Contribution à la qualité de service dans les réseaux d'accès sans-fil

Mohamad El Masri

Université de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l' *Institut National des Sciences Appliquées*

Discipline : Systèmes Informatiques

Présentée et soutenue par

Mohamad El Masri

Le 9 juillet 2009

Contribution à la qualité de service dans les réseaux d'accès sans fil

---

JURY

Président :	Michel Diaz	LAAS-CNRS
Rapporteurs :	Andrzej Duda	Grenoble INP-Ensimag
	Francis Lepage	Université Henri Poincaré, Nancy
Examinateurs :	Slim Abdellatif	INSA Toulouse
	Fethi Filali	Eurecom
	Guy Juanole	Université Paul Sabatier

École Doctorale : École Doctorale Systèmes
Unité de Recherche : Laboratoire d'Analyse et d'Architectures des Systèmes - LAAS - CNRS
Directeurs de Thèse : Slim Abdellatif, Guy Juanole

À chaque fois que la question était posée, je disais que je travaillais pour l'Humanité.
C'est donc à Elle que je dédis ce manuscrit :p

فقلْ لمنْ يدَّعي في العلمِ فلسفةً
حفِظْتَ شَيئًا، وغابَتْ عنك أشياءُ
أبو نواس

Dit à celui qui, de la science se prétend philosophe,
Tu as appris une chose, et tu en ignores beaucoup
Abu Nawas

Avant propos..

Beaucoup des personnes citées ici mériteraient chacune une page de remerciements, beaucoup d'autres qui n'y sont pas mériteraient d'y apparaître.

Je tiens d'abord à exprimer ma profonde gratitude à Monsieur le Professeur émérite Guy Juanole. Monsieur Juanole votre disponibilité, des fois tard le soir, tôt le matin ou en week-end, a été exceptionnelle. Vos remarques, vos idées, vos propositions sont le rocher sur lequel ce travail est bâti.

Je souhaite remercier de tout mon cœur Monsieur Slim Abdellatif. Tant sur le plan humain que scientifique, tu as été d'un encouragement sans cesse. Ce travail n'aurait pu être ce qu'il est si ce n'est pour toi. Je suis ravi de t'avoir connu, en tant qu'enseignant puis en tant qu'encadrant, collègue et même adversaire au badminton.

Je tiens à remercier Messieurs Francis Lepage et Andrzej Duda d'avoir consacré une partie de leur temps pour juger ce travail et pour leurs remarques, Monsieur Fethi Filali d'avoir accepter de faire partie du jury de soutenance et Monsieur Michel Diaz d'avoir présider le jury de soutenance.

Des personnes ont contribué aux résultats exposés dans ce manuscrit. Je tiens à remercier particulièrement Mokhtaria Meftah, Hilal Daezly, Arnaud Limouzin et Boris Compaore.

Je tiens à remercier Gina Briand, Sophie Achte, Yves Crouzet et Daniel Daurat pour la logistique de dernière minute, ainsi que le service documentation du LAAS, le groupe OLC et ses membres, le LAAS et ses membres

Ces années passées au LAAS ont été pour moi pleines de rencontres. Des personnes qui ont fait que ce temps passé au laboratoire a été aussi riche humainement qu'il l'était scientifiquement, je tiens à les remercier pour tous les moments de bonne humeur. Les colocataires de bureaux : Ahmad, Florent, Ihsan, Karim, Rodrigo et Usman. Les testeurs du réseau : Batak, Florin, Fred, i2, Ion, Pascal, Nabil et Thierry entre autres. Ceux avec qui j'ai partagé des moments autour d'un café ou d'un gâteau au frigo ou ailleurs : François, Géraldine, German O'Neal, Joe, Jorge O'Neal, Mathieu, Matthias, Nicolas, Racha, Rodrigo, Roxy, Rym et tous ceux qui précédent et tous ceux qui suivent. D'autres avec qui j'ai eu beaucoup de discussions socio-politico-acharnées sur la *situation* de ce côté de la Méditerranée ou de bien d'autres côtés : Baptiste, Florent, François, Ismael, Riyad.

J'ai rencontré en début de thèse un groupe de personnes avec qui j'ai partagé de nombreuses soirées (de travail des fois) et périples. Nous avons organisé deux journées d'étude sur le thème obscur de la *serendipity*, journées qui étaient (IMHO) bien plus intéressantes que des

workshops sur la QoS ou sur les réseaux auxquelles j'ai assisté. Je vous salue, les premiers serendipitesques, bien bas : Alin, Anne, Aurélie, Benjamin, Magali, Sophie et Wafaa.

Il y a des gens qui ont réussi à me supporter (au sens *to support* comme au sens *to bear*), certains pendant des années, à supporter mes sauts d'humeur, mes absences répétées et prolongées des ondes et mes coups de tête, à m'accompagner dans cette entreprise difficile qu'est la thèse. Mes amis, je vous remercie : Adil (le voici ton sudoku), Badr, Bassem, Fatim, Georges et sa jolie famille, Ghida, Jérôme, Kawt, Layale (arigato), Moe Ali, Ossama (7abibi w ibn 7abibi), Sandy (ç), Sam.

Ghassan et Wissam, les meilleurs frères qu'on puisse avoir, ils ont été là quand il le fallait, à m'épauler voire plus. Mes parents à qui je dois ce que je suis. Les mots ne sauront exprimer ce que j'ai envie de leur dire.

May the wind always be at your back, and the sun always upon your face, and may the wings of destiny carry you aloft to dance with the stars.

Table des matières

Liste des figures

Introduction...

Moss : *This, Jen, is the Internet ...*
Jen : *Hang on, it doesn't have any wires or anything!*
Moss : *It's wireless!*
Jen : *Oh yes. Everything is wireless nowadays, isn't it?*

"The IT Crowd" saison 3, épisode 4

Ces dernières années ont connu une démocratisation des réseaux sans fil. Les réseaux sans fil font maintenant partie du quotidien de l'homme de par l'utilisation accrue des téléphones portables, la multiplication des *hotspots* WiFi dans les villes et l'installation de réseaux domestiques gravitant autour d'une *box* fournissant aux utilisateurs aussi bien un accès filaire que sans fil. Les réseaux industriels, utilisant traditionnellement des technologies filaires se sont tournés vers le sans-fil à cause de la facilité d'installation et de la flexibilité que ces technologies apportent. Il aurait été intéressant d'étudier ce phénomène sous beaucoup d'angles : d'une part, les stratégies Marketing employées par les grands groupes de télécommunications pour atteindre un tel emballement ; d'autre part, l'effet qu'a cette montée de l'utilisation des technologies sans fil sur le comportement de l'être humain ou même l'effet physiologique que cela peut avoir. Ceci ne sera pas le propos de ce manuscrit, nous nous situerons strictement dans un contexte technique. Il sera pour nous question de contribuer aux mécanismes permettant de satisfaire les besoins en qualité de service (QoS pour *Quality of Service* ou QdS) des applications (et donc des utilisateurs) lorsque ceux-ci utilisent diverses technologies d'accès sans fil (et, par voie de conséquence, aider les groupes de télécommunications à mieux vendre, satisfaire les exigences des utilisateurs et influer éventuellement sur leurs comportements).

Cette dernière décennie a connu une croissance fulgurante de l'utilisation des technologies sans fil, à commencer par le GSM (comme le montrent les chiffres fournis par l'autorité de régulation des communications électroniques et des postes - l'ARCEP (ARCEP 09)) et puis ensuite avec WiFi. Une étude réalisée pour le compte de l'ARCEP en 2006 dit qu'en 3 ans (entre 2003 et 2006) environ 37000 *hotspots* ont été installés en France. Les puces WiFi ne sont plus limitées aux seuls ordinateurs portables tel que c'était le cas ces dernières années. Aujourd'hui la plupart des téléphones mobiles permettent un accès WiFi. De même des baladeurs multimédia ainsi que d'autres équipements grand public commencent à en être équipés. Certaines prévisions (chiffres *ABIresearch* (ABIResearch 09)) parlent d'un milliard de puces WiFi vendues dans le monde pour la seule année 2012.

En parallèle avec cette modification profonde des technologies habituellement utilisées pour l'accès, l'univers des réseaux de données a connu une autre révolution au niveau applicatif. Internet qui servait, en ses débuts, à transporter du texte et des images devient maintenant porteur de communications téléphoniques, de vidéo, de données exigeant, entre autres, des limites sur les délais ou sur le taux de pertes introduisant ainsi la nécessité d'offrir des services de communication avec une qualité de service. La qualité de service (QoS) peut être définie comme étant la capacité d'un réseau à transporter les données d'une application de façon à ce que cette application satisfasse les exigences de l'utilisateur. Techniquement, la

QoS se définit par des critères que le réseau doit respecter sur les délais, la gigue ou la bande passante entre autres.

Au niveau de la couche Réseau du modèle OSI (typiquement IP), des solutions ont été conçues à partir du milieu des années 90 pour répondre aux exigences de QoS des applications émergentes. IntServ (*Integrated Services*) et DiffServ (*Differentiated Services*) sont les modèles les plus marquants. IntServ, couplé avec RSVP (*Resource ReSerVation Protocol*), est un système permettant d'offrir de la qualité de service avec une granularité fine. Chaque application peut exprimer, par le biais de sa spécification (TSPEC - *Traffic SPECification*), un besoin de **réservation de ressources** nécessaires à son bon fonctionnement sur tout le chemin traversé. Le modèle de DiffServ contraste avec IntServ par sa granularité moins fine. DiffServ présente un cadre permettant de construire des services avec QoS. Il définit la notion de classes de trafic dans lesquelles sont classées les applications suivant leurs exigences en termes de QoS. Chaque classe est associée au niveau d'un domaine DiffServ à un comportement spécifique (PHB - *Per Hop Behavior*) permettant d'établir **une priorité** entre les classes et d'assurer aux applications ainsi classifiées les garanties qu'elles exigent. Ces modèles sont souvent couplés à des protocoles de routage prenant en compte la QoS et des techniques d'étiquetage et de commutation telles que MPLS (*MultiProtocol Label Switching*) fonctionnant au niveau des routeurs.

