 jours et n'est donc pas une approche valide.

La Figure 6.6a montre un sous-ensemble choisi au hasard des configurations de l'espace de conception pour l'application multiplication de matrice (Matmult) étudiée par la suite. La figure représente la consommation d'énergie en fonction du pourcentage d'utilisation du HW, correspondant à la proportion des tuiles calculée sur la partie HW. La Figure 6.6b montre, pour les mêmes configurations explorées, la consommation d'énergie en fonction du pourcentage de ressources utilisées de la partie HW. Ces figures montrent que l'espace de conception est très grand et que l'exploration ne garantit pas la validité d'implémentation des configurations obtenues. Une configuration non valide dans la Figure 6.6 est une configuration qui nécessite davantage de ressources que celles disponibles sur le circuit.

Pour résoudre ce problème, la section suivante présente une méthode d'optimisation basée sur la formulation MILP, qui résout le problème défini par (6.2) et trouve la meilleure solution dans l'espace de conception valide. Les Figures 6.6a et 6.6b donnent également la solution obtenue grâce à l'optimisation MILP (configuration tracée avec le symbole $\times$ en rouge).

Formulation MILP

La programmation linéaire en nombres entiers mixtes (MILP) est une solution connue pour la résolution des problèmes de partitionnement. Avec cette approche, les contraintes sont définies comme un ensemble d'inégalités avec des variables booléennes, entières (discrètes) et non-entières (continues). Les solutions peuvent être ensuite déterminées en utilisant des solveurs commerciaux ou open source. L'optimisation est définie par une fonction de coût linéaire. L'intersection des contraintes d'inégalité représente un polyèdre des solutions réalisables. La solution optimale se

trouve à l'intersection de la fonction de coût et de l'espace des solutions réalisables. Cette partie présente l'ensemble des inégalités décrivant les contraintes de notre problème. Ensuite, deux fonctions de coût sont définies.

Contraintes du modèle Trois contraintes sont définies afin d'obtenir une configuration valide :

- Contrainte de couverture : elle permet de garantir que chaque tuile ne sera exécutée qu'une seule fois :

$$N_{tiles} = \sum_{i=1}^{N_{SW}} tiles_{SW}^{i} + \sum_{j=1}^{N_{HW}} tiles_{HW}^{j}. \tag{6.9}$$

- Contrainte d'unicité : pour s'assurer qu'une seule implémentation matérielle est utilisée pour chaque thread HW, la variable $usedImpl_{type}^{j}$ est définie par :

$$usedImpl_{type}^{j} = \begin{cases} 1, & \text{si } impl_j = type \\ 0, & \text{sinon} \end{cases} \tag{6.10}$$

Cette variable binaire peut être multipliée par une autre variable sans introduire de non-linéarité insoluble dans le modèle. La contrainte d'unicité garantit qu'une seule implémentation est utilisée dans chaque cœur HW et est définie par :

$$\forall j \in (1 \ldots N_{HW}) : \sum_{type=1}^{N_{impl}} usedImpl_{type}^{j} \leq 1. \tag{6.11}$$

- Contrainte de ressource : la partie HW (PL) de l'architecture ciblée est contrainte par la quantité de ressources disponibles. Ces ressources sont séparées en quatre catégories : les blocs RAM (BRAM), les blocs DSP (DSP), les Flip-Flops (FF) et les Look Up Tables (LUT). La contrainte de ressource garantit la disponibilité des ressources et est définie par :

$$\forall r \in \{\text{BRAM, DSP, FF, LUT}\} : \sum_{j=1}^{N_{HW}} \sum_{type=1}^{N_{impl}} usedImpl_{type}^{j} \times RscCost_{r}^{type} \leq avlbRsc_r, \tag{6.12}$$

avec $RscCost_{r}^{type}$ le coût d'une ressource r pour l'implémentation $type$ et $avlbRsc_r$ les ressources disponibles de type r de l'architecture.

Fonctions de coûts Le solveur construit d'abord le vecteur de configuration $\vec{C}$ en utilisant les contraintes, puis sélectionne la meilleure solution en utilisant une fonction de coût dépendant de l'objectif d'optimisation. Deux objectifs sont définis : minimiser l'énergie d'exécution globale en utilisant la fonction de coût (6.7) ou minimiser le temps d'exécution global en utilisant les fonctions de coût (6.5) et (6.4). Ces fonctions objectif sont définies par :

$$\vec{C}_{energy} = \min_{\forall \vec{C}} \left[E_t(\vec{C}) \right] \tag{6.13}$$

$$\vec{C}_{time} = \min_{\forall \vec{C}} \left[\max \left[T_{SW}(\vec{C}), T_{HW}(\vec{C}) \right] \right] \tag{6.14}$$

L'optimisation MILP donne le meilleur vecteur de configuration en fonction de l'objectif et une estimation de ses caractéristiques (consommation d'énergie et temps d'exécution). La Figure 6.6 donne la solution obtenue avec la méthode MILP pour la multiplication de matrice. Cette solution est toujours optimale en termes de coûts et est calculée en un temps limité.

6.3 Résultats expérimentaux d'exploration

Cette section présente les deux applications utilisées, l'architecture cible et l'outil de mesure de puissance. Ensuite, les implémentations matérielles utilisées dans chaque cas sont décrites. Enfin, les résultats de l'extraction des paramètres sont présentés ainsi les résultats et la précision de l'optimisation MILP.

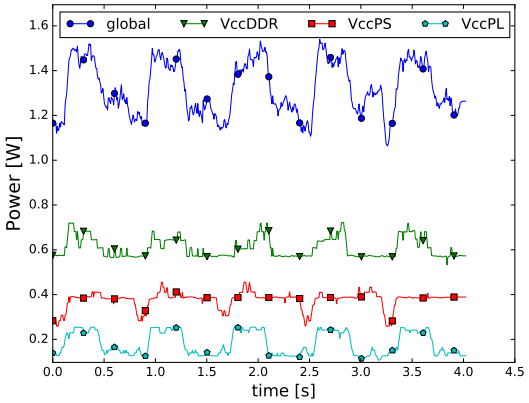


FIGURE 6.7 – Traces de mesure de puissance pour l'application Matmult, avec 37.5 % des calculs en HW.

6.3.1 Environnement d'expérimentation

Noyaux d'application

Le premier noyau est une application de multiplication matricielle (Matmult). Les matrices d'entrée et de sortie sont de taille 512×512. Le calcul matriciel est divisé en 256 tuiles de taille 32×32. Chaque tuile lit des lignes ou des colonnes entières des matrices d'entrée et écrit les résultats dans la tuile de sortie. Le deuxième noyau est un calcul de Stencil basé sur un filtre de taille 4×4, qui est successivement appliqué 10 fois. Les matrices d'entrée et de sortie sont de taille 542×542 et 512×512, respectivement. L'application est tuilée avec un chevauchement afin d'empêcher la dépendance entre les tuiles, ce qui entraîne le calcul de 256 tâches indépendantes.

Infrastructure de mesure de consommation

Les expérimentations sont réalisées sur l'architecture Zynq (carte Zynq Zc702) de Xilinx [112] composée de deux processeurs ARM cortex A9 associés à une partie FPGA. La consommation d'énergie est mesurée via le bus *Power Management Bus (PMBus)* intégré sur la carte. Dans nos expérimentations, un composant HW échantillonne à intervalles réguliers 7 rails d'alimentation du Zynq. Ces mesures permettent de calculer la puissance instantanée et la variation d'énergie pendant l'exécution de l'application. La Figure 6.7 affiche les courbes de puissance obtenues pour l'application Matmult avec 37.5 % des calculs effectués dans la partie HW. Les courbes montrent les trois rails d'alimentation les plus représentatifs : *VccPS* alimente les processeurs ARM, *VccPL* le noyau de calcul dans le partie PL et *VccDDR* la mémoire DDR externe. La dernière courbe *global* affiche la consommation totale sur les 7 rails. Ces courbes montrent que la résolution temporelle de la mesure de puissance permet de distinguer précisément chaque étape de calcul. Avec cette répartition SW/HW, le calcul HW se termine avant celui SW pour chaque exécution. Le calcul SW est donc le goulot d'étranglement de cette configuration.

Implémentations matérielles

Une exploration matérielle a été réalisée afin d'obtenir plusieurs implémentations HW avec des caractéristiques différentes en termes de ressources utilisées et de latence de calcul. Seules des implémentations représentatives ont été sélectionnées parmi tout l'espace de conception. Les implémentations matérielles ont été générées à l'aide de Vivado HLS tool *(2015.4 HLX Edition)* [113].

La Table 6.1 donne la latence et les ressources matérielles pour chaque implémentation du calcul d'une tuile de l'application Matmult. Toutes ces implémentations sont pipelinées, mais le degré de parallélisme des nids de boucle est différent. Le nom de l'implémentation dans la table exprime le degré de parallélisme des trois boucles internes. LnP est synonyme de *Loop-nest Parallelism*. Par exemple, LnP 248 signifie que les 2 itérations de la troisième boucle sont réalisées en parallèle, 4 pour la deuxième et 8 pour la première.

Le même type d'exploration est réalisé pour le calcul d'une tuile de l'application Stencil et les implémentations obtenues sont également données dans la Table 6.1.

Apps.	Impl.	Lat. [Mcycle]	Resources [%]			
			BRAM	DSP	FF	LUT
Matmult	LnP 114	140	25%	2%	2%	5%
	LnP 118	73	25%	4%	2%	6%
	LnP 128	40	25%	9%	3%	10%
	LnP 148	23	25%	18%	6%	16%
	LnP 248	14	26%	36%	11%	30%
	LnP 448	12	30%	59%	19%	47%
Stencil	LnP 114	12	24%	39%	10%	42%
	LnP 244	7	21%	75%	20%	72%
	LnP 384	6	24%	100%	28%	96%

TABLE 6.1 – Implémentations HW pour l'application Matmult et Stencil.

Implémentations logicielles

L'implémentation logicielle est écrite afin d'exposer le parallélisme des instructions et de maximiser la réutilisation du cache. Compilé avec gcc et le flag *-Ofast*, le code binaire obtenu utilise les instructions NEON dédiées à l'architecture ARM. Le SW obtenu prend environ 8.5×10^{-3} s pour calculer une tuile pour l'application Mamult et 2×10^{-3} s pour l'application Stencil sur un cœur ARM fonctionnant à 800 MHz.

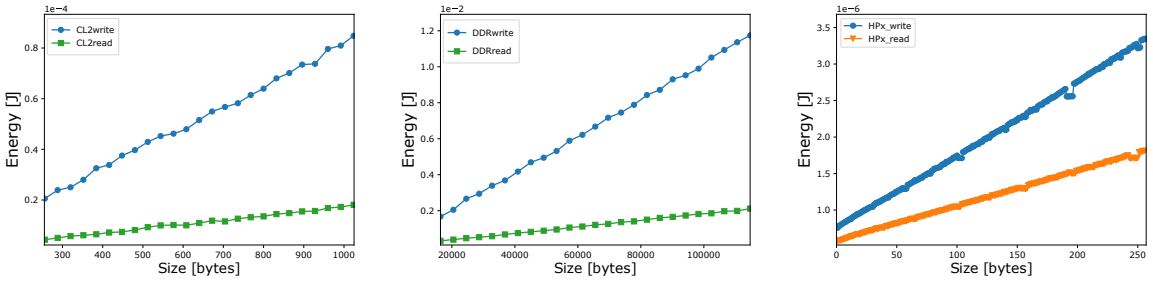
6.3.2 Extraction des paramètres

Paramètres de communication

Introduit dans [roux16mcsoc], un ensemble de micro-benchmarks a été conçu pour extraire les paramètres de communication de la plateforme Zynq. Pour les deux applications, les cœurs SW et HW ont accès aux données situées dans la mémoire DDR. Les cœurs SW accèdent donc à la fois au cache L2 et à la mémoire DDR par un accès mémoire standard, tandis que les cœurs HW accèdent à la mémoire DDR en utilisant les quatre ports HP. Afin d'extraire le coût de communication de ces canaux, les trois micro-benchmarks suivants sont utilisés :

— CL2 : pour le canal entre les cœurs SW et le deuxième niveau de mémoire. Il génère des opérations de lecture ou d'écriture à partir du cœur du SW dans un tableau situé dans le cache L2.