Offrir de la QoS au niveau de la couche réseau (notamment IP) n'est pas suffisant. En effet, les réseaux physiques sur lesquels reposent les réseaux IP doivent à leur tour offrir une QoS en conformité avec la QoS exigée par l'application, ne serait ce que pour garantir une certaine QoS dans le transport de données entre deux nœuds de l'Internet (entre un hôte et un routeur ou entre deux routeurs). Ce problème se pose particulièrement sur les réseaux d'extrémité. Pour les réseaux d'extrémité de type Ethernet commuté qui constitue la référence pour les réseaux filaires, IEEE 802.1D a proposé des solutions afin de permettre la maîtrise de la QoS offerte sur ce type de réseau. Il s'agit de solutions assez similaires dans leurs concepts à celles présentées au niveau réseau, avec la définition de classes de trafic identifiées par un UP (*User Priority*) et la possibilité d'y associer des traitements adéquats. Ces solutions ne sont pas directement applicables aux réseaux sans-fil tels que WiFi ou WiMAX pour plusieurs raisons : les ressources et les performances de ce type de réseaux sont moindres, elles sont de plus variables et quelque peu imprévisibles. Elles sont également partagées entre des stations du réseau compliquant le contrôle sur les ressources que peut avoir une solution de QoS, celle-ci devant composer avec la technique d'accès multiple utilisée. De nouvelles solutions, assez différentes de celles des réseaux filaires, sont à concevoir. C'est dans ce cadre que s'inscrivent les travaux de cette thèse.

Notre objectif est d'apporter notre brique dans les travaux qui, ces dernières années, ont eu pour but de proposer et d'évaluer les mécanismes à employer pour offrir de la qualité de service au niveau des réseaux d'accès sans fil. Nous nous intéressons spécifiquement à deux technologies réseaux complémentaires : WiFi et WiMAX ; WiFi pour la raison que nous avons exposée ci-dessus : le "futur" très positif qu'on prévoit à cette technologie ; WiMAX car cette technologie représente une alternative prometteuse et évidente aux technologies d'accès DSL (malgré le retard que prend la mise en place de cette technologie en France pour diverses raisons essentiellement non-techniques). Un des scénarios d'utilisation qu'on peut imaginer est par exemple celui d'une région rurale où la mise en place d'accès xDSL peut s'avérer coûteuse (contrastant les coûts de travaux Génie Civil avec le marché ciblé). Dans ce genre de régions, WiMAX peut s'avérer une solution peu coûteuse permettant d'apporter un accès large bande aux utilisateurs finaux. On peut même imaginer un cas où WiMAX fournirait

à des *hotspots* WiFi, et donc aux utilisateurs, un accès à Internet. Cette solution permet de profiter de l'ubiquité des puces WiFi. Un tel scénario a d'ailleurs été considéré par le *WiMAX Forum*, un consortium d'industriels cherchant à promouvoir l'utilisation de WiMAX. Il est essentiel dans un tel contexte que des mécanismes de qualité de service soient mis en place au niveau des deux réseaux d'accès afin de satisfaire les exigences des utilisateurs. Nos travaux présentent des contributions à plusieurs niveaux de ces deux réseaux. Nous présentons dans ce manuscrit l'ensemble de ces contributions.

Ce manuscrit est organisé de la façon suivante :

- Le *premier chapitre* présente un état de l'art des technologies et des mécanismes que nous étudions.
- Le *deuxième chapitre* expose un modèle d'EDCA (*Enhanced Distributed Channel Access*) de la norme IEEE 802.11e que nous avons développé ainsi qu'un cas d'utilisation de ce modèle. Nous avons développé (sur la base de modèles existants) un modèle analytique en chaîne de Markov du mécanisme d'accès distribué avec différentiation de WiFi. L'originalité de ce modèle est la prise en compte explicite du phénomène de collision virtuelle. Nous présentons le modèle sous une forme détaillée que nous appelons le modèle général. Nous avons également réalisé une version réduite du modèle général qui lui est équivalente. Ce modèle peut être employé dans divers cas : une des utilisations que nous en faisons est l'évaluation de la performance d'une modification apportée au mécanisme d'accès EDCA. Une annexe en fin du manuscrit vient présenter des calculs rentrant dans le développement du modèle.
- Le *troisième chapitre* présente un algorithme de contrôle d'admission pour EDCA que nous proposons. Cet algorithme utilise le modèle réduit d'EDCA dans son processus de décision. L'algorithme que nous avons développé a été évalué par voie de simulation. Nous l'avons implémenté sur une plateforme réseau au sein du LAAS en vue d'une évaluation *"in vivo"*.
- Le *quatrième chapitre* de ce manuscrit s'intéresse à notre contribution autour de WiMAX. Nous avons, à ce niveau, développé une modalité de gestion de la bande passante pour WiMAX. Cette modalité est basée sur une gestion dite agrégée de la bande passante. Nous avons évalué notre contribution par voie de simulation.

Les résultats de cette thèse ont donné lieu à huit publications : (Masri 06), (Masri 07a), (Masri 07b), (Masri 08b), (Masri 08c), (Masri 08a), (Masri 09b) et (Masri 09a).

État de l'art

1

"The wireless telegraph is not difficult to understand. The ordinary telegraph is like a very long cat. You pull the tail in New York, and it meows in Los Angeles. The wireless is exactly the same, only without the cat."
Attribué à Albert Einstein

Sommaire

Nous parcourons dans ce chapitre l'ensemble de l'état de l'art. Nous couvrirons les prérequis techniques nécessaires à la compréhension de nos contributions, il s'agit des aspects techniques des mécanismes d'accès sans fil que nous étudions : l'accès pour les WLAN utilisant le standard IEEE 802.11 (appelé aussi WiFi par abus de langage) et l'accès pour les WMAN utilisant le standard IEEE 802.16 (appelé aussi WiMAX par abus de langage). Pour chacun des deux standards, nous nous intéresserons aux différents mécanismes qu'ils proposent pour fournir une qualité de service de niveau liaison. Nous présenterons aussi les travaux de recherche qui gravitent autour de ce sujet.

1.1 IEEE 802.11

1.1.1 Introduction aux standards IEEE 802.11

Le standard IEEE 802.11-2007 (80211 07) définit les spécifications d'un contrôle d'accès au médium et de plusieurs couches physiques pour la connectivité sans fil de stations fixes ou mobiles dans une "zone locale" (*local area* dans le standard). Il est le résultat de l'intégration au premier standard IEEE 802.11-1997 (80211 97) (communément appelé 802.11-*legacy* ; publié en 1999 et confirmé en 2003) des différentes modifications qui sont apparues au fil des années jusqu'à la date de publication de la nouvelle révision. 802.11-*legacy*, rendu obsolète par les modifications apportées, permettait un fonctionnement à des débits physiques de 1 Mbit/s ou 2 Mbit/s utilisant 3 technologies au niveau physique au choix : l'infrarouge ; les radio-fréquences avec étalement de spectre par saut de fréquence (FHSS *Frequency-Hopping Spread Spectrum*) ou avec étalement de spectre à séquence directe (DSSS *Direct-Sequence Spread Spectrum*). Les technologies à radio-fréquences fonctionnent dans la bande ISM *Industrial Scientific Medical*) autour de 2,4 GHz et la bande U-NII (*Unlicensed National Information Infrastructure*) autour de 5 GHz. Les modifications qu'intègre cette nouvelle révision sont les suivantes :

IEEE 802.11a-1999 permettant un fonctionnement dans la bande de fréquence 5 Ghz à un débit physique allant jusqu'à 54 Mbit/s grâce à l'utilisation d'une technique de codage OFDM (*Orthogonal Frequency-Division Multiplexing*) ;

IEEE 802.11b-1999 ainsi que la correction **IEEE 802.11b-1999 Corrigendum 1-2001** permettant d'augmenter le débit physique dans la bande de 2,4 GHz jusqu'à 11 Mbit/s ;

IEEE 802.11d-2001 ajoutant des mécanismes permettant aux stations de respecter les réglementations locales en termes de puissances de transmission et de plages de fréquences autorisées ;

IEEE 802.11g-2003 introduit la technique de codage OFDM dans la bande de fréquence 2,4 GHz permettant ainsi des débit allant jusqu'à 54 Mbit/s ;

IEEE 802.11h-2003 harmonisant la modification 802.11a avec les contraintes réglementaires de la communauté européenne ;

IEEE 802.11i-2004 spécifiant des mécanismes de sécurité ;

IEEE 802.11j-2004 établissant la conformité de l'utilisation de la bande 4,9-5 GHz avec les règles japonaises ;

IEEE 802.11e-2005 spécifiant des mécanismes de Qualité de Service sur lesquelles porte une partie de notre travail ; les modifications apportées par le groupe de travail 802.11e sont détaillées dans le document (802.11e 05).