— DDR : pour le canal entre les cœurs logiciels et la mémoire externe. Il génère des opérations de lecture ou d'écriture à partir du cœur SW dans un tableau situé dans la mémoire DDR.

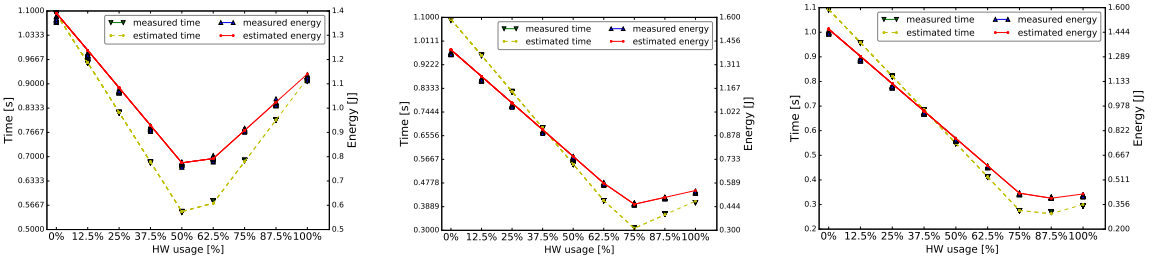


(a) Ecriture/lecture entre le cœur SW et le cacle L2.

(b) Ecriture/lecture entre le cœur SW et la mémoire DDR.

(c) Ecriture/lecture entre le cœur HW et la mémoire DDR.

FIGURE 6.8 – Energie consommée moyenne en fonction de la taille des données (N en Octets) pour différents canaux de communication.



(a) Matmult Impl. LnP 118. (b) Matmult Impl. LnP 148. (c) Matmult Impl. LnP 448.

FIGURE 6.9 – Temps d'exécution et consommation d'énergie en fonction de l'utilisation HW : valeurs mesurées et estimées.

— HPx : pour le canal entre un cœur matériel et la mémoire externe. Il génère des opérations de lecture et d'écriture à partir du cœur HW dans un tableau situé dans la mémoire DDR.

La Figure 6.8 montre la consommation moyenne d'énergie pour l'accès (lecture et écriture) au cache L2 et à la mémoire DDR. Les résultats sont donnés pour différentes tailles N des données (en octets) et moyennés sur 20 exécutions des micro-benchmarks. Les résultats montrent que la variance des mesures est assez faible. Le coût de la consommation d'énergie de ces communications peut être approximé comme une fonction linéaire $f(N) = e_{0_c} \times N + e_{1_c}$, où les valeurs e_{0_c} et e_{1_c} sont respectivement les parties dynamiques et statiques.

Paramètres de calcul

En utilisant la méthode introduite dans la partie 6.2.2, un sous-ensemble des configurations a été exécuté 20 fois pour chaque implémentation HW (une seule implémentation SW est utilisée). La consommation énergétique moyenne et le temps d'exécution ont ensuite été extraits des mesures. Tout d'abord, une estimation du coût de communication est calculée pour chaque configuration. Ensuite, les fonctions de coût temporelles (6.5) et (6.4) et énergétiques (6.7) sont utilisées avec un algorithme LSR pour extraire les paramètres de coût en temps et en énergie.

L'extraction et l'estimation ont été effectuées pour chaque implémentation HW des applications Matmult et Stencil. Les mesures sont entièrement automatisées et prennent 336 s pour le Matmult et 122 s pour le Stencil. Ce temps de mesure est vraiment acceptable car l'extraction de paramètres est exécutée une seule fois. La Figure 6.9 donne un exemple de résultats pour trois implémentations HW de l'application Matmult. Ces courbes affichent les valeurs mesurées et estimées de consommation d'énergie et de temps d'exécution pour différents partitionnements HW/SW de l'ensemble des tuiles. Ces résultats montrent que l'erreur entre les valeurs mesurées et estimées est faible. Les paramètres de coût SW (i.e. HW usage = 0%) sont indépendants de la mise en œuvre HW associée. Enfin, ces courbes montrent que la distribution de tuiles HW/SW qui minimise la consommation d'énergie dépend de l'implémentation utilisée.

6.3.3 Performances de l'exploration MILP

Optimisation MILP

Les paramètres de coût précédemment extraits sont intégrés dans les modèles MILP pour les 2 applications et le problème d'optimisation est résolu pour la fonction de coût énergétique (6.13) à l'aide du solveur MILP *Gurobi* [114] sur un processeur *Intel i7 Haswell-ult* fonctionnant à 2.10 GHz. La solution MILP est obtenue après 18980 itérations du simplexe en $\simeq 0.73$ s pour l'application Matmult et après 348 itérations en $\simeq 0.03$ s pour l'application Stencil. Cette différence est principalement due au plus grand nombre d'implémentations HW disponibles pour l'application Matmult.

La Table 6.2 donnent les configurations MILP obtenues avec la répartition des tuiles sur les cœurs et les implémentations utilisées. Pour l'application Matmult, 18 tuiles sont réparties équitablement sur les cœurs SW, les tuiles restantes sont réparties sur trois noyaux HW. L'implémentation *LnP 248* est utilisée sur deux cœurs HW, le troisième utilise une implémentation plus petite pour occuper l'espace restant dans la partie PL du FPGA. Dans ce cas, le nombre de cœurs matériels

Conf.		Thread ID					
		SW0	SW1	HW0	HW1	HW2	HW3
Mat.	Impl.	SW	SW	LnP 248	LnP 248	LnP 148	none
	N_{tiles}	9	9	87	85	66	0
Stenc.	Impl.	SW	SW	LnP 114	LnP 114	none	none
	N_{tiles}	25	24	104	103	0	0

Table 6.2 – Configurations obtenues par l'optimisation MILP.

utilisés est limité par les ressources du FPGA disponibles et non par l'accès à la mémoire. Pour l'application Stencil, seuls deux cœurs HW sont utilisés, les implémentations HW du Stencil utilisant plus de ressources que celles de Matmult.

Précision et gain

Les performances de la configuration MILP sont comparées à la meilleure configuration *Sample Conf.* du sous-ensemble de l'espace de conception (cf. section 6.2.3) et à deux configurations de base : *Full SW* et *Full HW*. Pour les applications Matmult et Stencil, la Table 6.3 donne les ressources utilisées par chaque configuration et les estimations du temps d'exécution et de la consommation d'énergie. Comme la configuration MILP peut utiliser plus d'un cœur matériel, la quantité totale de ressources utilisées peut être supérieure à celle de la configuration *Sample Conf.*. C'est le cas pour l'application Matmult. D'autre part, la configuration MILP pour l'application Stencil utilise deux implémentations plus petites que celle utilisée dans la configuration *Sample Conf.*, ce qui entraîne une utilisation importante de BRAM, mais moins de blocs DSP. Le gain d'accélération et la réduction d'énergie sont le rapport entre une configuration donnée et la configuration *Sample Conf.* la plus économe en énergie, prise comme référence. La configuration MILP permet une réduction d'énergie d'environ 34 % pour l'application Matmult et d'environ 12 % pour l'application Stencil.

Afin de valider l'approche proposée, les estimations du temps d'exécution et de l'énergie consommée de la configuration MILP sont comparées aux mesures réelles de la même configuration sur la carte Zynq. Les valeurs mesurées sont des moyennes sur 40 exécutions indépendantes de la configuration MILP. La Table 6.4 donne les valeurs estimées et mesurées pour les deux applications. L'erreur moyenne pour l'application Matmult est d'environ 5 %, ce qui est relativement faible pour des mesures de puissance sur une architecture réelle. Pour l'application Stencil, l'erreur moyenne est plus élevée. Son temps d'exécution étant plus court que celui de la multiplication de matrice, l'application Stencil est plus sensible à l'imprécision de la mesure de puissance.

Conf.		Resources [%]				Time [s]	acc. gain	Energy [J]	energy red.
		BRAM	DSP	FF	LUT				
Matmult	Full SW	0%	0%	0%	0%	1.21	−307.7%	1.59	−199.4%
	Full HW	30%	59%	19%	47%	4.1×10^{-1}	−5.1%	5.6×10^{-1}	−3.7%
	Sample Conf.	30%	59%	19%	47%	3.9×10^{-1}	0%	5.4×10^{-1}	0%
	MILP	76%	90%	28%	76%	2.290×10^{-1}	**41.3%**	3.558×10^{-1}	**34.1%**
Stencil	Full SW	0%	0%	0%	0%	3.2×10^{-1}	−150%	4.2×10^{-1}	−135.9%
	Full HW	24%	100%	28%	96%	1.28×10^{-1}	0%	1.78×10^{-1}	0%
	Sample Conf.	24%	100%	28%	96%	1.28×10^{-1}	0%	1.78×10^{-1}	0%
	MILP	48%	78%	20%	84%	1.121×10^{-1}	**12.4%**	1.566×10^{-1}	**12%**

Table 6.3 – Performances de l'exploration MILP.

Application		Time [s]	err. [%]	Energy [J]	err. [%]
Mat.	Est.	2.29×10^{-1}	3.35%	3.558×10^{-1}	5.03%
	Meas.	2.369×10^{-1}		3.746×10^{-1}	
Sten.	Est.	1.121×10^{-1}	9.58%	1.566×10^{-1}	10.25%
	Meas.	1.24×10^{-1}		1.745×10^{-1}	

Table 6.4 – Précision de l'exploration MILP : comparaison des performances estimées et mesurées.

6.4 Autres contributions

6.4.1 Plateforme d'émulation de HMpSoC

Les expérimentations précédentes ont soulevé des problèmes annexes au flot de conception proposé avec notamment les difficultés liées au test des algorithmes d'exploration sur une large gamme d'architectures et d'applications. En effet, pour effectuer la validation sur une nouvelle architecture, il est nécessaire de développer une infrastructure de mesure permettant la récupération des informations de consommation. Cette étape nécessite du temps et requiert la présence de capteurs sur la cible. Au niveau des applications, de nombreux jeux de benchmarks existent avec diverses caractéristiques, mais de la même manière le portage de ces applications sur l'architecture cible demande beaucoup de temps.

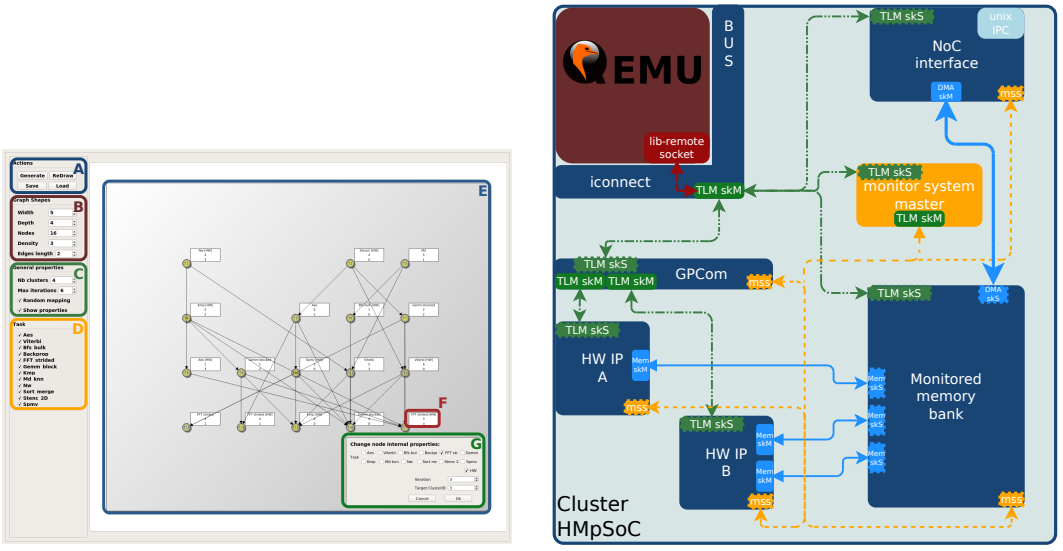
Pour pallier ces problèmes, nous avons introduit une plateforme d'émulation d'architectures HMpSoC configurable qui permet de simuler un grand nombre d'architectures. Cette dernière associe les technologies QEMU et SystemC TLM-2.0 permettant d'atteindre de très bonnes performances en termes de temps de simulation. La structure interne des clusters utilisés dans la plateforme d'émulation est présentée dans la Figure 6.10b. Cette plateforme d'émulation intègre un système d'estimation de l'énergie et un système de génération d'applications représentatives basé sur l'association aléatoire de divers noyaux de calcul. Les noyaux utilisés reposent sur le concept des "Dwarf" introduit par des chercheurs de Berkeley [115]. Un dwarf est un ensemble de noyaux de calcul représentatif permettant de décrire un très grand nombre d'applications parallèles. Le benchmark MachSuite est utilisé pour l'implémentation des dwarfs, il est composé de 12 noyaux de calculs représentatifs [116]. La Figure 6.10a montre l'interface du générateur avec un exemple d'application générée.

Cette architecture va permettre le test automatisé et à grande échelle d'algorithmes de partitionnement HW/SW et de placement de tâches sur des architectures HMpSoC, et ainsi faciliter l'adoption de ces nouvelles architectures.

6.5 Conclusion

Les travaux présentés se concentrent sur l'élaboration d'un flot de conception efficace en énergie pour des applications de radio logicielle embarquées hautes performances ciblant des architectures multi-cœurs hétérogènes. Nous avons proposé une formalisation générique de la structure des architectures cibles ainsi que des applications. Sur ces bases, nous avons proposé un flot de conception ciblant l'efficacité énergétique.

Dans un premier temps, nous avons proposé un modèle de consommation calculable rapidement, basé sur les communications, qui permet de traiter la phase de placement de tâches sur une architecture hétérogène. Ensuite, nous avons proposé une formalisation MILP, permettant de trouver la configuration optimale en terme d'énergie pour le placement d'applications tuilées sur ce type d'architecture. Enfin, nous avons introduit une plateforme d'émulation d'architectures HMpSoC, embarquant un modèle de consommation ainsi qu'un générateur d'applications. Cette plateforme ouvre de nouvelles opportunités dans le domaine des algorithmes de partitionnement HW/SW et de placement des taches en permettant leur validation à grande échelle. Ces travaux ont donc mis en évidence les problématiques limitant l'adoption des architectures HMpSoC et proposent des



(a) Interface de l'application de génération de graphes.

(b) Structure des clusters de la plateforme d'émulation.

FIGURE 6.10 – Plateforme d'émulation de HMpSoC.

outils pour répondre à ces problèmes et faciliter les futures recherches.

La suite directe de ces travaux est l'application de l'algorithme de partitionnement aux cœurs de calcul de type dwarfs et son extension à des applications complètes. Une large gamme d'architectures et d'applications peut ensuite être évaluée avec la plateforme d'émulation, ainsi que différents algorithmes d'exploration. Une autre perspective est l'orchestration dynamique des ressources HW et SW. En effet, la flexibilité temps-réel de l'application (e.g. pour les télécommunications) requiert une architecture matérielle (e.g. un partitionnement HW/SW) pouvant évoluer au cours du temps. Or, cette dynamicité n'est pas permise par l'algorithme d'optimisation proposé. Une solution est une exploration mixte de l'espace de conception, avec une pré-optimisation couvrant l'espace complet et un manager embarqué gérant un espace restreint en cours d'exécution.

Chapitre 7

Conclusions et perspectives

Après l'exposé des différentes contributions scientifiques dans les chapitres précédents, ce chapitre dresse le bilan scientifique de mes activités scientifiques depuis mon arrivée à l'IRISA en 2011. Tout d'abord, un bilan des résultats autour des deux axes de recherches sera dressé ainsi qu'un bilan plus personnel. Enfin, des perspectives à plus long terme seront données et nous verrons qu'il n'est pas exclu de faire converger ces recherches ou de les appliquer à d'autres domaines.

Sommaire

7.1 Bilan scientifique 105
7.1.1 Axes de recherche 105
7.1.2 Bilan personnel 106
7.2 Perspectives de recherche 107
7.2.1 Smart wake-up radio 107
7.2.2 Near sensor computing 108
7.2.3 Radio logicielle pour l'interception de signaux compromettants 108
7.2.4 Partitionnement efficace en énergie pour les architectures Cloud-RAN . 109

7.1 Bilan scientifique

7.1.1 Axes de recherche

Ces dernières décennies, la consommation d'énergie est devenue primordiale à la conception des systèmes embarqués communicants tout en maintenant une bonne qualité de service et une capacité de calcul importante. L'efficacité énergétique du système traduit directement ce compromis entre les ressources utilisées et les performances atteintes et son optimisation est au cœur de nos travaux de recherche. Ainsi, même les plus petits objets communicants pourront traiter des données issues de capteurs générant un flux important de données ou réaliser une partie d'un traitement distribué sur plusieurs systèmes.

Un autre élément important est la variabilité de l'environnement dans lequel évolue l'objet communicant. La topologie du réseau, les ressources énergétiques et/ou la qualité de service (dictée par une application) peuvent évoluer au cours du temps. Dans ce contexte, un dimensionnement statique du système est clairement inefficace. Une adaptation dynamique des différents éléments du système permet de faire face aux variations de l'environnement. Ainsi, des algorithmiques d'adaptation des couches protocolaires (e.g. l'intervalle de réveil) et physiques (e.g. la puissance d'émission) permettent d'améliorer l'efficacité énergétique des systèmes. Au niveau du matériel, une reconfiguration dynamique de l'architecture (e.g. partitionnement SW/HW, dimensionnement virgule fixe, ...) permet également de s'adapter à l'environnement. Une approche inter-couche permettrait d'optimiser globalement le système.

Différents travaux ont été menés sur l'optimisation de la consommation énergétique des objets communicants très faible consommation d'énergie. Les résultats obtenus permettent de croire en l'autonomie énergétique des objets tout en maximisant l'efficacité d'utilisation de l'énergie récupérée. Cette haute efficacité est obtenue par la conception d'algorithmes d'adaptation et de protocoles de communication dédiés. De plus, nous avons montré le potentiel des wake-up radios pour réduire significativement l'énergie consommée et la latence des communications pour des applications ne

nécessitant pas de longue portée (e.g. les réseaux corporels, les bâtiments intelligents, la collecte d'information, ...). Nous avons notamment montré expérimentalement comment les wake-up radios sont parfaitement adaptées aux trafics variables et permettent d'obtenir des communications efficaces en énergie et voire des réseaux autonomes lorsqu'elles sont combinées à la récupération d'énergie.

Dans le domaine de la conception d'architectures matérielles, nos travaux ont cherché à augmenter le niveau d'abstraction des architectures reconfigurables afin d'améliorer leur programmabilité tout en conservant leur capacité de calcul et de reconfiguration. Une partie de nos travaux repose sur les outils de synthèse de haut niveau et nous avons montré que leur utilisation permet d'obtenir des performances intéressantes (temps de développement, ressources, latence, flexibilité) en augmentant le niveau d'abstraction de la cible matérielle. Cependant, ces performances sont obtenues après une certaine pratique des outils, leur utilisation nécessite donc une certaine expertise et une connaissance des contraintes d'implémentations. Ensuite, avec l'augmentation de la complexité des architectures associant plusieurs types de cœur de calcul, nous avons montré l'intérêt de ces architectures pour des plateformes radio logicielle haute performance combinant capacités de calcul élevées et flexibilité matérielle. L'utilisation de ce type de plateforme passe par une augmentation du niveau de description de l'application pour un prototypage rapide et par une prise en compte au plus tôt de l'énergie dans le flot de conception.

7.1.2 Bilan personnel

Ce document dresse le bilan de 7 années de recherche depuis ma prise de fonction de maître de conférences en 2011. Initiés au sein de l'équipe CAIRN de l'IRISA, ces travaux ont ensuite été poursuivis dans l'équipe GRANIT suite à sa création en 2005. La création de l'équipe a été une expérience enrichissante. Si des thématiques scientifiques étaient bien définies, une activité dynamique et conséquente autour des ces thématiques était entièrement à construire par des financements de thèses et le montage de projets collaboratifs et au travers du rayonnement de l'équipe.

Les travaux développés au cours de ces années sont au croisement des compétences entre les systèmes embarqués, les communications numériques, les architectures de calcul et le traitement du signal, le but étant d'améliorer globalement l'efficacité énergétique des systèmes communicants en combinant l'ensemble de ces compétences. Pour atteindre cet objectif, nous développons les performances théoriques des systèmes de transmission ou des algorithmes que nous proposons grâce à des expressions analytiques, puis nous affinons ces modèles avec des micro-benchmarks et enfin nous validons les performances par l'implémentation des techniques proposées sur des plateformes matérielles. Cette approche de modélisation hybride des systèmes est une approche que je souhaite continuer à développer afin d'optimiser leur dimensionnement : des modélisations analytiques pour explorer au mieux l'espace de conception et des modèles expérimentaux pour affiner et valider les modèles théoriques.

Les innovations proposées dans nos travaux sont donc très souvent validées et implémentées sur des plateformes matérielles et donnent lieu à des prototypes de démonstrations. Nous essayons de garder un contact permanent avec la réalité industrielle afin de proposer des solutions pouvant être directement appliquées. Ces travaux de recherche émanent donc de problématiques industrielles permettant une très forte interaction avec le tissu industriel local et national. Ces interactions sont essentielles au développement de nos activités de recherches par leur financement direct au travers de projets collaboratifs ou de contrats bilatéraux.

Ces travaux couvrent 5 thèses : 2 visant les réseaux de capteurs ultra-faible consommation et 3 sur les radios logicielles efficaces en énergie. Ils ont donné lieu à des publications dans 7 revues internationales et 26 conférences internationales. Quatre démonstrations dans des conférences ont également été présentées. La qualité des travaux proposés repose en grande partie sur le travail des doctorants et sur la visibilité de l'équipe Granit au niveau national. Ce rayonnement a permis d'attirer de bons étudiants et de collaborer avec de nombreux partenaires académiques et industriels. Un objectif pour les prochaines années est l'amélioration de cette visibilité au niveau international.

Pour conclure, l'emploi du temps d'un enseignant-chercheur est réparti entre les tâches pédagogiques, les activités de recherche et les responsabilités administratives. Sur l'intégrale de la carrière, un investissement dans toutes les composantes du métier doit être réalisé. Cet équilibre est selon moi un point essentiel de ce métier.

7.2 Perspectives de recherche

Mes travaux de recherche traitent la conception d'algorithmes d'adaptation des systèmes de transmission et leur mise en œuvre sur des architectures reconfigurables avec comme critère principal la faible consommation d'énergie. La démarche repose notamment sur une modélisation hybride des systèmes (de communication et de calcul) alliant expressions analytiques et mesures expérimentales. Nous avons ensuite à cœur de valider ces modèles avec des mises en œuvre complètes des solutions sur des plateformes matérielles. Un point essentiel à explorer dans le futur est la maîtrise de l'énergie consommée par le système. En effet, les systèmes de transmission s'adaptant à leur environnement (retransmission, changement de débit, ...) et les architectures associées étant reconfigurables, l'énergie consommée évolue au cours du temps et sa connaissance est primordiale pour limiter le gaspillage d'énergie avec un dimensionnement au pire cas. Des modélisations fines de cette énergie sont donc nécessaires et peuvent être obtenues par une modélisation hybride.

Une autre solution est l'utilisation de méthodes d'apprentissage pour apprendre au fil de l'exécution les variations de l'écosystème (énergie consommée, énergie récupérée, canal de transmission, ...). Cette approche a été testée partiellement dans ce mémoire pour les réseaux de capteurs à récupération d'énergie. Une prolongation de ces travaux à des réseaux ad-hocs avec des traitements pouvant être distribués dans le réseau permettra l'optimisation globale du réseau.