D'autres groupes de travail et groupes d'études ont été mis en place par le comité de travail 802.11 afin de mettre en place des modifications à apporter au standard 802.11, parmi ceux-là :

IEEE 802.11k-2008 fournit aux couches de plus haut niveau des interfaces d'accès à des mesures radio, cette modification a été publiée en 2008 (802.11k 08) ;

IEEE 802.11n étudie la possibilité de fournir des débits utiles allant au delà des 100 Mbit/s. Le groupe de travail 802.11n est toujours actif avec un projet de modification proposant l'utilisation de la technologie MIMO (*Multiple-Input Multiple-Output*). La standardisation de 802.11n est prévue pour Novembre 2009 ;

IEEE 802.11p étudie la possibilité de modifier le standard afin de fournir la possibilité de communiquer entre un véhicule et une entité de bord de route ou entre deux véhicules

circulant à une vitesse allant jusqu'à 200 Km/h avec une portée allant jusqu'à 1000 m, le travail de ce groupe se pose dans le cadre d'un système de transport intelligent, le groupe de travail est actif ;

IEEE 802.11r-2008 permet une connectivité continue des entités mobiles en utilisant des mécanismes de transition rapides, transparents et sécurisés : *Fast Basic Service Set Transition*, cette modification a été publiée en 2008 (802.11r 08) ;

IEEE 802.11aa ou VTS (*Video Transport System*) a pour but de spécifier des mécanismes permettant le transport robuste des flux audio et vidéo sur 802.11, le travail de ce groupe a débuté en 2007 et est toujours actif ;

1.1.2 La couche physique

Nous exposerons rapidement dans cette section les différents aspects de la couche physique tels que détaillés dans le standard (80211 07). Ce standard présente 5 spécifications de couches physiques qui diffèrent par les techniques de codage et de modulation utilisées, par la bande de fréquence préconisée et par conséquent par le débit binaire maximal possible. Ces différentes spécifications sont :

- **FHSS** (*Frequency-Hopping Spread Spectrum*) ce mode utilise une modulation par déplacement de phase Gaussienne GFSK (*Gaussian Frequecy Shift Keying*) avec un étalement de spectre par saut de fréquences dans la bande 2.4 GHz. Ce mode permet un débit physique de 1 Mbit/s (si la modulation utilisée est une 2GFSK) ou 2 Mbit/s (si la modulation utilisée est une 4GFSK).
- **DSSS** (*Direct-Sequence Spread Spectrum*) ce mode utilise une modulation par déplacement de phase Différentiel DPSK (*Differential Phase Shift Keying*) avec un étalement de spectre à séquence direct dans la bande 2.4 GHz. Ce mode permet un débit physique de 1 Mbit/s (si le déplacement de phase est binaire DBPSK) ou 2 Mbit/s (si le déplacement de phase est quadratique DQPSK).
- **IR** (*InfraRed*) cette spécification de la couche physique n'est plus développée. Elle utilise un mode de diffusion infrarouge avec des longueurs d'onde qui varient entre 850 et 950 nm. Une modulation d'impulsions par positions est utilisée donnant des débits physiques de 1 Mbit/s ou 2 Mbit/s suivant le type d'encodage.
- **OFDM** (*Orthogonal Frequency -Division Multiplexing*) cette spécification de la couche physique fonctionne par répartition en fréquences orthogonales dans la bande 5 GHz. Elle définit plusieurs modes de fonctionnement caractérisés par le type de modulation utilisé : BPSK, QPSK, 16 QAM (*16 Quadratic Amplitude Modulation*, une modulation d'amplitude en quadrature) ou 64 QAM (*64 Quadrature Amplitude Modulation*) ; le ratio de correction d'erreur FEC (*Forward Error Correction*) utilisé :1/2, 2/3 ou 3/4 ; et la largeur des canaux de fréquences : 5, 10 ou 20 MHz donnant ainsi une spécification de la couche physique pouvant offrir un débit physique allant de 1.5 Mb/s à 54 Mb/s.
- **HR/DSSS** (*High Rate DSSS*) ce mode se propose d'étendre le mode DSSS afin d'offrir des débits physiques plus importants. Une modulation code complémentaire (CCK *Complementary Code Keying*) sera utilisée permettant d'obtenir des débits de 5.5 Mb/s et de 11 Mb/s.
- **ERP** (*Extended Rate Physical Layer*) combine un ensemble de mécanismes introduits par les spécifications précédentes (notamment l'introduction de l'OFDM) pour améliorer les débits physiques (allant de 5.5 Mb/s à 54 Mb/s) dans la bande de fréquence 2.4 GHz.

Un certain nombre de valeurs utilisées dans les procédures d'accès dépendent de la couche physique utilisée. Nous noterons spécifiquement les valeurs *aSIFSTime* et *aSlotTime*. *aSIFSTime* est défini comme étant le temps nominal entre la réception du dernier symbole d'une trame par la couche PHY et le début d'envoi du premier symbole de la trame réponse, en comptant le temps nécessaire au traitement au niveau MAC. *aSlotTime* est la durée d'un *slot*, le *slot* correspond à la durée au bout de laquelle un accès au médium effectué par une station du réseau est détecté par toutes les autres stations.

1.1.3 La couche MAC

1.1.3.1 Premières définitions

Le standard divise les stations (STA) en deux types :

non-QoS STA sont les stations ne supportant pas les mécanismes de qualité de service préconisés par la modification du standard apportée par IEEE 802.11e (802.11e 05) ;

QoS STA ou QSTA sont les stations supportant les mécanismes de qualité de service de la modification 802.11e.

Le standard définit une BSS (*Basic Service Set*) comme étant le bloc constituant d'un réseau local 802.11. Une BSS est typiquement composée d'un point d'accès (AP, *Access Point*) et d'une ou plusieurs STA. Une IBSS (*Independant Basic Service Set*) est un bloc constituant d'un réseau ad hoc, il est donc constitué d'au moins deux stations sans infrastructure fixe.

1.1.3.2 Architecture MAC

L'architecture générale du plan de données de la couche MAC de IEEE 802.11 (80211 07) est présentée dans la figure 1.1. Elle est constituée des fonctions suivantes :

DCF (*Distirbuted Coordination Function*) est la fonction permettant l'accès avec contention. DCF est basée sur un schéma CSMA/CA (*Carrier Sense Multiple Access with Collision Avoidance*). L'implémentation de DCF est obligatoire pour les non-QoS STA et pour les QoS STA.

PCF (*Point Coordination Function*) est la fonction gérant l'accès sans contention pour les non-QoS STA. PCF est basée sur un système de scrutation par un point central. Sa présence est optionnelle.

HCF (*Hybrid Coordination Function*) est une fonction introduite par la modification 802.11e apportant deux nouveaux mécanismes d'accès : **EDCA** (*Enhanced Distributed Channel Access*) qui fournit un service d'accès avec contention (basé sur CSMA/CA) avec différenciation de trafic ; **HCCA** (*HCF Controlled Channel Access*) qui fournit un accès sans contention (par scrutation) pour un service avec QoS paramétrée. La fonction HCF est obligatoire dans les QoS STA.

Les fonctions PCF et HCF se basent sur la fonction DCF pour leur accès, en d'autres mots, PCF et HCF utilisent les services de DCF pour accéder au médium. La différenciation entre les fonctions d'accès se fait essetiellement au moyen de l'emploi de différents intervalles inter-trames (IFS-*Inter-Frame Spaces*) que nous définissons dans le paragraphe 1.1.4.1. Nous présentons dans la suite le fonctionnement de chaque fonction d'accès.

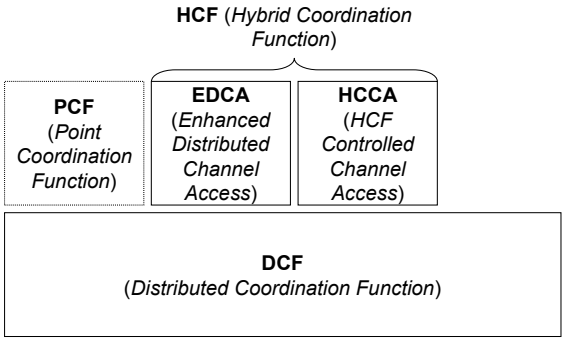

FIG. 1.1 – *L'architecture de la sous-couche MAC de 802.11*

1.1.4 DCF

DCF est le protocole d'accès de base introduit par IEEE 802.11. Il est basé sur un schéma CSMA/CA (*Carrier Sense Multiple Access with Collision Avoidance*), le schéma d'accès est couplé à un méchanisme optionel RTS/CTS (*Request To Send/Clear To Send*, mécanisme aidant à la résolution des problèmes du terminal caché et du terminal exposé et à la réduction de l'effet des collisions). Les stations du réseau souhaitant accéder au réseau vont d'abord s'assurer que le médium est libre pendant un certain temps IFS (*InterFrame Space*), cette partie de la procédure est appelée CS (*Carrier Sense*) ; si le médium est libre, la station accède au médium, sinon la station diffère sa transmission jusqu'à ce que le médium devienne libre. Dès que le médium devient libre, une station souhaitant transmettre doit procéder au *backoff* en attendant un temps aléatoire tiré dans l'intervalle $[0, CW]$ (CW étant la valeur de la fenêtre de contention) ; ceci permet de réduire la probabilité de collision. À la réception correcte d'une trame qui lui est destinée, une station doit envoyer un acquittement (ACK), si aucun acquittement n'est reçu par la station émettrice à l'expiration d'une temporisation interne, la station émettrice déduit qu'une collision a eu lieu. Une collision a principalement lieu lorsque la procedure de backoff aléatoire de deux fonctions DCF de deux stations du réseau expirent en même temps. Ainsi, afin de réduire le taux de collisions, la valeur de la fenêtre de contention d'une station subissant une collision est doublée dans la limite de l'intervalle de contention $[CW_{min}, CW_{max}]$ (CW_{min} et CW_{max} étant des paramètres protocolaires décrits ci-aprés), la trame peut être retransmise jusqu'à atteinte de la limite de retransmissions.

1.1.4.1 IFS - *InterFrame Space*

Le standard définit cinq types d'espace entre trames (IFS) permettant d'instaurer des priorités entre les accès, ces temporisations sont utilisées par DCF et par les autres fonctions d'accès :

SIFS (*Short InterFrame Space*) sera utilisé pour séparer deux trames faisant partie d'un même échange atomique ; par exemple entre la transmission d'une trame et la transmission de l'acquittement correspondant, entre la transmission d'une trame RTS et la trame CTS correspondante, entre deux trames faisant partie d'une même rafale. SIFS est le plus petit des IFS, il permet ainsi d'empêcher qu'un échange en cours ne soit interrompu par un autre échange. Sa durée est définie par la valeur de *aSIFSTime*.