Enfin, l'utilisation de processeurs de calcul hétérogènes (avec des cœurs logiciels et matériels) pour la radio logicielle ouvre des perspectives intéressantes en termes d'applications et de méthodes de conception associées. Ce type de radio logicielle haute performance est déjà commercialisé (e.g. l'USRP E310 [117] embarque un Zynq 7020) et leur utilisation dans le cadre de nos recherches doit permettre d'adresser de nouvelles applications. Notamment, la reconfiguration temps-réel et sous contrainte de l'application est une problématique ouverte lorsque des architectures hétérogènes et distribuées (dans le cloud et dans les systèmes embarqués) sont envisagées.

Il reste donc un travail considérable pour faire converger toutes nos activités vers une implantation efficace. Si de nombreuses perspectives émanent directement de mes différents axes de recherche, cette partie détaille quatre axes qui seront développés ces prochaines années.

7.2.1 Smart wake-up radio

La faible consommation d'énergie des wake-up radios est souvent due à une conception spécifique dédiée à une application (bande de fréquence, portée, débit, latence, ...). La continuité de nos travaux sur les wake-up radios est de proposer une nouvelle méthodologie pour leur conception et de définir une plateforme matérielle capable de s'adapter à différentes applications. L'utilisation des capacités de calcul ultra faible consommation embarquées dans la wake-up et une gestion de son énergie sont nécessaires afin d'obtenir une wake-up radio intelligente, capable de s'adapter à son environnement et reconfigurable en fonction des besoins applicatifs.

Un défi non négligeable consiste à concevoir des algorithmes d'adaptation efficaces, adaptés aux ressources limitées fournies par les wake-up radios en termes de mémoire, de puissance de calcul et de stockage d'énergie. Un contrôle précis de l'énergie dépensée en fonction des configurations possibles du nœud est donc nécessaire et peut être associé à une modélisation énergétique précise du nœud ou à des méthodes d'apprentissage. D'autre part, de nouveaux protocoles d'accès et de réseau doivent être mis en œuvre pour tirer parti de cette technologie notamment lorsque des réseaux hétérogènes, utilisant plusieurs protocoles radio, sont considérés.

D'autres pistes autour de cette activité peuvent être envisagées. Si la wake-up radio est par définition sensible au canal radio, l'utilisation d'autres événements peut permettre aussi de réveiller le système principal. Par exemple, la reconnaissance de motif spécifique dans l'information issue d'un capteur peut aussi initier une communication. Nous parlons alors de wake-up sensors dont l'association à la wake-up radio peut permettre d'adresser d'autres types d'applications.
Une autre piste intéressante est l'association de la wake-up radio aux techniques de communication backscatter. En effet, un des défauts des wake-up radios est l'utilisation d'un transmetteur principal pour envoyer les trames de réveil. L'utilisation d'un backscatterer [118] permet d'utiliser l'énergie ambiante pour transmettre l'information à un coût énergétique très faible. Le système résultant, recevant et transmettant à très faible énergie, pourrait être alimenté par récupération d'énergie, et utilisé pour des applications RFID.

7.2.2 Near sensor computing

Le problème de la consommation d'énergie pour les systèmes embarqués de l'IoT devient particulièrement critique quand la quantité de données produites augmente considérablement (en utilisant par exemple des capteurs avancés, un microphone, une caméra, ...). La quantité de données a en effet un impact direct sur le nombre de communications et sur la complexité des traitements [119]. Dans ce contexte, une question importante est la localisation des traitements dans le réseau : centralisée ou décentralisée. Dans une approche centralisée, toutes les données détectées sont transmises à un nœud principal alors que la décentralisation vise à inclure un traitement proche des capteurs. Dans la deuxième approche, seules les informations synthétiques ou utiles sont envoyées au nœud principal.

En exécutant des traitements (exploitation partielle, compression) sur les nœuds ou groupes de nœuds voisins, il sera ainsi possible de diminuer la quantité d'information transitant sur le réseau. Cependant, ces traitements peuvent nécessiter un surplus important de calculs embarqués et donc engendrer un coût énergétique non négligeable. Le but de ce travail serait d'étudier le bilan énergétique entre le traitement, la communication et le stockage des réseaux de capteurs sans fil au travers d'une exploration architecturale.

Une approche intéressante serait aussi de caractériser différemment la qualité des données transmises en ajoutant des dimensions inhérentes aux systèmes. L'utilisation de métriques inter-couches (du protocole au matériel) pourrait qualifier la donnée reçue et donc compenser la réduction du nombre d'informations transmises.

7.2.3 Radio logicielle pour l'interception de signaux compromettants

La sécurité des systèmes d'informations est un enjeu crucial à de multiples échelles et la protection des données sensibles (mécanisme de défense) comme l'interception des données (mécanisme d'attaque) présentent des intérêts stratégiques. Ces dernières années, une nouvelle menace à la confidentialité de l'information a émergé avec la détection en amont du processus de cryptage des informations. Cette détection par canaux auxiliaires s'effectue dans l'espace privé des données et doit donc être non-invasive. L'attaque consiste à détecter à distance des signaux parasites compromettants issus de canaux cachés et est illustrée par la Figure 7.1. Ces canaux peuvent être de nature électrique, électromagnétique, mécanique ou acoustique [120]. L'étude des signaux parasites compromettants entre dans le contexte général TEMPEST défini par la National Security Agency (NSA) qui fait référence aux mesures et aux standards liés aux émissions compromettantes des systèmes traitant de l'information [121].

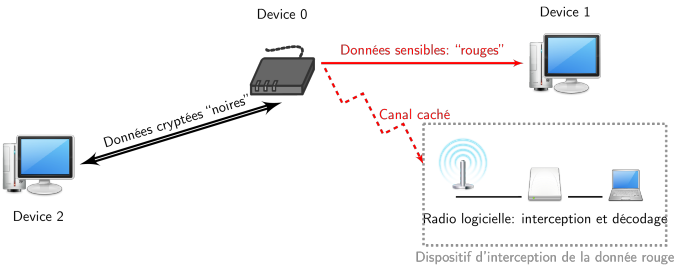


FIGURE 7.1 – Interception de signaux dans un contexte TEMPEST.

Dans un contexte pratique d'interception avec des fortes contraintes temps-réel, il est nécessaire de construire des méthodes qui soient i) embarquées, capables de fonctionner sans devoir nécessairement déporter un flux de données massif vers un centre de traitement ii) rapides, de manière à garantir l'interception des données dans une fenêtre temporelle restreinte iii) reconfigurables, de manière à pouvoir adapter la stratégie d'intervention à la volée. Il y a cependant plusieurs verrous technologiques et pratiques qu'il convient d'adresser pour bâtir des méthodes à la fois embarquées, rapides et reconfigurables. L'objectif de ces travaux est donc triple :

— Concevoir des méthodes de détections rapides et large bande des canaux cachés,
— Développer des méthodes de décodage adaptées à des transmissions non-stationnaires,
— Proposer des méthodes de prototypage rapides qui permettent de déployer les méthodes sur des terminaux reconfigurables, compacts et à faible empreinte énergétique.

7.2.4 Partitionnement efficace en énergie pour les architectures Cloud-RAN

Les réseaux d'accès radio dans le cloud (C-RANs - Cloud Radio Access Networks) ont été introduits pour réduire les coûts de déploiement et de gestion des stations de base [122] [123]. L'approche C-RAN est souvent considérée comme une évolution architecturale du concept de stations de base distribuées, qui permet de séparer l'emplacement physique de la tête radio distante (RRH - Remote Radio Head) de l'unité bande de base numérique (BBU - Digital Baseband Unit). Cela facilite la coordination entre les stations de base potentiellement interférentes, mais permet également le partage des ressources entre les stations de base. L'approche C-RAN est susceptible de réduire les coûts et la consommation d'énergie par rapport aux approches traditionnelles et devrait donc être largement adoptée par les systèmes 5G.

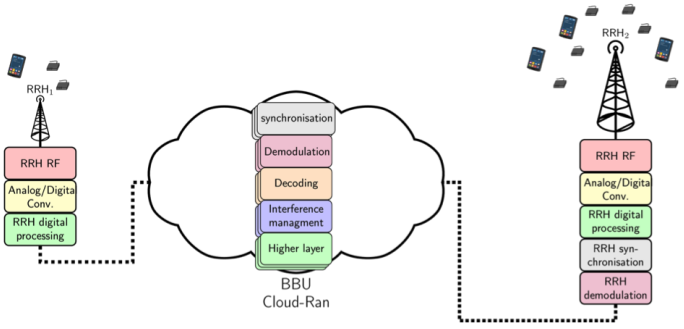


FIGURE 7.2 – Topologie d'une architecture C-RAN.

Cependant, un des principaux défis consiste à déterminer la séparation des traitements entre la station de base et le cloud. Comme illustré dans la Figure 7.2, le partitionnement RRH/BBU peut être différent d'une RRH à une autre. De plus, un partitionnement fonctionnel peut être redistribué dans le temps en fonction de plusieurs critères liés à la consommation d'énergie, à la congestion du réseau, à la latence et au débit. Il faut donc envisager une station de base adaptative capable de faire face à diverses configurations et la mise en œuvre d'un orchestrateur, responsable de la séparation dynamique RRH/BBU.

Afin d'évaluer l'efficacité d'un partitionnement RRH/BBU, il est nécessaire de proposer des modèles précis de la consommation d'énergie, du débit et de la latence, tant au niveau de la RRH que de la BBU. De plus, différentes cibles matérielles de traitement doivent être prises en compte. Ces modèles aideront l'orchestrateur à réaliser un partitionnement efficace et un prototypage rapide.

Publications personnelles

Articles dans revues internationales à comité de lecture indexées JCR :

[aitaoudia18sensors] F. Ait Aoudia, M. Gautier, M. Magno, O. Berder, and L. Benini, "Leveraging Energy Harvesting and Wake-Up Receivers for Long-Term Wireless Sensor Networks," *Sensors*, vol. 5, no. 18, pp. 1–27, May 2018.

[aitaoudia18gcn] F. Ait Aoudia, M. Gautier, and O. Berder, "RLMan : an Energy Manager Based on Reinforcement Learning for Energy Harvesting Wireless Sensor Networks," *IEEE Transactions on Green Communications and Networking*, vol. 2, no. 2, pp. 408–417, February 2018.

[aitaoudia17micpro] F. Ait Aoudia, M. Gautier, M. Magno, M. Le Gentil, O. Berder, and L. Benini, "Long-Short Range Communication Network Leveraging LoRa and Wake-up Receiver," *Microprocessors and Microsystems : Embedded Hardware Design (MICPRO)*, vol. 56, pp. 184 – 192, December 2017.

[aitaoudia17wcl] F. Ait Aoudia, M. Gautier, and O. Berder, "Distributed computation of fair packet rates in energy harvesting wireless sensor networks," *IEEE Wireless Communications Letters*, vol. 6, no. 5, pp. 626–629, October 2017.

[aitaoudia17ton] F. Ait Aoudia, M. Gautier, M. Magno, O. Berder, and L. Benini, "A generic framework for modeling mac protocols in wireless sensor networks," *IEEE/ACM Transactions on Networking*, vol. 5, no. 3, pp. 1489–1500, June 2017.

[aitaoudia16js] F. Ait Aoudia, M. Gautier, and O. Berder, "OPWUM : Opportunistic MAC Protocol Leveraging Wake-Up Receivers in WSNs," *Journal of Sensors*, vol. 2016, pp. 1–9, 2016.

[ouedraogo14jasp] G.-S. Ouedraogo, M. Gautier, and O. Sentieys, "A frame-based Domain-Specific Language for rapid prototyping of FPGA-based Software-Defined Radios," *EURASIP Journal on Advances in Signal Processing*, vol. 2014, no. 1, p. 164, November 2014.

[noguet11jwcn] D. Noguet, M. Gautier, and V. Berg, "Advances in Opportunistic Radio Technologies for TVWS," *EURASIP Journal on Wireless Communications and Networking*, vol. 2011, no. 1, p. 170, November 2011. **Best Paper Award**.

[gautier10ijmwt] M. Gautier, G. Villemaud, and I. Burciu, "The multi-antenna code multiplexing front-end : theory and performance," *International Journal of Microwave and Wireless Technologies*, vol. 2, pp. 515–522, december 2010.

[burciu10ijmwt] I. Burciu, G. Villemaud, J. Verdier, and M. Gautier, "Low power front-end architecture dedicated to the multistandard simultaneous reception," *International Journal of Microwave and Wireless Technologies*, vol. 2, pp. 505–514, december 2010.

[gautier07annals] M. Gautier and J. Lienard, "Application de la modulation multiporteuse par paquets d'ondelettes aux transmissions sans fil," vol. 62, no. 7-8, pp. 871–893, July/August 2007.

Conférences internationales avec actes et comité de lecture :

[djidi18icecs] N.E.H. Djidi, A. Courtay, M. Gautier, and O. Berder, "Adaptive relaying for wireless sensor networks leveraging wake-up receiver," in *IEEE International Conference on Electronics Circuits and Systems (ICECS)*, Bordeaux, France, Dec. 2018.

[gleonec18ict] P.-D. Gleonec, J. Ardouin, M. Gautier, and O. Berder, "A Real-World Evaluation of Energy Budget Estimation Algorithms for Autonomous Long Range IoT Nodes," in *International Conference on Telecommunications (ICT)*, Saint-Malo, France, June 2018.

[le18ict] X.-C. Le, B. Vrigneau, M. Gautier, M. Mabon and O. Berder, "Energy/Reliability Trade-off of LoRa Communications over Fading Channels," in *International Conference on Telecommunications (ICT)*, Saint-Malo, France, June 2018.

[kazdoghli18vtc] M. Kazdoghli Lagha, F. Ait Aoudia, M. Gautier, and O. Berder, "Feature Selection Framework for Multi-source Energy Harvesting Wireless Sensor Networks," in *IEEE Vehicular Technology Conference (VTC-Spring)*, Porto, Portugal, June 2018.

[magno17date] M. Magno, F. Ait Aoudia, M. Gautier, O. Berder, and L. Benini, "WULoRa : An Energy Efficient IoT End-Node for Energy Harvesting and Heterogeneous Communication," in *IEEE/ACM Design, Automation & Test in Europe Conference & Exhibition (DATE)*, Lausanne, Switzerland, March 2017.

[aitaoudia17icc] F. Ait Aoudia, M. Gautier, and O. Berder, "Learning to Survive : Achieving Energy Neutrality in Wireless Sensor Networks Using Reinforcement Learning," in *IEEE International Conference on Communications (ICC)*, Paris, France, May 2017.

[aitaoudia16scube] F. Ait Aoudia, M. Gautier, M. Magno, O. Berder, and L. Benini, "SNW-MAC : an Asynchronous Protocol Leveraging Wake-up Receivers for Data Gathering in Star Networks," in *EAI International Conference on Sensor Systems and Software (S-cube)*, Nice, France, Dec. 2016.

[nguyen16globecom] T. Nguyen-Ti, M. Gautier, P. Scalart, O. Berder, T.-H. Nguyen, and F. Ait Aoudia, "Blind I/Q Imbalance Compensation for M-QAM Optical Coherent Systems based on Pseudo-rotation," in *IEEE Global Communications Conference (Globecom)*, Washington, United States, Dec. 2016.

[aitaoudia16globecom] F. Ait Aoudia, M. Magno, M. Gautier, O. Berder, and L. Benini, "Analytical and Experimental Evaluation of Wake-up Receivers based protocols," in *IEEE Global Communications Conference (Globecom)*, Washington, United States, Dec. 2016.

[roux16mcsoc] B. Roux, M. Gautier, O. Sentieys, and S. Derrien, "Communication-Based Power Modelling for Heterogeneous Multiprocessor Architecture," in *IEEE 10th International Symposium on Embedded Multicore /Many-core Systems-on-Chip (MCSoC)*, Lyon, France, Sep. 2016.

[gleonec16ogc] P.-D. Gleonec, J. Ardouin, M. Gautier, and O. Berder, "Multi-source Energy Harvesting for IoT nodes," in *IEEE Online Conference on Green Communications (Online GreenComm)*, n/a, Nov. 2016.

[aitaoudia16dsd] F. Ait Aoudia, M. Magno, M. Gautier, O. Berder, and L. Benini, "A Low Latency and Energy Efficient Communication Architecture for Heterogeneous Long-Short Range Communication," in *Euromicro Conference on Digital System Design (DSD)*, Limassol, Cyprus, Aug. 2016.

[aitaoudia16icc] F. Ait Aoudia, M. Gautier, and O. Berder, "Fuzzy Power Management for Energy Harvesting Wireless Sensor Nodes," in *IEEE International*

Conference on Communications (ICC), Kuala Lumpur, Malaysia, May 2016, pp. 1 – 6.

[nguyen16icc] T. H. Nguyen, P. Scalart, M. Gay, L. Bramerie, C. Peucheret, T. Nguyen-Ti, M. Gautier, O. Sentieys, J.-C. Simon, and M. Joindot, "Blind Adaptive Transmitter IQ Imbalance Compensation in M-QAM Optical Coherent Systems," in *IEEE International Conference on Communications (ICC)*, Kuala Lumpur, Malaysia, May 2016, pp. 1 – 6.

[tran16crowncom] M.-T. Tran, M. Gautier, and E. Casseau, "On the FPGA-based implementation of a flexible waveform from a high-level description : Application to LTE FFT case study," in *EAI International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, Grenoble, France, May 2016, pp. 1–12.

[nguyen16crowncom] V.-T. Nguyen, M. Gautier, and O. Berder, "FTA-MAC : Fast Traffic Adaptive energy efficient MAC protocol for Wireless Sensor Networks," in *EAI International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, Grenoble, France, May 2016, pp. 1–12.

[aitaoudia15pimrc] F. Ait Aoudia, M. Gautier, and O. Berder, "GRAPMAN : Gradual power manager for consistent throughput of energy harvesting wireless sensor nodes," in *IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Hong Kong, China, September 2015, pp. 954 – 959.

[cladera15isvlsi] F. Cladera, M. Gautier, and O. Sentieys, "Energy-Aware Computing via Adaptive Precision under Performance Constraints in OFDM Wireless Receivers," in *IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, Montpellier, France, July 2015, pp. 591–596. **Best Paper Award**.

[cladera15eusipco] F. Cladera, M. Gautier, and O. Sentieys, "Channel-Aware Energy Optimization of OFDM Receivers Using Dynamic Precision Scaling in FPGAs," in *EURASIP European Signal Processing Conference (EUSIPCO)*, Nice, France, August 2015, pp. 1–5.

[hoang15icc] G.-M. Hoang, M. Gautier, and A. Courtay, "Cooperative-cum-Constrained Maximum Likelihood algorithm for UWB-based localization in wireless BANs," in *IEEE International Conference on Communications (ICC)*, London, UK, June 2015, pp. 2704–2709.

[nguyen14omnet] V.-T. Nguyen, M. Gautier, and O. Berder, "Implementation of an adaptive energy-efficient MAC protocol in OMNeT++/MiXiM," in *1st OMNeT++ Community Summit*, Hamburg, Germany, September 2014, pp. 1–4.

[gautier14dsd] M. Gautier, G. Ouedraogo, and O. Sentieys, "Design Space Exploration in an FPGA-Based Software Defined Radio," in *Euromicro Conference on Digital System Design (DSD)*, Verona, Italy, August 2014, pp. 22–27.

[nhan14crowncom] N.-Q. Nhan, M. Gautier, and O. Berder, "Asynchronous MAC Protocol for Spectrum Agility in Wireless Body Area Sensor Networks," in *EAI International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, Oulu, Finland, June 2014, pp. 203–208.

[ouedraogo14crowncom] G. Ouedraogo, M. Gautier, and O. Sentieys, "Frame-based Modeling for Automatic Synthesis of FPGA-Software Defined Radio," in *EAI International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, Oulu, Finland, June 2014, pp. 341–346.

[bhatnagar13vtc] V. Bhatnagar, G. Ouedraogo, M. Gautier, A. Carer, and O. Sentieys, "An FPGA Software Defined Radio Platform with a High-Level Synthesis Design Flow," in *IEEE Vehicular Technology Conference (VTC-Spring)*, Dresden, Germany, June 2013, pp. 1–5.

[gautier13vtc] M. Gautier and D. Noguet, "Signal Detection Using Watermark Insertion," in *IEEE Vehicular Technology Conference (VTC-Spring)*, Dresden, Germany, June 2013, pp. 1–5.

[gautier12crowncom] M. Gautier, V. Berg, and D. Noguet, "Wideband frequency domain detection using Teager-Kaiser energy operator," in *EAI International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, Stockholm, Sweden, June 2012, pp. 332–337.

[datta11crowncom] R. Datta, M. Gautier, V. Berg, Y. Futatsugi, M. Ariyoshi, M. Schuhler, Z. Kollar, P. Horvath, D. Noguet, and G. Fettweis, "Flexible multicarrier PHY design for cognitive radio in White Space," in *EAI International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, Osaka, Japan, June 2011, pp. 141–145.

[gautier11cocora] M. Gautier, M. Laugeois, and P. Hostiou, "Cyclostationarity Detection of DVB-T Signal : Testbed and Measurement," in *International Conference on Advances in Cognitive Radio (COCORA)*, Budapest, Hungary, April 2011. **Best Paper Award**.

[noguet11vitae] D. Noguet, R. Datta, P. Lehne, M. Gautier, and G. Fettweis, "TVWS regulation and QoSMOS requirements," in *IEEE International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace Electronic Systems Technology (Wireless VITAE)*, India, February 2011, pp. 1–5.

[burciu11rws] I. Burciu, M. Gautier, G. Villemaud, and J. Verdier, "Candidate architecture for MIMO LTE-Advanced receivers with multiple channels capabilities and reduced complexity and cost," in *IEEE Radio and Wireless Symposium (RWS)*, Phoenix, USA, January 2011, pp. 263–266.

[gautier10crowncom] M. Gautier, M. Laugeois, and D. Noguet, "Teager-Kaiser energy detector for narrowband wireless microphone spectrum sensing," in *EAI International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, Cannes, France, June 2010, pp. 1–5.

[gautier09vtc] M. Gautier, P.-F. Morlat, and G. Villemaud, "IQ Imbalance Reduction in a SMI Multi-Antenna Receiver by Using a Code Multiplexing Front-End," in *IEEE Vehicular Technology Conference (VTC-Spring)*, Barcelona, Spain, April 2009, pp. 1–5.

[burciu09vtc] I. Burciu, G. Villemaud, J. Verdier, and M. Gautier, "A 802.11g and UMTS Simultaneous Reception Front-End Architecture using a Double IQ Structure," in *IEEE Vehicular Technology Conference (VTC-Spring)*, Barcelona, Spain, April 2009, pp. 1–5.

[gautier09wcnc] M. Gautier and G. Villemaud, "Low Complexity Antenna Diversity Front-End : Use of Code Multiplexing," in *IEEE Wireless Communications and Networking Conference (WCNC)*, Budapest, Hungary, April 2009, pp. 1–6.

[gautier09eucap] M. Gautier, I. Burciu, and G. Villemaud, "New antenna diversity front-end using code multiplexing," in *European Conference on Antennas and Propagation (Eucap)*, Berlin, Germany, March 2009.

[gautier08isccsp] M. Gautier, C. Lereau, M. Arndt, and J. Lienard, "PAPR analysis in wavelet packet modulation," in *International Symposium on Communications, Control and Signal Processing (ISCCSP)*, Malta, March 2008, pp. 799–803.

[gautier07aict] M. Gautier, M. Arndt, and J. Lienard, "Efficient wavelet packet modulation for wireless communication," in *Advanced International Conference on Telecommunications (AICT)*, Mauritius, May 2007, pp. 1–5.

[gautier06norsig] M. Gautier and J. Lienard, "Performances of Complex Wavelet Packet Based Multicarrier Transmission Through Double Dispersive Channel," in *IEEE Nordic Signal Processing Symposium (NORSIG)*, Iceland, June 2006, pp. 74–77.