PIFS (*PCF InterFrame Space*) est la temporisation utilisée par l'élément central du réseau (le point d'accès AP) pour intervenir au profit d'une scrutation. Cette temporisation est

utilisée par la fonction PCF et par HCCA. Elle est d'une durée supérieure à SIFS (afin d'éviter qu'une scrutation n'interrompe un échange en cours) mais inférieure à DIFS (afin de rendre une scrutation plus prioritaire qu'un accès avec contention).

DIFS (*DCF InterFrame Space*) est la temporisation minimale de médium libre qu'une fonction d'accès avec contention (notamment DCF) doit attendre avant de pouvoir entamer une procédure d'accès.

AIFS (*Arbitration InterFrame Space*) est utilisé par EDCA, il représente l'équivalent de DIFS dans DCF, à savoir une temporisation de séparation des trames représentant une partie du mécanisme de différenciation introduit par EDCA.

EIFS (*Extended InterFrame Space*) est une temporisation, d'une durée supérieure à DIFS, utilisée suite à une réception erronée par une station afin d'éventuellement permettre à une autre station du réseau d'acquitter la réception du paquet qui, pour la première station, contenait des erreurs.

1.1.4.2 BEB - *Binary Exponential Backoff*

Une procédure de backoff est mise en place suite à la détection par la fonction d'accès DCF d'un état de médium libre pour une durée supérieure à DIFS. Cette procédure permet aux stations de réduire la probabilité de collisions. Le backoff correspond à l'attente pendant une durée aléatoire avant de procéder à l'envoi. Ce temps aléatoire ($TempsDeBackoff$) est choisi de la façon suivante :

$$TempsDeBackoff = Uniforme(0, CW) \times aSlotTime$$

$Uniforme(a, b)$ étant la fonction de tirage aléatoire uniforme d'un entier dans l'intervalle $[a, b]$; CW étant la valeur en cours de la fenêtre de contention. La valeur de CW évolue dans l'intervalle $[CW_{min}, CW_{max}]$ défini par le standard, la valeur de CW est initialisée à CW_{min} lorsqu'un paquet vient d'être envoyé avec succès ou lorsqu'un paquet est rejeté suite au dépassement de la limite de retransmissions. Suite à une collision, la valeur de CW est augmentée de façon exponentielle jusqu'à atteindre la borne maximale CW_{max} afin de réduire le taux de collisions :

$$CW = min(CW \times 2 - 1, CW_{max})$$

Une fois la valeur du temps de Backoff tirée, elle est décrémentée de 1 à chaque *slot* libre observé par la fonction d'accès. Lorsque la temps de backoff atteint 0, et si le médium est toujours libre, la fonction d'accès tente l'envoi sur le médium. Si en cours de décrémentation du temps de backoff le médium devient occupé, la valeur en cours du temps de backoff est mémorisée, la décrémentation reprendra au point où elle s'était arrêtée lorsque la fonction observera à nouveau un intervalle DIFS d'inoccupation du médium.

1.1.5 PCF

PCF permet aux stations d'accéder au médium sans contention. Une entité centralisée appelée le PC (*Point Coordinator*) est située au niveau du point d'accès (AP). Lorsqu'un PC est actif au niveau du point d'accès, le temps d'accès au médium est divisé en deux périodes : la CFP (*Contention Free Period*), période où l'accès se fait exclusivement en PCF ; et la CP (*Contention Period*), période où l'accès se fait exclusivement en DCF. Le temps est donc divisé, lors de l'emploi de PCF, en supertrames périodiques contenant chacune une CFP suivie d'une CP. Le début d'une supertrame (et donc de la CFP) est marqué par l'envoi par le PC d'un paquet

Beacon. Le paquet Beacon doit être envoyé à une allure régulière, une temporisation PIFS est employée pour l'envoi du *Beacon*. La division est schématisée par la figure 1.2.

L'accès par PCF pendant les CFP se fait sur la base d'une scrutation. Le PC envoie aux

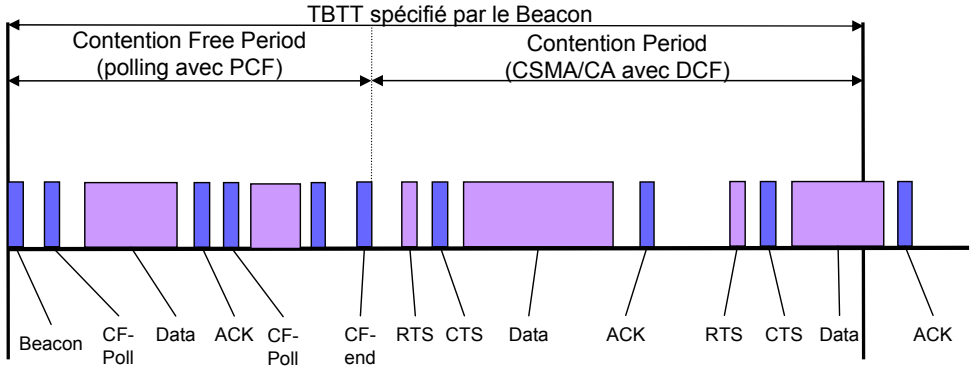

FIG. 1.2 – *Exemple d'une supertrame 802.11*

stations supportant PCF (une telle station est appelée dans le standard une *CF-Pollable STA*) une trame CF-Poll l'invitant à accéder au médium, une temporisation SIFS est employée pour l'envoi des CF-Poll. Une *CF-Pollable STA* recevant une trame CF-Poll peut envoyer une trame données MPDU (*MAC Protocol Data Unit*) avec une temporisation SIFS. Si l'envoi n'est pas acquitté, la station ne peut procéder à une retransmission sauf si elle est à nouveau scrutée ou pendant la période CP. PCF offre des possibilités de *piggybacking* : une trame envoyée par le PC peut être un CF-Poll et contenir un acquittement, des données ou les deux à la fois ; une trame envoyée par une station peut contenir des données et l'acquittement d'un paquet reçu auparavant. PCF a rarement été implémenté. Il est maintenant tombé en désuétude.

1.1.6 HCF : une fonction d'accés au médium avec QoS

HCF (*Hybrid Coordination Function*) est utilisée uniquement dans ce que le standard appelle un réseau QoS (c'est le réseau où le point d'accès met en place un HC-*Hybrid Coordinator*). La modification IEEE 802.11e (802.11e 05) a introduit cette nouvelle méthode ainsi que d'autres mécanismes afin d'apporter certaines propriétés QoS au niveau de l'accès. HCF introduit des modifications à DCF et PCF ainsi qu'un certain nombre de mécanismes et de types de trames permettant la mise en place de transferts avec qualité de service pendant la CP et la CFP. HCF introduit la notion d'opportunité de transmission (TXOP) qu'une QoS-STA peut obtenir en utilisant l'une des méthodes d'accès d'HCF : la méthode d'accès avec contention EDCA (*Enhanced Distributed Channel Access*) ou la méthode d'accès par scrutation (*HCF Controlled Channel Access*). L'obtention d'un TXOP peut permettre l'envoi d'une ou plusieurs trames. Si TXOP vaut 0, une seule trame données peut être envoyée par opportunité de transmission.

1.1.6.1 EDCA

EDCA (représentée par la figure 1.3) apporte à l'accès par contention de 802.11 la différenciation. La méthode définit quatre catégories d'accès (AC) qui ont des propriétés différentes à l'accès. Une proposition de correspondance est établie entre la valeur UP (*User Priority* définie par 802.1D) et les 4 ACs, elle est détaillée dans le tableau 1.1.

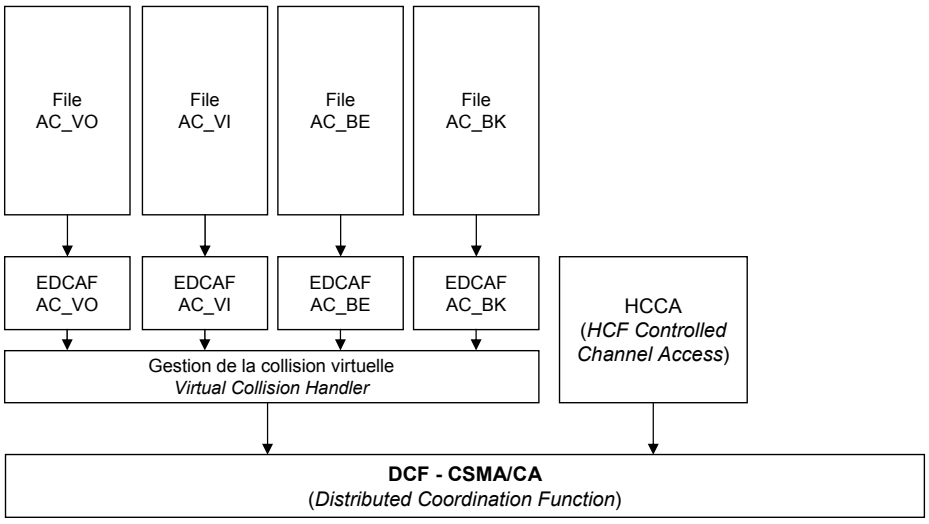

FIG. 1.3 – *Schéma d'une station avec EDCA.*

User Priority	Appellation 802.1D	Catégorie d'Accès
1	BK (*Background*)	AC_BK (*Background*)
2	– (*Spare*)	AC_BK (*Background*)
0	BE (*Best Effort*)	AC_BE (*Best Effort*)
3	EE (*Excellent Effort*)	AC_BE (*Best Effort*)
4	CL (*Controlled Load*)	AC_VI (*Video*)
5	VI (*Video*)	AC_VI (*Video*)
6	VO (*Voice*)	AC_VO (*Voice*)
7	NC (*Network Control*)	AC_VO (*Voice*)