Conférences nationales avec actes et comité de lecture :

[aitaoudia18cores] F. Ait Aoudia, M. Gautier, and O. Berder, "Evaluation analytique des wake-up radios pour les réseaux de capteurs sans fil," in *Rencontres Francophones sur la Conception de Protocoles, l'Évaluation de Performance et l'Expérimentation des Réseaux de Communication (CoRes)*, Roscoff, France, Mai 2018.

[aitaoudia17gretsi] F. Ait Aoudia, M. Gautier, and O. Berder, "Modélisation logicielle/matérielle par chaînes de markov absorbantes des protocoles pour les réseaux de capteurs sans-fil," in *26ème colloque GRETSI sur le traitement du signal et des images*, Juan-les-pins, France, Septembre 2017.

[aitaoudia16ursi] F. Ait Aoudia, M. Gautier, and O. Berder, "Comment bénéficier de la wake-up radio pour les réseaux de capteurs à récupération d'énergie ?" in *Actes des Journées Scientifiques URSI France 2016 - Energie et radiosciences*, Rennes, France, Mars 2016, pp. 1 – 6.

[aitaoudia15gretsi] F. Ait Aoudia, M. Gautier, and O. Berder, "Adaptation dynamique de la qualité de service dans les réseaux de capteurs avec récupération d'énergie," in *25ème colloque GRETSI sur le traitement du signal et des images*, Lyon, France, Septembre 2015.

[gautier15gretsi] M. Gautier, A. Courtay, and G.-M. Hoang, "Estimation cooperative et sous contraintes pour la localisation dans les réseaux de capteurs corporels," in *25ème colloque GRETSI sur le traitement du signal et des images*, Lyon, France, Septembre 2015.

[ouedraogo13gretsi] G. S. Ouedraogo, M. Gautier, and O. Sentieys, "Description haut niveau de formes d'ondes pour la radio logicielle sur architectures reconfigurables," in *24ème colloque GRETSI sur le traitement du signal et des images*, Brest, France, Septembre 2013.

[gautier09gretsi] M. Gautier, I. Burciu, and G. Villemaud, "Analyse du PAPR pour un récepteur multi-antennes à multiplexage par code," in *22ème colloque GRETSI sur le traitement du signal et des images*, Dijon, France, Septembre 2009.

[gautier09jnm] ——, "Frontal d'un récepteur à diversité d'antenne," in *16ème Journée Nationale Microondes (JNM)*, Grenoble, France, Mai 2009.

[gautier07gretsi] M. Gautier, M. Arndt, and J. Lienard, "Nouvelles formes d'ondes par paquets d'ondelettes pour les modulations multiporteuses," in *21ème colloque GRETSI sur le traitement du signal et des images*, Troyes, France, Septembre 2007.

Ouvrages collectifs :

[gautier11crcpress] M. Gautier, E. Casseau, H. Yviquel, G.-S. Ouedraogo, M. Raulet, and O. Sentieys, "Rapid Prototyping for Video Coding over Flexible Radio Links," in *Multimedia over Cognitive Radio Networks : Algorithms, Protocols, and Experiments*, Ed. CRC Press, December 2014.

[gautier11cambridge] M. Gautier, G. Villemaud, T. Risset, C. Lévy-Bencheton, and N. D., "Design of digital front-End for flexible wireless link," in *Digital Front-End in Wireless Communications and Broadcasting : Circuits and Signal Processing*, Ed. Cambridge University Press, November 2011.

[villemaud11cambridge] G. Villemaud, J. Verdier, M. Gautier, I. Burciu, and P.-F. Morlat, "Front-end Architectures and Impairment corrections in multimode and multi-antenna systems," in *Digital Front-End in Wireless Communications and Broadcasting : Circuits and Signal Processing*, Ed. Cambridge University Press, November 2011.

Brevets :

[gautier_berg13patent] M. Gautier and V. Berg, "Process and device for detection of a frequency sub-band in a frequency band and communications equipment comprising such a device," US Patent, US9 008 708 B2, May, 2013.

[gautier_noguet13patent] M. Gautier and D. Noguet, "Method for Identifyng and Detecting a Radio Signal For a Cognitive Communication System," US Patent, US 8 964 910 B2, March, 2013.

[burciu10patent] I. Burciu, M. Gautier, G. Villemaud, and J. Verdier, "Method for Processing Two Signals Received by a Single Terminal," International Patent, WO/2010/031 944, March, 2010.

Autres publications (démonstrations, posters, colloques) :

[gleonec18ict_demo] P.-D. Gleonec, J. Ardouin, M. Gautier, and O. Berder, "Demo Abstract : Multi-source energy management for Long Range IoT nodes," in *International Conference on Telecommunications (ICT)*, June 2018.

[legentil17sips] M. Le Gentil, F. Ait Aoudia, M. Gautier, and O. Berder, "Demo Abstract : Online QoS Adaptation using Fuzzy Control for Indoor Light Energy Harvesting in Wireless Sensor Networks," in *IEEE International Workshop on Signal Processing Systems (SiPS)*, Oct. 2017.

[roux17fccm] B. Roux, M. Gautier, O. Sentieys, and J.-P. Delahaye, "Poster : Fast and Energy-Driven Design Space Exploration for Heterogeneous Architectures," in *IEEE Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM)*, May 2017, poster.

[aitaoudia17gdrsocsip] F. Ait Aoudia, M. Gautier, A. Courtay, and O. Berder, "Etude du compromis latence-consommation d'énergie des radios longue portée," in *Colloque du GDR SoC-SiP, Juin 2017, Bordeaux, France*, Juin 2017, poster.

[tran16dasip] M.-T. Tran, E. Casseau, and M. Gautier, "Demo abstract : FPGA-based implementation of a flexible FFT dedicated to LTE standard," in *Conference on Design and Architectures for Signal and Image Processing (DASIP), Demo Night*, Oct. 2016.

[aitaoudia16dasip] F. Ait Aoudia, M. Gautier, M. Le Gentil, and O. Berder, "Demo Abstract : How Fuzzy Logic can enhance Energy Management in autonomous Wireless Sensor Nodes ?" in *Conference on Design and Architectures for Signal and Image Processing (DASIP), Demo Night*, Oct. 2016.

[gleonec16dasip] P.-D. Gleonec, J. Ardouin, M. Gautier, and O. Berder, "Poster : Multi-source Energy Harvesting for IoT nodes," in *Conference on Design and Architectures for Signal and Image Processing (DASIP)*, Oct. 2016, poster.

[gleonec16gdrsocsip] P.-D. Gleonec, J. Ardouin, M. Gautier, and O. Berder, "Système de commutation de sources d'énergie pour capteurs sans-fil," in *Colloque du GDR SoC-SiP*, Juin 2016, poster.

[aitaoudia16gdrsocsip] F. Ait Aoudia, M. Gautier, and O. Berder, "Protocoles de communication faibles latences et faibles consommations à base de wake-up radio," in *Colloque du GDR SoC-SiP*, Juin 2016, poster.

[aitaoudia16ipsn] F. Ait Aoudia, M. Magno, M. Gautier, O. Berder, and L. Benini, "Poster Abstract : Wake-Up Receivers for Energy Efficient and Low Latency Communication," in *ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, Apr. 2016, poster.

[ouedraogo13gdrsocsip] G. S. Ouedraogo, M. Gautier, and O. Sentieys, "Vers un language spécialisé pour la radio logicielle sur FPGA," in *Colloque national du GDR SoC-SiP*, Lyon, France, Juin 2013, poster.

[burciu09cost2100] I. Burciu, M. Gautier, G. Villemaud, and J. Verdier, "Candidate Architecture for MIMO LTE-Advanced Receivers with Multiple Channels Capabilities and Reduced Complexity and Cost," in *COST2100 Pervasive Mobile & Ambient Wireless Communications*, Valancia, Spain, May 2009.

[burciu08cost2100] I. Burciu, G. Villemaud, J. Verdier, and M. Gautier, "A Multistandard Simultaneous Reception Front-End Architecture," in *COST2100 Pervasive Mobile & Ambient Wireless Communications*, Lille, France, Oct. 2008.

[gautier06see] M. Gautier and J. Lienard, "Performance de la modulation multiporteuses par paquets d'ondelettes en présence d'un brouilleur bande étroite," in *Journées SEE - Nouvelles formes d'ondes agiles en imagerie, localisation et communication*, Paris, France, Mars 2006.

Références

[1] Nokia Solutions and Networks, "Looking ahead to 5G," *White paper*, 2013.

[2] A. M. Ionescu and C. Hierold, "Guardian Angels for a Smarter Life : Enabling a Zero-Power Technological Platform for Autonomous Smart Systems," in *European Future Technologies Conference and Exhibition 2011 (FET 11)*, vol. 7, Oct. 2011, pp. 43 – 46.

[3] R. J. M. Vullers, R. v. Schaijk, H. J. Visser, J. Penders, and C. V. Hoof, "Energy Harvesting for Autonomous Wireless Sensor Networks," *IEEE Solid-State Circuits Magazine*, vol. 2, no. 2, pp. 29–38, Spring 2010.

[4] P. Pirinen, "A brief overview of 5G research activities," in *International Conference on 5G for Ubiquitous Connectivity*, Nov. 2014, pp. 17 – 22.

[5] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What Will 5G Be ?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[6] G. P. Fettweis, "The Tactile Internet : Applications and Challenges," *IEEE Vehicular Technology Magazine*, vol. 9, no. 1, pp. 64–70, March 2014.

[7] T. Rault, A. Bouabdallah, and Y. Challal, "Energy Efficiency in Wireless Sensor Networks : A Top-Down Survey," *Computer Networks*, vol. 67, pp. 104–122, July 2014.

[8] Y. R. Thomas, M. Picot, A. Carer, O. Berder, O. Sentieys, and F. Barrière, "A single sediment-microbial fuel cell powering a wireless telecommunication system," *Journal of Power Sources*, vol. 241, pp. 703 – 708, 2013.

[9] P. Huang, L. Xiao, S. Soltani, M. W. Mutka, and N. Xi, "The Evolution of MAC Protocols in Wireless Sensor Networks : A Survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 1, pp. 101–120, January 2013.

[10] A. Bachir, M. Dohler, T. Watteyne, and K. Leung, "MAC Essentials for Wireless Sensor Networks," *IEEE Communications Surveys & Tutorials*, vol. 12, no. 2, pp. 222–248, Second 2010.

[11] N. S. Mazloum and O. Edfors, "Performance Analysis and Energy Optimization of Wake-Up Receiver Schemes for Wireless Low-Power Applications," *IEEE Transactions on Wireless Communications*, vol. 13, no. 12, pp. 7050–7061, December 2014.

[12] Y. Ammar, S. Bdiri, and F. Derbel, "An Ultra-Low Power Wake Up Receiver with Flip Flops Based Address Decoder," in *International Multi-Conference on Systems, Signals and Devices (SSD)*, March 2015, pp. 1–5.

[13] J. Oller, I. Demirkol, J. Casademont, and J. Paradells, "Design, Development, and Performance Evaluation of a Low-cost, Low-power Wake-up Radio System for Wireless Sensor Networks," *ACM Transaction on Sensor Networks*, vol. 10, no. 1, December 2013.

[14] L. Chen, J. Warner, P. L. Yung, D. Zhou, W. Heinzelman, I. Demirkol, U. Muncuk, K. Chowdhury, and S. Basagni, "REACH2-Mote : A Range-Extending Passive Wake-Up Wireless Sensor Node," *ACM Transactions On Sensor Networks*, vol. 11, no. 4, December 2015.

[15] D. Spenza, M. Magno, S. Basagni, L. Benini, M. Paoli, and C. Petrioli, "Beyond Duty Cycling : Wake-up Radio with Selective Awakenings for Long-Lived Wireless Sensing Systems," in *IEEE International Conference on Computer Communications (INFOCOM)*, April 2015, pp. 522–530.

[16] F. Sutton, B. Buchli, J. Beutel, and L. Thiele, "Zippy : On-Demand Network Flooding," in *ACM Conference on Embedded Networked Sensor Systems (SenSys)*, November 2015, pp. 45–58.