TAB. 1.1 – *Correspondance entre la valeur UP et la catégorie d'accès*

A chaque catégorie d'accès est attachée une fonction logique semblable à DCF, appelée EDCAF (*Enhanced Distributed Channel Access Function*) qui participera à la contention afin d'obtenir une TXOP au profit de l'AC concernée et lui permettre ainsi d'accéder au médium. L'EDCAF fonctionne comme DCF, suivant un schéma CSMA/CA avec BEB, elle a ceci de particulier qu'elle utilise un lot de paramètres de contention qui lui sont propres (une valeur de AIFS et un intervalle de fenêtre de contention $[CW_{min}, CW_{max}]$ ainsi qu'une limite sur la taille d'une TXOP). Ces paramètres sont soit fixés par défaut soit déterminés et annoncés par le point d'accès. Les paramètres de contention permettent de mettre en place la différenciation à l'accès d'EDCA. Les valeurs par défaut (tableau 1.2) le montrent clairement. L'introduction de fonctions d'accès concurrentes et indépendantes au niveau des AC de chaque station amène une nouvelle notion, celle de la collision virtuelle. En effet, les fonctions d'accès procédant à leur activité de backoff de façon indépendantes, il est possible que deux (ou plus) ACs d'une même station voient leurs compteurs de backoff expirer au même moment. Dans ce cas, la catégorie d'accès de plus haute priorité (AC_VO > AC_VI > AC_BK > AC_BE) gagne l'op-

Catégorie d'accès	AIFS	CW_{min}	CW_{max}
AC_VO	$SIFS + 2 \times aSlotTime$	$\frac{aCW_{min}+1}{4} - 1 ; 3$	$\frac{aCW_{min}+1}{2} - 1 ; 7$
AC_VI	$SIFS + 2 \times aSlotTime$	$\frac{aCW_{min}+1}{2} - 1 ; 7$	$aCW_{min} ; 15$
AC_BE	$SIFS + 3 \times aSlotTime$	$aCW_{min} ; 15$	$aCW_{max} ; 1023$
AC_BK	$SIFS + 7 \times aSlotTime$	$aCW_{min} ; 15$	$aCW_{max} ; 1023$

TAB. 1.2 – *Calcul des Valeurs par défaut des paramètres de contention EDCA et exemple avec* $aCW_{min} = 15$ *et* $aCW_{max} = 1023$

portunité de transmission, les éventuelles autres se comportent comme si une collision réelle s'était produite.

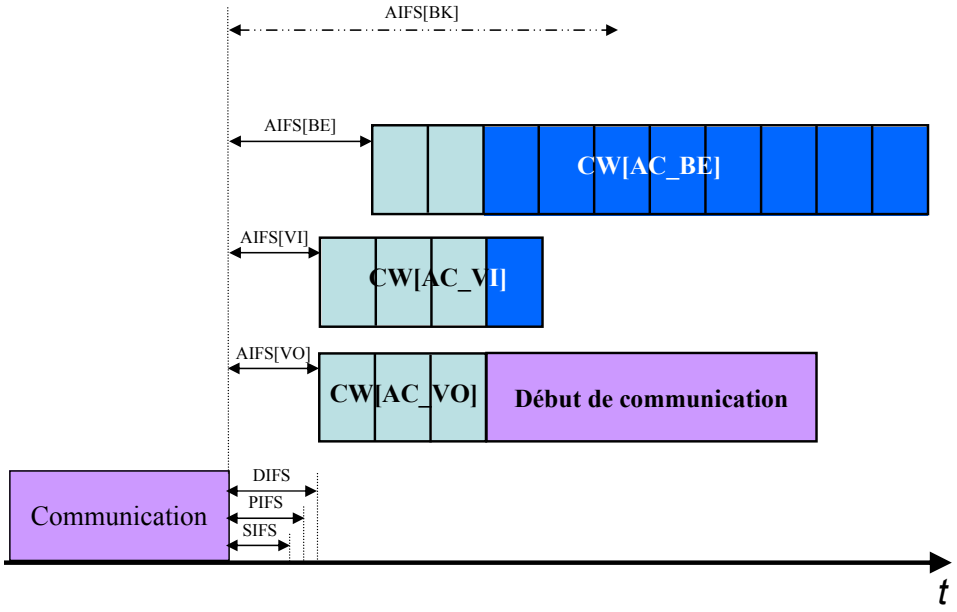

FIG. 1.4 – *Chronologie de l'accès par EDCA : procédure de backoff des différentes ACs.*

Exemple La figure 1.4 présente un exemple de fonctionnement des fonctions d'accès EDCA. Dans l'exemple, nous nous situons au niveau d'une station. Toutes les catégories d'accès de la station souhaitent accéder au médium. Les ACs, voyant une occupation en cours du médium, retardent leurs transmissions. Lorsque l'occupation du médium prend fin, chaque AC utilise une temporisation qui lui est propre avant de se lancer dans la procédure de backoff, chaque AC attend que le médium devienne libre pour sa propre période AIFS à la fin de quoi, l'EDCAF choisira de façon aléatoire un temps de backoff dans la fenêtre de contention. L'EDCAF dont la procédure de backoff expire en premier se voit accorder un TXOP. Les autres garderont en mémoire la valeur de temps de backoff à laquelle elles sont parvenues afin de réactiver la procédure de backoff à ce niveau lorsque le médium sera à nouveau libre.

1.1.6.2 HCCA

HCCA est une implémentation d'un accès par scrutation. Il met en place des améliorations par rapport à PCF. La division de la supertrame en CP et CFP cesse d'être importante lorsque la fonction HCCA est employée. Cette division continue à exister afin de permettre aux non-QoS STA de fonctionner dans un réseau QoS. Cependant, HCCA permet au HC d'intervenir durant une CP ou une CFP pour mettre en place une scrutation. L'accès du HC pour ce genre de procédure se faisant avec une temporisation PIFS qui est inférieure aux AIFS, cet accès devient ainsi plus prioritaire que les accès par EDCA. La scrutation est mise en place grâce à une connaissance supposée du HC de l'état des files des stations au sein de la BSS.
Durant la CP, une QSTA peut accéder au médium si elle y parvient par EDCA, ou si elle y est explicitement invitée par un paquet CF-POLL envoyé par le HC. Le HC peut accéder au médium afin de scruter les QSTA (envoie d'un paquet CF-POLL) suite à une période PIFS d'inactivité du médium. Durant la CFP, les QSTA ne peuvent accéder au médium à moins qu'elles y soient explicitement invitées par un paquet CF-POLL envoyé par le HC. Le HC peut s'accorder un TXOP pour un envoi sur la voie descendante ou peut envoyer un paquet de scrutation CF-POLL suite à une période PIFS d'inactivité du médium.
Un paquet CF-POLL contient une valeur TXOPLimit, indiquant à la station scrutée la durée d'utilisation du médium à ne pas dépasser.

1.1.7 Les mécanismes supplémentaires de QoS

Le standard (80211 07) présente différents mécanismes, complémmentaires à HCF, permettant d'offrir une QoS pour l'accès 802.11. L'essentiel de ces mécanismes fut introduit par la modification IEEE 802.11e (802.11e 05). Nous en exposons certains dans ce paragraphe, celui qui nous intéresse étant essentiellement le contrôle d'admission.

1.1.7.1 Le contrôle d'admission

Un cadre pour le contrôle d'admission a été mis en place par la modification 802.11e. Ce cadre concerne l'accès par HCF avec ou sans contention (par EDCA ou par HCCA). Le contrôle d'admission servira à la gestion et la régulation de la bande passante disponible. Une QSTA souhaitant avoir des garanties de QoS (sur les délais d'accès, sur les débits ou sur le taux de pertes par exemple) devra passer par le contrôle d'admission. Les algorithmes de contrôle d'admission ne sont pas définis par le standard, le choix de l'algorithme utilisé est laissé à l'équipementier. Le standard définit cependant un cadre et un certain nombre de règles que les algorithmes devront respecter.

Pour EDCA Le point d'accès peut annoncer les ACs qui sont soumises au contrôle d'admission (en réglant le champs ACM - *Admission Control Mandatory* dans les annonces des paramètres EDCA des ACs). Il est à noter que ce champs doit rester statique, le choix des ACs soumises à un contrôle d'admission doit rester le même tout au long de la vie de la BSS. Si une AC est soumise au contrôle d'admission, alors toutes les ACs qui sont de plus haute priorité le sont aussi. Lorsqu' une QSTA souhaite utiliser une AC soumise au contrôle d'admission pour un nouveau flux, la QSTA envoie une requête ADDTS (*ADD Traffic Stream*) précisant l'AC utilisée, le sens du flux (montant, descendant ou bidirectionel) et la spécification du nouveau flux (TSPEC). En réponse à une requête ADDTS, l'AP renvoie une réponse ADDTS précisant si le nouveau flux est accepté ou refusé. Les règles de prise de décision d'ac-

cepter ou de refuser un nouveau flux (autrement dit, l'algorithme de contrôle d'admission) n'est pas spécifié par le standard. Un exemple d'algorithme est donné en annexe du standard , cet exemple n'est pas obligatoire. Si un flux se voit refuser l'accès, il doit soit utiliser une AC de moindre priorité, non soumise au contrôle d'admission soit renégocier l'admission en spécifiant un nouveau TSPEC. Lorsque le flux arrive à sa fin d'activité, la QSTA envoie à l'AP un message DELTS (*DELete Traffic Strem*) le lui signalant. À chaque EDCAF sont associées deux variables qui devront être mises à jour par l'EDCAF. Ces variables sont *admitted_time* (représentant le temps médium que le contrôle d'admission a accordé à cette AC) et *used_time* (représentant le temps médium consommé par l'EDCAF). L'EDCAF est contrainte de respecter la limite spécifiée par la variable *admitted_time*. Si les conditions du réseau se dégradent emmenant la station à réduire son débit physique, les EDCAF des AC soumises au contrôle d'admission se doivent de continuer à respecter la contrainte *admitted_time*, une nouvelle demande d'admission peut être envoyée afin de palier au manque de bande passante induit.

Pour HCCA Le contrôle d'admission au niveau de HCCA est essentiel pour que HCCA puisse fournir un service avec une qualité de service paramétrée. Le contrôle d'admission est dans le cas de HCCA couplé à une politique d'ordonnancement qui assurera la qualité de service demandée. La procédure d'ajout ou de suppression de flux est identique à celle d'EDCA. L'algorithme de contrôle d'admission n'est pas spécifié dans le standard mais des règles d'usage sont fixées. Un HC ayant accepté un nouveau flux ne peut en aucun cas le modifier ni l'annuler sans qu'il y ait eu expressément une demande de la part de la QSTA concernée. Le but de HCCA est d'offrir des garanties à des flux (en considérant que le réseau évolue dans un environnement contrôlé), il est donc important que le contrôle d'admission et l'algorithme d'ordonnancement y veillent.

1.1.7.2 Autres mécanismes de QoS introduits par IEEE 802.11e

Block ACK Cette procédure optionnelle permet d'améliorer l'utilisation du médium. En effet, elle permet à une station d'envoyer plusieurs paquets sans que ceux-là soient acquittés individuellement. Le bloc de paquets pourra être acquitté à la fin de l'envoi du bloc ou dans un TXOP ultérieur. L'utilisation du réseau s'en trouve ainsi améliorée.

Direct Link Protocol Si DLP (*Direct Link Protocol*) est activé au sein d'une BSS, les QSTA peuvent communiquer directement sans avoir à passer nécessairement par l'AP. DLP permet d'améliorer l'utilisation de la bande passante.

Respect des échéances D'autres modifications ont été introduites par IEEE 802.11e permettant d'améliorer le respect des échéances en contraignant les durées d'accès des stations :

- Les stations utilisant le médium sont contraintes de respecter le TBTT annoncé par le paquet Beacon. Une station voulant accéder au médium doit vérifier que la transmission entamée (jusqu'à la réception éventuelle du ACK) ne doit pas dépasser le TBTT annoncé. Le CFP ne sera, par conséquent, pas retardé par les stations accédant au médium.
- Chaque accès au médium se fait dans la limite de l'opportunité de transmission (TXOP) accordée. La valeur du TXOP est fixée par le HC.

1.1.8 Revue de la littérature

Depuis le début du processus de standardisation de 802.11, beaucoup de travaux de recherche se sont intéressés à différents aspects de ce standard, dans la volonté de l'évaluer et d'en améliorer la performance. Nous nous intéresserons dans cette revue de la littérature aux travaux récents qui ont eu pour objet les améliorations du standard IEEE 802.11 se rapportant à la qualité de service (essentiellement celles apportées par la modification IEEE 802.11e (802.11e 05)). Certains des travaux présentés ci-dessous n'ont pas forcément de rapport direct avec la contribution de cette thèse, ils ont cependant apporté une brique à la réflexion et méritent donc pour cette raison d'être cités. Ainsi, étant donnée la nature large de cette revue de littérature, le placement de nos travaux par rapport à la littérature se fera dans les chapitres présentant nos contributions. Il est possible de diviser les travaux en plusieurs catégories :

- Les travaux sur l'évaluation de la performance des techniques d'accès du standard. L'évaluation de performances a été réalisée, soit avec des modèles analytiques, soit avec la simulation, soit sur des plateformes de test.
- Les travaux sur les aspects *cross-layering*.
- Les travaux sur les mécanismes de QoS, couvrant à la fois des travaux sur l'amélioration du schéma d'accès, des travaux sur l'ordonnancement ou des travaux sur le contrôle d'admission. Certains autres travaux sont allés jusqu'à proposer des architectures de QoS pour IEEE 802.11.

1.1.8.1 Travaux sur l'évaluation de performances

Plusieurs modèles analytiques ont été réalisés avec comme principal but l'analyse de performances de 802.11. À commencer par (Bianchi 00) qui proposa l'un des premiers modèles pour DCF. Vinrent ensuite (Kong 04), (Wang 05a), (Engelstad 05), (Clifford 06), (Sharma 06), (Wu 06) entre autres travaux qui proposèrent des modèles d'EDCA et analysèrent ses performances. Nous présenterons quelques uns de ces modèles dans le paragraphe 2.1.
D'autres travaux d'analyse de performances ont utilisé des outils de simulation. (Mangold 02) analyse l'équité d'EDCA dans un environnement avec chevauchement de BSS. (Mangold 03) est l'un des premiers papiers à proposer une analyse complète par simulation de HCF. Ce papier montre l'amélioration qu'apporte HCF du point de vue de la QoS par rapport aux schémas d'accès classiques. Il compare aussi les performances d'EDCA à celles d'HCCA dans différentes conditions réseau. (Garg 03) évalue la performance de EDCA et de HCCA et étudie la possibilité d'un changement des paramètres d'accès d'EDCA.
Plus récemment, (Ng 05) et (Dangerfield 06) ont réalisé une étude sur plateforme de test de IEEE 802.11e avec des applications TCP ou Voix sur IP.

1.1.8.2 Travaux sur les aspects *cross-layering*

Au niveau du *cross-layering*, les travaux ont intégré, dans le fonctionnement des schémas d'accès 802.11 des informations venant de la couche physique ou des couches supérieures. (Chen 04) propose un système modifiant le comportement de la couche MAC en se basant sur des mesures de l'état du canal ; (Zhai 05) étudie l'impact des conditions canal sur la Qualité de Service perçue. (Casetti 05) explore l'utilisation d'EDCA et de HCCA pour des applications voix sur IP.

1.1.8.3 Travaux sur les mécanismes de Qualité de Service

Travaux sur l'amélioration du schéma d'accès Les premiers travaux d'amélioration de IEEE 802.11e se sont intéressés au schéma d'accès distribué EDCA et aux différents paramètres qui rentrent en jeu à ce niveau. (Romdhani 03) et (Malli 04) proposent d'apporter des modifications au fonctionnement du BEB : soit en modifiant le comportement des tailles de fenêtres de contention après une transmission avec succès, soit en modifiant la vitesse de décompte du backoff en l'accélérant au fil du décompte. (Xiao 04) propose un système à deux niveaux permettant de protéger les flux voix et vidéo (utilisant EDCA) des autres flux QoS (en adoptant un contrôle d'admission) et des flux non-QoS (en adaptant au vol les paramètres d'accès EDCA). (Ramos 05) donne un aperçu des directions de recherche nécessaires pour assurer une qualité de service sur les réseaux d'accès 802.11. Il propose une solution basée sur l'adaptation des paramètres d'EDCA (seuil de fragmentation, facteur de persistence, la vitesse de décompte du backoff). Le système ainsi conçu a pour but de pallier aux changements des conditions canal et de permettre à la fonction d'accès de s'adapter au profil applicatif. (Lee 05) propose de modifier les paramètres d'accès afin d'assurer à la fois la priorité et une équité proportionnelle entre les catégories d'accès EDCA. (Yang 06) propose lui d'adapter la taille de la fenêtre de contention aux besoins en termes de délai des flux actifs de façon distribuée. (Velayutham 06) propose de réserver certaines supertrames aux flux QoS permettant ainsi une protection accrue de ces flux. (Kuo 08) rajoute un système de requêtes de bande passante à l'ajustement des variables EDCA afin de permettre la transmission de flux multimedia et de flux *best effort* dans des conditions de réseaux changeantes. Enfin, (Loyola 08) propose de mettre en place un système de retransmission coopératif permettant aux stations ayant une meilleure condition canal de retransmettre les paquets envoyés par des stations ayant une mauvaise qualité de canal.

Sur les algorithmes d'ordonnancement Les travaux de recherche se sont intéressés à plusieurs aspects : l'ordonnancement des scrutations HCCA en est un. (Ansel 04) propose un schéma d'ordonnancement améliorant le traitement des flux VBR en intégrant une estimation de l'état des files des stations au niveau du point d'accès et une distribution équitable du temps accordé à une station parmi les flux le nécessitant. (Lim 04) propose lui aussi un ordonnanceur pour IEEE 802.11e utilisé en tandem avec EDCA et HCCA pour améliorer la QoS perçue. (Fallah 04) et (Fallah 05) présentent différentes approches pour l'ordonnancement des scrutations sous HCCA en employant un système d'émulation virtuelle des stations au niveau de l'AP. (Ramos 07) propose un système d'ordonnancement utilisant à la fois EDCA et HCCA de façon dynamique afin d'améliorer la qualité de service perçue par les applications voix et vidéo.

Sur les algorithmes de contrôle d'admission (Gao 05) présente une revue des contrôles d'admission pour IEEE 802.11e EDCA et HCCA. (Chou 05) utilise une approche basée sur le temps d'occupation du médium pour établir son contrôle d'admission, il met aussi en place un système de signalisation pour HCCA pour permettre au contrôle d'admission de fonctionner correctement. (Bai 06) présente un algorithme d'admission basé sur une adaptation de la variable *admitted_time* aux conditions du réseau. (Bai 07) présente un algorithme de contrôle d'admission basé sur un modèle simple de l'utilisation d'un *timeslot* en EDCA. (Lee 07) propose un autre système de calcul pour le contrôle d'admission s'adaptant au débit physique. (Assi 08) propose un algorithme de contrôle d'admission par flux ainsi qu'un système d'adaptation des paramètres d'accès EDCA par flux afin de satisfaire les exigences des flux.