[17] A. Kansal, J. Hsu, S. Zahedi, and M. B. Srivastava, "Power Management in Energy Harvesting Sensor Networks," *ACM Transactions on Embedded Computing Systems*, vol. 6, no. 4, September 2007.

[18] C. M. Vigorito, D. Ganesan, and A. G. Barto, "Adaptive Control of Duty Cycling in Energy-Harvesting Wireless Sensor Networks," in *IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks (SECON)*, June 2007.

[19] A. Castagnetti, A. Pegatoquet, C. Belleudy, and M. Auguin, "A Framework for Modeling and Simulating Energy Harvesting WSN Nodes with Efficient Power Management Policies," *EURASIP Journal on Embedded Systems*, no. 1, October 2012.

[20] T. N. Le, A. Pegatoquet, O. Berder, and O. Sentieys, "Energy-Efficient Power Manager and MAC Protocol for Multi-Hop Wireless Sensor Networks Powered by Periodic Energy Harvesting Sources," *IEEE Sensors Journal*, vol. 15, no. 12, pp. 7208–7220, 2015.

[21] C. Renner, S. Unterschütz, V. Turau, and K. Römer, "Perpetual Data Collection with Energy-Harvesting Sensor Networks," *ACM Transactions on Sensor Networks*, vol. 11, no. 1, September 2014.

[22] S. Peng and C. P. Low, "Prediction Free Energy Neutral Power Management for Energy Harvesting Wireless Sensor Nodes," *Ad Hoc Networks*, vol. 13, Part B, 2014.

[23] T. N. Le, O. Sentieys, O. Berder, A. Pegatoquet, and C. Belleudy, "Power Manager with PID Controller in Energy Harvesting Wireless Sensor Networks," in *IEEE International Conference on Green Computing and Communications (GreenCom)*, 2012.

[24] S. Yang, X. Yang, J. A. McCann, T. Zhang, G. Liu, and Z. Liu, "Distributed Networking in Autonomic Solar Powered Wireless Sensor Networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 12, pp. 750–761, December 2013.

[25] R. C. Hsu, C. T. Liu, and H. L. Wang, "A Reinforcement Learning-Based ToD Provisioning Dynamic Power Management for Sustainable Operation of Energy Harvesting Wireless Sensor Node," *IEEE Transactions on Emerging Topics in Computing*, vol. 2, no. 2, pp. 181–191, June 2014.

[26] J. R. Piorno, C. Bergonzini, D. Atienza, and T. S. Rosing, "Prediction and Management in Energy Harvested Wireless Sensor Nodes," in *International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace Electronic Systems Technology (Wireless VITAE)*, May 2009.

[27] F. K. Shaikh and S. Zeadally, "Energy harvesting in wireless sensor networks : A comprehensive review," *Renewable and Sustainable Energy Reviews*, vol. 55, pp. 1041 – 1054, 2016.

[28] G. M. d. Araújo, A. R. Pinto, J. Kaiser, and L. B. Becker, *Genetic Machine Learning Approach for Link Quality Prediction in Mobile Wireless Sensor Networks.* Berlin, Heidelberg : Springer Berlin Heidelberg, 2014, pp. 1–18.

[29] C. Bergonzini, D. Brunelli, and L. Benini, "Algorithms for Harvested Energy Prediction in Batteryless Wireless Sensor Networks," in *International Workshop on Advances in sensors and Interfaces*, June 2009, pp. 144–149.

[30] A. Cammarano, C. Petrioli, and D. Spenza, "Pro-Energy : A novel energy prediction model for solar and wind energy-harvesting wireless sensor networks," in *IEEE International Conference on Mobile Adhoc and Sensor Systems (MASS)*, October 2012, pp. 75–83.

[31] G. Chen and T. T. Pham, *Introduction to Fuzzy Sets, Fuzzy Logic, and Fuzzy Control Systems.* CRC Press, 2000.

[32] "PowWow : Power Optimized Hardware and Software FrameWork for Wireless Motes," http://powwow.gforge.inria.fr/, 2014, [Online ; accessed January-2019].

[33] J. Paradiso and T. Starner, "Energy Scavenging for Mobile and Wireless Electronics," *IEEE Pervasive Computing*, vol. 4, no. 1, pp. 18–27, January 2005.

[34] C. Ó. Mathúna, T. O'Donnell, R. V. Martinez-Catala, J. Rohan, and B. O'Flynn, "Energy scavenging for long-term deployable wireless sensor networks," *Talanta*, vol. 75, no. 3, pp. 613 – 623, 2008, Special Section : Remote Sensing.

[35] D. Benedetti, C. Petrioli, and D. Spenza, "GreenCastalia : An Energy-harvesting-enabled Framework for the Castalia Simulator," in *ACM International Workshop on Energy Neutral Sensing Systems (ENSsys)*, 2013.

[36] A. Boulis, "Castalia : Revealing Pitfalls in Designing Distributed Algorithms in WSN," in *ACM International Conference on Embedded Networked Sensor Systems (SenSys)*, 2007.

[37] M. A. Green, K. Emery, Y. Hishikawa, W. Warta, and E. D. Dunlop, "Solar cell efficiency tables (version 45)," *Progress in Photovoltaics : Research and Applications*, vol. 23, no. 1, pp. 1–9, 2015.

[38] G. Anastasi, M. Conti, M. D. Francesco, and A. Passarella, "Energy conservation in wireless sensor networks : A survey," *Ad Hoc Networks*, vol. 7, no. 3, pp. 537 – 568, 2009.

[39] X. Fafoutis and N. Dragoni, "Analytical Comparison of MAC Schemes for Energy Harvesting - Wireless Sensor Networks," in *International Conference on Networked Sensing Systems (INSS)*, June 2012.

[40] H. Van Hasselt, "Reinforcement Learning in Continuous State and Action Spaces," in *Reinforcement Learning.* Springer, 2012, pp. 207–251.

[41] I. Grondman, L. Busoniu, G. A. D. Lopes, and R. Babuska, "A Survey of Actor-Critic Reinforcement Learning : Standard and Natural Policy Gradients," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 6, pp. 1291–1307, November 2012.

[42] R. S. Sutton, D. A. McAllester, S. P. Singh, and Y. Mansour, "Policy Gradient Methods for Reinforcement Learning with Function Approximation." in *Advances in Neural Information Processing Systems 12*, vol. 99, 2000, pp. 1057–1063.

[43] R. S. Sutton, "Learning to Predict by the Methods of Temporal Differences," *Machine Learning*, vol. 3, no. 1, pp. 9–44, 1988.

[44] O. Berder and O. Sentieys, "PowWow : Power Optimized Hardware/Software Framework for Wireless Motes," in *International Conference on Architecture of Computing Systems (ARCS)*, February 2010.

[45] M. Gorlatova, A. Wallwater, and G. Zussman, "Networking Low-Power Energy Harvesting Devices : Measurements and Algorithms," in *IEEE International Conference on Computer Communications (INFOCOM)*, April 2011.

[46] "NREL : MIDC - National Wind Technology Center," http ://www.nrel.gov/midc/nwtc_m2/.

[47] J. Oller, I. Demirkol, J. Casademont, J. Paradells, G. U. Gamm, and L. Reindl, "Has Time Come to Switch From Duty-Cycled MAC Protocols to Wake-Up Radio for Wireless Sensor Networks ?" *IEEE/ACM Transactions on Networking*, vol. 24, no. 2, pp. 674–687, April 2016.

[48] R. Piyare, A. L. Murphy, C. Kiraly, P. Tosato, and D. Brunelli, "Ultra Low Power Wake-Up Radios : A Hardware and Networking Survey," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2117–2157, 2017.

[49] D. Ghose, F. Y. Li, and V. Pla, "MAC Protocols for Wake-Up Radio : Principles, Modeling and Performance Analysis," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 5, pp. 2294–2306, May 2018.

[50] F. Z. Djiroun and D. Djenouri, "MAC Protocols With Wake-Up Radio for Wireless Sensor Networks : A Review," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 587–618, Firstquarter 2017.

[51] M. Magno and L. Benini, "An Ultra Low Power High Sensitivity Wake-up Radio Receiver with Addressing Capability," in *IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, October 2014.

[52] C. M. Grinstead and J. L. Snell, *Grinstead and Snell's Introduction to Probability.* Chance Project, 2006.

[53] E.-Y. Lin, J. Rabaey, and A. Wolisz, "Power-efficient Rendez-vous Schemes for Dense Wireless Sensor Networks," in *IEEE International Conference on Communications (ICC)*, June 2004, pp. 3769–3776.

[54] L. Tang, Y. Sun, O. Gurewitz, and D. B. Johnson, "PW-MAC : An Energy-Efficient Predictive-Wakeup MAC Protocol for Wireless Sensor Networks," in *IEEE International Conference on Computer Communications (INFOCOM)*, April 2011, pp. 1305–1313.

[55] M. Buettner, G. V. Yee, E. Anderson, and R. Han, "X-MAC : A Short Preamble MAC Protocol for Duty-cycled Wireless Sensor Networks," in *International Conference on Embedded Networked Sensor Systems (SenSys)*, Oct. 2006.

[56] T. Watteyne, A. Bachir, M. Dohler, D. Barthel, and I. Auge-Blum, "1-hopMAC : An Energy-Efficient MAC Protocol for Avoiding 1 -hop Neighborhood Knowledge," in *IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks (SECON)*, vol. 2, Sept. 2006, pp. 639–644.

[57] T. N. Le, M. Magno, A. Pegatoquet, O. Berder, O. Sentieys, and E. Popovici, "Ultra Low Power Asynchronous MAC Protocol Using Wake-up Radio for Energy Neutral WSN," in *ACM International Workshop on Energy Neutral Sensing Systems (ENSsys)*, November 2013.

[58] M. Magno, V. Jelicic, B. Srbinovski, V. Bilas, E. Popovici, and L. Benini, "Design, Implementation, and Performance Evaluation of a Flexible Low-Latency Nanowatt Wake-Up Radio Receiver," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 2, pp. 633–644, April 2016.

[59] K. Klues, G. Hackmann, O. Chipara, and C. Lu, "A Component-based Architecture for Power-efficient Media Access Control in Wireless Sensor Networks," in *International Conference on Embedded Networked Sensor Systems (SenSys)*, November 2007.

[60] T. N. Le, A. Pegatoquet, O. Berder, O. Sentieys, and A. Carer, "Energy-Neutral Design Framework for Supercapacitor-Based Autonomous Wireless Sensor Networks," *ACM Journal on Emerging Technologies in Computing Systems (JETC)*, vol. 12, no. 2, pp. 19 :1–19 :21, September 2015.

[61] A. Kailas, D. Brunelli, and M. A. Weitnauer, "Comparison of Energy Update Models for Wireless Sensor Nodes with Supercapacitors," in *ACM International Workshop on Energy Neutral Sensing Systems (ENSsys)*, 2013.

[62] "LoRaWAN^{TM}, Specification v1.0, LoRa Alliance, Inc. 2400 Camino Ramon, Suite 375 San Ramon, CA 94583 (2015)," LoRa Alliance, Tech. Rep., 2015.

[63] C. Goursaud and J.-M. Gorce, "Dedicated networks for IoT : PHY / MAC state of the art and challenges," *EAI endorsed transactions on Internet of Things*, October 2015.

[64] J. Palicot, *De la radio logicielle à la radio intelligente.* Hermes Science Publications, 2010.

[65] J. Mitola, "Software radios : Survey, Critical Evaluation and Future Directions," *IEEE Aerospace and Electronic Systems Magazine, 8(4) :25-36*, April 1993.

[66] J. Mitola III and G. Q. Maguire Jr, "Cognitive Radio : Making Software Radios More Personal," *IEEE Personal Communications*, vol. 6, no. 4, pp. 13–18, 1999.

[67] A. D. Stefano, G. Fiscelli, and C. G. Giaconia, "An FPGA-Based Software Defined Radio Pplatform for the 2.4 GHz ISM Band," *Research in Microelectronics and Electronics, pp.73,76*, 2006.

[68] T. Ulversoy, "Software Defined Radio : Challenges and Opportunities," *IEEE Communications Surveys & Tutorials*, vol. 12, no. 4, pp. 531–550, Fourth 2010.