Des propositions d'architectures D'autres travaux se sont intéressés à l'intégration des mécanismes de qualité de service IEEE 802.11 dans une architecture de bout en bout. (Park 03) propose et étudie une architecture QoS de bout en bout avec interaction entre DiffServ et IEEE 802.11e. (Li 04) propose un cadre de QoS de bout en bout dans un réseau hybride filaire-sans fil, il combine les mécanismes de IEEE 802.11 avec RSVP dans la partie filaire (*Resource Reservation Protocol*).

1.2 IEEE 802.16

1.2.1 Introduction aux standards IEEE 802.16

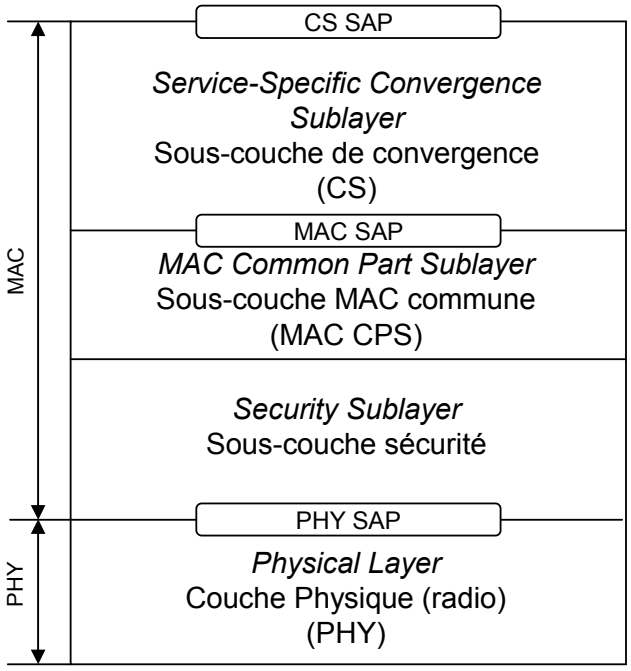

FIG. 1.5 – *Architecture de IEEE 802.16*

IEEE 802.16 (802.16 04) est le standard spécifiant plusieurs interfaces physiques ainsi que la sous-couche MAC permettant de mettre en place un système d'accès sans fil à large bande (BWA - *Broadband Wireless Access*) orienté connexion. Un réseau IEEE 802.16 est typiquement constitué d'une station de base (BS - *Base Station*), élément central de contrôle du réseau et d'au moins une station (SS - *Subscriber Station*). Une telle topologie est dite PMP (*Point to MultiPoint*). Une topologie maillée (MESH) a aussi été introduite ; dans celle-ci, la communication se fait directement entre les SS. Le standard définit un modèle de référence de IEEE 802.16 que nous présentons dans la figure 1.5. Pour la couche PHY, le standard définit plusieurs spécifications, avec pour chacune une bande de fréquence, un débit et les propriétés qui en découlent. La couche MAC comprend trois sous-couches :

– La sous-couche de convergence (CS) dont le rôle est de fournir un service de correspondance et de transformation de l'ensemble des données reçues des couches de plus haut niveau (la couche réseau) à travers le point d'accès au service (SAP - *Service Access Point*) en données que la sous-couche sous-jascente comprend. Ceci concerne, entre autres, les spécifications de trafic, les identifiants des connexions et les paquets de niveau réseau.

- La sous-couche MAC commune (CPS) contient l'ensemble des fonctionnalités d'accès : l'établissement et la maintenance des connexions, l'accès suivant les classes de service WiMAX, la gestion de la bande passante.
- La sous-couche de sécurité assure l'authentification, l'échange de clefs et le chiffrement/déchiffrement des trames.

1.2.2 La couche physique

Le standard IEEE 802.16 définit plusieurs spécifications possibles de la couche physique. Ces spécifications dépendent du type de modulation utilisé ainsi que de la bande de fréquence, le choix de la spécification de la couche physique est couplé au type d'application/déploiement visé.

WirelessMAN-SC est une spécification de couche physique à porteuse unique fonctionnant dans la bande 10-66 GHz (bande soumise à une licence) et nécessitant une ligne de vue (LOS - *Line Of Sight*) entre les éléments communicants du réseau (typiquement entre la BS et chaque SS du réseau).

WirelessMAN-SCa est une spécification de couche physique à porteuse unique fonctionnant dans la bande de fréquence inférieure à 11 GHz (bande soumise à une licence). Il est ainsi possible de se passer de la ligne de vue entre les éléments communicants, on passe en mode NLOS (*Non Line Of Sight*).

WirelessMAN-OFDM utilise une modulation OFDM (*Orthogonal Frequency-Division Multiplexing*) dans la bande de fréquence inférieure à 11 GHz en mode NLOS (bande soumise à une licence).

WirelessMAN-OFDMA utilise une modulation OFDMA (Yin 06) (*Orthogonal Frequency-Division Multiple Access*) dans la bande de fréquence inférieure à 11 GHz en mode NLOS (bande soumise à une licence).

WirelessHUMAN fonctionne dans la bande 5-6 GHz, libre d'utilisation. Cette spécification se base sur l'une des trois spécifications qui précèdent (à savoir SCa, OFDM ou OFDMA) en modifiant la division des canaux et le masque spectral de transmission.

1.2.2.1 Le duplexage

Comme pour 802.11, IEEE 802.16 utilise la notion de supertrame dans la gestion physique du temps : le temps est divisé en supertrames. Au sein de chaque supertrame, le standard supporte deux modes de duplexage des voies montante et descendante : le duplexage temporel (TDD - *Time Division Duplexing*) et le duplexage fréquentiel (FDD - *Frequency Division Duplexing*).
Dans le mode TDD (figure 1.6), la sous-trame descendante (DL - *Dowlink Subframe* servant à la communication descendante de la BS vers les SS) et la sous-trame montante (UL - *Uplink Subframe* servant à la communication montante des SS vers la BS) utilisent la même bande de fréquence en alternance. Dans le mode FDD (figure 1.7), les deux sous-trames utilisent des bandes de fréquence différentes simultanément. Il est possible dans ce mode d'avoir simultanément dans le même réseau des SS fonctionnant en mode *full-duplex* (capable d'envoyer et de recevoir en même temps) et des SS fonctionnant en mode *half-duplex* (ne pouvant pas envoyer et recevoir en même temps). En cas de présence de SS *half-duplex*, l'organisation de l'accès au médium devra en tenir compte afin qu'il n'y ait pas chevauchement entre les périodes d'en-

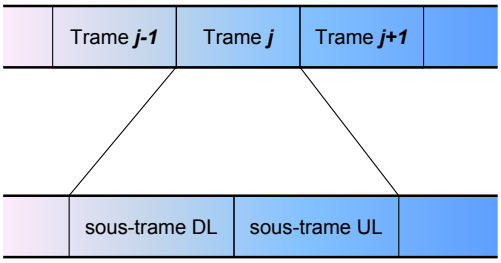

FIG. 1.6 – *Duplexage d'une supertrame en mode TDD : chaque sous-trame est transportée sur un intervalle de temps séparé.*

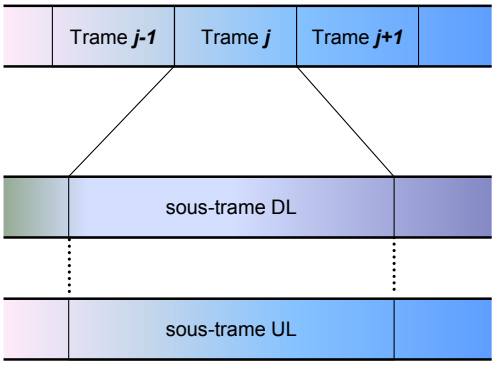

FIG. 1.7 – *Duplexage d'une supertrame en mode FDD : chaque sous-trame est transportée sur une bande de fréquence séparée.*

voi et de réception de ces stations. La construction des sous-trames et leurs contenus seront détaillés dans le paragraphe suivant.

1.2.3 La technique d'accès

L'organisation du contenu des sous-trames DL et UL est de la responsabilité de la BS et est annoncée par deux messages envoyés par la BS en début de chaque sous-trame DL, ces messages sont le DL-MAP et l'UL-MAP. La sous-trame DL est envoyée en TDM (*Time Division Multiplexing*), la sous-trame UL est organisée en *slots* où l'accès se fera par TDMA (*Time-Division Multiple Access*). Dans certains cas la sous-trame DL peut contenir une partie où l'accès (descendant) devient en TDMA, ce sont des périodes de réception DL en fin de supertrame. C'est le cas notamment en mode FDD avec des SS *half-duplex* : si de telles SS ont leurs *slots* UL dédiés en début de supertrames, elles se voient spécifier des périodes de réception DL en fin de supertrame avec préambule de synchronisation afin de leur permettre de se resynchroniser après passage du mode transmission en mode réception. Les figures 1.8 et 1.9 montrent le détail du contenu des sous-trames DL et UL pour le mode TDD et pour le mode FDD. L'exemple que nous donnons ici concerne une couche physique avec porteuse unique, l'extrapolation de cette représentation pour des couches physiques multi-porteuse consiste à multiplier la dimension de la figure pour prendre en compte l'utilisation des différentes porteuses.