[69] M. Dardaillon, K. Marquet, T. Risset, and A. Scherrer, "Software Defined Radio Architecture Survey for Cognitive Testbeds," in *IEEE International Wireless Communications and Mobile Computing Conference (IWCMC)*, August 2012.

[70] O. Anjum, T. Ahonen, F. Garzia, J. Nurmi, C. Brunelli, and H. Berg, "State of the Art Baseband DSP Platforms for Software Defined Radio : A Survey," *EURASIP Journal on Wireless Communications and Networking*, June 2011.

[71] http ://www.ettus.com.

[72] G. J. Minden, J. B. Evans, L. Searl, D. DePardo, V. R. Petty, R. Rajbanshi, T. Newman, Q. Chen, F. Weidling, J. Guffey, D. Datla, B. Barker, M. Peck, B. Cordill, A. M. Wyglinski, and A. Agah, "KUAR : A Flexible Software-Defined Radio Development Platform," in *IEEE New Frontiers in Dynamic Spectrum Access Networks*, April 2007, pp. 428–439.

[73] B. Bougard, B. De Sutter, D. Verkest, L. V. der Perre, and R. Lauwereins, "A Coarse-grained Array Accelerator for Software-Defined Radio," *IEEE Micro*, vol. 28, no. 4, pp. 41–50, 2008.

[74] K. Van Berkel, F. Heinle, P. P. E. Meuwissen, K. Moerman, and M. Weiss, "Vector processing as an enabler for software-defined radio in handheld devices," *EURASIP Journal on Advances in Signal Processing*, January 2005.

[75] T. Limberg, M. Winter, M. Bimberg, R. Klemm, E. Matus, M. Tavares, G. Fettweis, H. Ahlendorf, and P. Robelly, "A Fully Programmable 40 GOPS SDR Single Chip Baseband for LTE/WiMAX Terminals," in *European Solid-State Circuits Conference (ESSCIRC)*, September 2008, pp. 466–469.

[76] "Embb, a generic hardware and software architecture for digital signal processing," http ://embb.telecom-paristech.fr/.

[77] F. Clermidy, R. Lemaire, X. Popon, D. Ktenas, and Y. Thonnart, "An Open and Reconfigurable Platform for 4G Telecommunication : Concepts and Application," in *Euromicro Conference on Digital System Design, Architectures, Methods and Tools (DSD)*, August 2009, pp. 449–456.

[78] http ://warp.rice.edu.

[79] http ://www.nutaq.com.

[80] G. Jianxin, Y. Xiaohui, G. Jun, and L. Quan, "The Software Communication Architecture : Evolutions and Trends," in *IEEE Conference on Computational Intelligence and Industrial Applications (PACIIA)*, November 2009.

[81] R. Reinhart *et al.*, "Space telecommunications radio system (strs) architecture standard," *NASA glenn research center, Clevelend, TM 2010-216809*, 2010.

[82] E. D. Willink, "The Waveform Description Language : Moving from Implementation to Specification," in *IEEE Military Communications Conference (MILCOM 2001)*, October 2001.

[83] Y. Lin, R. Mullenix, M. Woh, S. Mahlke, T. Mudge, A. Reid, and K. Flautner, "SPEX : A Programming Language for Software Defined Radio," in *In Software Defined Radio Technical Conference and Product Exposition (SDR-Forum)*, November 2006.

[84] J. Gonzalez-Pina, R. Ameur-Boulifa, and R. Pacalet, "DiplodocusDF, a Domain-Specific Modelling Language for Software Defined Radio Applications," in *Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, September 2012, pp. 1–8.

[85] A. Gelonch, X. Revès, V. Marojevik, and R. Ferrús, "P-HAL : a Middleware for SDR applications," in *SDR Forum Technical Conference*, November 2005.

[86] E. Grayver, H. S. Gree, and J. L. Roberson, "SDRPHY - XML Description for SDR Physical Layer," in *Military Communications Conference (MILCOM)*, October 2010, pp. 1140–1146.

[87] "GNU Radio, The free and open software radio ecosystem," www.gnuradio.org.

[88] J. Delahaye, J. Palicot, C. Moy, and P. Leray, "Partial Reconfiguration of FPGAs for Dynamical Reconfiguration of a Software Radio Platform," in *IST Mobile and Wireless Communications Summit*, July 2007, pp. 1–5.

[89] A. Massouri and T. Risset, "FPGA-based Implementation of Multiple PHY Layers of IEEE 802.15.4 Targeting SDR Platform," in *Wireless Innovation Forum (SDR-WInnComm)*, April 2014, pp. 90–98.

[90] C.-J. Tseng and D. P. Siewiorek, "Automated Synthesis of Data Paths in Digital Systems," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 5, no. 3, pp. 379–395, 1986.

[91] C. H. Gebotys and M. I. Elmasry, "VLSI Design Synthesis with Testability," in *ACM/IEEE Design Automation Conference (DAC)*, June 1988, pp. 16–21.

[92] K. Wakabayashi, "C-based Behavioral Synthesis and Verification Analysis on Industrial Design Examples," in *Asia and South Pacific Design Automation Conference (ASP-DAC)*, January 2004, pp. 344–348.

[93] M. Fowler and R. Parsons, *Domain-Specific Languages.* The Addison-Wisley Signature Series, 2011.

[94] J. Schafer and D. Klein, "Implementing Situation Awareness for Car-to-X Applications using Domain Specific Languages," in *IEEE Vehicular Technology Conference (VTC-Spring)*, June 2013.

[95] G. S. Ouedraogo, "Automatic Synthesis of Hardware Accelerators from High-Level Specifications of Physical Layers for Flexible Radio," Thèse de doctorat, Université de Rennes 1, Décembre 2014.

[96] E. Casseau and B. Le Gal, "Design of Multi-Mode Application-Specific Cores Based on High-Level Synthesis," *Integration, the VLSI Journal*, vol. 45, no. 1, pp. 9–21, 2012.

[97] E. Lemoine and D. Merceron, "Run Time Reconfiguration of FPGA for Scanning Genomic Databases," in *IEEE Symposium on Field-Programmable Custom Computing Machines (FCCM)*, April 1995, pp. 90–98.

[98] "Partial Reconfiguration User Guide, UG702 (v14.1), www.xilinx.com," 2012.

[99] "Increasing Design Functionality with Partial and Dynamic Reconfiguration in 28-nm FPGAs, Altera white paper, WP-01137-1.0, www.altera.com."

[100] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, "Julia : A fresh approach to numerical computing," *SIAM Review*, vol. 59, pp. 65–98, 2017.

[101] G. E. Moore, "Cramming More Components onto Integrated Circuits," *IEEE Solid-State Circuits Society Newsletter*, April 1965.

[102] R. H. Dennard, F. H. Gaensslen, V. L. Rideout, E. Bassous, and A. R. LeBlanc, "Design of Ion-implanted MOSFET's with Very Small Physical Dimensions," *IEEE Journal of Solid-State Circuits*, October 1974.

[103] H. Esmaeilzadeh, E. Blem, R. St. Amant, K. Sankaralingam, and D. Burger, "Dark Silicon and the End of Multicore Scaling," in *International Symposium on Computer Architecture (ISCA)*, June 2011, pp. 365–376.

[104] D. Cordes, P. Marwedel, and M. Arindam, "Automatic parallelization of embedded software using hierarchical task graphs and integer linear programming," in *IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis (CODES+ISSS)*, October 2010.

[105] J. Ceng, J. Castrillon, W. Sheng, H. Scharwächter, R. Leupers, G. Ascheid, H. Meyr, T. Isshiki, and H. Kunieda, "MAPS : An Integrated Framework for MPSoC Application Parallelization," in *ACM/IEEE Design Automation Conference (DAC)*, June 2008.

[106] M. Amini, B. Creusillet, S. Even, R. Keryell, O. Goubier, S. Guelton, J. O. Mcmahon, F.-X. Pasquier, G. Péan, and P. Villalon, "Par4All : From Convex Array Regions to Heterogeneous Computing," in *International Workshop on Polyhedral Compilation Techniques (HiPEAC)*, January 2012.

[107] A. Floch, T. Yuki, A. El-Moussawi, A. Morvan, K. Martin, M. Naullet, M. Alle, L. L'Hours, N. Simon, S. Derrien, F. Charot, C. Wolinski, and O. Sentieys, "GeCoS : A framework for prototyping custom hardware design flows," in *IEEE International Working Conference on Source Code Analysis and Manipulation (SCAM)*, September 2013.

[108] Y. S. Shao, B. Reagen, G. Y. Wei, and D. Brooks, "The aladdin approach to accelerator design and modeling," *IEEE Micro*, pp. 58–70, May 2015.

[109] M. E. Wolf and M. S. Lam, "A Loop Transformation Theory and an Algorithm to Maximize Parallelism," *IEEE Transactions on Parallel and Distributed Systems*, vol. 2, no. 4, pp. 452–471, October 1991.

[110] C. Bienia, S. Kumar, J. P. Singh, and K. Li, "The PARSEC Benchmark Suite : Characterization and Architectural Implications," in *International Conference on Parallel Architectures and Compilation Techniques (PACT)*, October 2008.

[111] S. Verdoolaege, *isl : An Integer Set Library for the Polyhedral Model.* Springer Berlin Heidelberg, 2010, pp. 299–302.

[112] L. H. Crockett, R. A. Elliot, M. A. Enderwitz, and R. W. Stewart, *The Zynq Book : Embedded Processing with the Arm Cortex-A9 on the Xilinx Zynq-7000 All Programmable Soc.* Strathclyde Academic Media, 2014.

[113] Xilinx, "Vivado Design Suite - HLX edition," 2015. [Online]. Available : http://www.xilinx.com/products/design-tools/vivado.html

[114] G. Optimization, "Gurobi optimizer reference manual," 2016. [Online]. Available : http://www.gurobi.com

[115] K. Asanovic, R. Bodik, B. C. Catanzaro, J. J. Gebis, P. Husbands, K. Keutzer, D. A. Patterson, W. L. Plishker, J. Shalf, S. W. Williams *et al.*, "The landscape of parallel computing research : A view from berkeley," Technical Report UCB/EECS-2006-183, EECS Department, University of California, Berkeley, Tech. Rep., December 2006.

[116] B. Reagen, R. Adolf, Y. S. Shao, G.-Y. Wei, and D. Brooks, "Machsuite : Benchmarks for Accelerator Design and Customized Architectures," in *IEEE International Symposium on Workload Characterization (IISWC)*, October 2014.

[117] E. Research, "USRP X310 Software Defined Radio (SDR) - Ettus Research," 2018. [Online]. Available : https://www.ettus.com/product/details/X310-KIT

[118] V. Liu, A. Parks, V. Talla, S. Gollakota, D. Wetherall, and J. R. Smith, "Ambient Backscatter : Wireless Communication out of Thin Air," in *ACM Conference on SIGCOMM*, August 2013, pp. 39–50.

[119] M. Tahir and R. Farrell, "Optimal Communication-Computation Tradeoff for Wireless Multimedia Sensor Network Lifetime Maximization," in *IEEE Wireless Communications and Networking Conference (WCNC)*, April 2009, pp. 1–6.

[120] Y. Hayashi, N. Homma, Y. Toriumi, K. Takaya, and T. Aoki, "Remote Visualization of Screen Images Using a Pseudo-Antenna That Blends Into the Mobile Environment," *IEEE Transactions on Electromagnetic Compatibility*, vol. 59, no. 1, pp. 24–33, February 2017.

[121] H. Tanaka, "Information Leakage via Electromagnetic Emanations and Evaluation of Tempest Countermeasures," in *3rd International Conference on Information Systems Security (ICISS)*, Berlin, Heidelberg, 2007, pp. 167–179.

[122] A. Checko and et al., "Cloud RAN for Mobile Networks ; A Technology Overview," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 405 – 426, 2015.

[123] S. Barbarossa, S. Sardellitti, and P. D. Lorenzo, "Communicating While Computing : Distributed mobile cloud computing over 5G heterogeneous networks," *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 45–55, November 2014.