1.2.3.1 La sous-trame DL

La sous-trame DL débute par un préambule servant à établir la synchronisation en début de sous-trame. Le préambule est suivi par deux messages établissant l'ordonnancement de

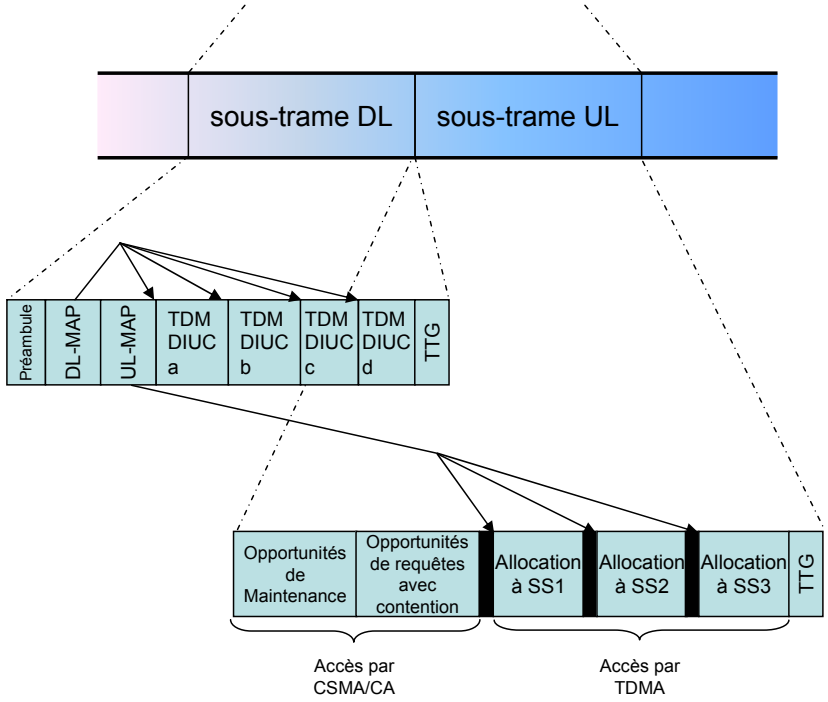

FIG. 1.8 – *Détail des sous-trames DL et UL en mode TDD.*

la sous-trame DL (message DL-MAP) et de la sous-trame UL (message UL-MAP). Ces deux messages indiquent l'organisation de chacune des sous-trames et les instants de début de chaque partie des sous-trames. Ces deux messages sont suivis de la partie d'accès TDM, cette partie est organisée par profil d'envoi, du plus robuste ou moins robuste. Dans la partie TDM, la BS va tour à tour envoyer les trames dans le sens descendant. En mode TDD, la partie TDM est conclue par un intervalle TTG (Transmit/recieve Transition Gap) permettant de séparer la partie d'envoi de la partie réception. En mode FDD, étant donné qu'il faut assurer la gestion des SS en *half-duplex*, la partie TDM (organisée de la même façon qu'en TDD) est utilisée pour l'envoi à destination :

- des SS *full-duplex* ;
- des SS *half-duplex* transmettant dans la même supertrame à un instant ultérieur à celui de la réception ;
- des SS *half-duplex* n'ayant pas de transmissions prévues pendant la même supertrame.

La partie TDM est suivie d'un ou de plusieurs intervalles DL (les intervalles TDMA) où les transmissions sont destinées aux SS *half-duplex* transmettant avant de recevoir dans la supertrame. Cette division permet à une SS de pouvoir décoder une partie de la sous-trame DL sans qu'elle n'ait eu à décoder l'ensemble de la sous-trame DL. Les intervalles TDMA sont séparés par des intervalles "préambule" permettant de rétablir la synchronisation de phase avec les SS concernées.

1.2.3.2 La sous-trame UL

La sous-trame UL est organisée de la même façon en mode TDD ou en mode FDD. Elle débute habituellement par deux périodes où l'accès se fait par CSMA/CA. Il s'agit d'une première période où la BS alloue des opportunités de maintenance initiale (*Initial Ranging Opportunities*)

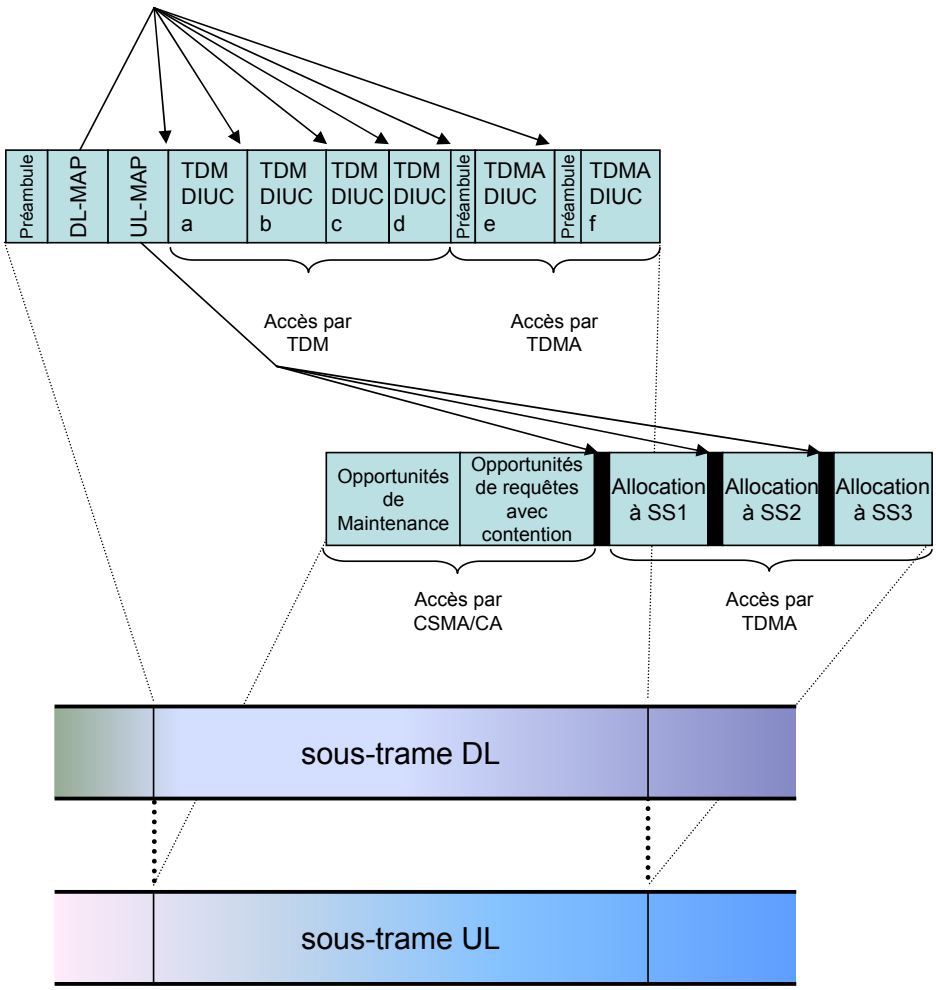

FIG. 1.9 – *Détail des sous-trames DL et UL en mode FDD.*

permettant l'entrée au réseau de nouvelles SS. La deuxième période concerne les opportunités d'envoi de requêtes de bande passante *multicast* ou *broadcast* (nous détaillerons cet aspect dans le paragraphe 1.2.4.2). Ces deux premières périodes peuvent être combinées en une seule. La sous-trame UL est ensuite organisée en périodes allouées spécifiquement aux SS, la taille et l'ordre de ces périodes sont spécifiés par l'algorithme d'ordonnancement *Uplink* au niveau de la BS. Ces périodes sont séparées par des intervalles SSTG (*SS Transition Gap*) permettant à la BS de se synchroniser avec la profil de la SS suivante. Durant les périodes allouées, une SS peut accéder avec son profil d'envoi pour envoyer des données utilisateurs ou des requêtes de bande passante.

1.2.4 QoS pour IEEE 802.16 : la gestion de la bande passante montante

IEEE 802.16 intègre divers mécanismes de qualité de service. Sa gestion de l'accès sur la voie montante se fait en tenant compte des besoins QoS des flux servis. Nous présentons dans cette section le cadre de la gestion de la bande passante montante proposé par le standard IEEE 802.16.

Sous IEEE 802.16, la gestion de la bande passante se fait en mode orienté connexion.

Chaque connexion, identifiée par son CID (*Connection Identifier*), est associée à un type de service. À chaque type de service sont associés une procédure spécifique et un ensemble de mécanismes lui permettant d'exprimer ses besoins en termes de bande passante. Nous rappelons dans le paragraphe qui suit les différents types de service définis par le standard. Nous détaillons ensuite les différents mécanismes d'expression de besoins et la manière préconisée par IEEE 802.16 pour allouer la bande passante.

1.2.4.1 Rappel des différents types de service

Le standard spécifie quatre types de services destinés à servir des connexions au niveau MAC. Chaque type de service correspond à un type d'application ayant un profil de trafic particulier. Ces services sont :

- **Unsolicited Grant Service (UGS)** : Ce type de service correspond à des flux avec contraintes temporelles qui génèrent périodiquement des données de la même taille. Les flux utilisant ce type de service n'ont pas à explicitement demander que leur soit accordée la chance de transmettre, ceci se fait automatiquement et en corrélation avec la spécification du flux donnée à l'établissement de la connexion.
- **Real-Time Polling Service (rtPS)** : Ce type de service correspond à des flux ayant un profil de trafic irrégulier et ayant des besoins "temps réel". La communication des besoins des flux de ce type se fait par *polling* (scrutation). La station de base accorde à ce type de flux des intervalles réguliers où ils peuvent communiquer leurs besoins par voie de requête. La fréquence de scrutation est déterminée à l'aide de la spécification du flux donnée lors de la demande de connexion. Les requêtes sont transmises au système d'ordonnancement qui les utilise pour mettre en place l'organisation de la voie montante.
- **Non-Real-Time Polling Service (nrtPS)** : Ce type de service correspond à des flux peu sensibles aux délais et à la gigue. Ces flux sont scrutés pour communiquer leurs besoins, cette scrutation doit être faite avec une période de l'ordre de la seconde. La différence avec le service rtPS est l'absence, ici, de contrainte sur les délais.
- **Best Effort (BE)** : Ce type de service correspond à des flux sans besoins spécifiques en termes de garanties de qualité de service. Ils seront donc servis au mieux que puisse le faire le gestionnaire de ressource.

La modification au standard IEEE 802.16e (802.16e 05), traitant de la mobilité inclut un cinquième type de service, l'**Extended Real-Time Polling Service** (ertPS). Ce service se place entre le service UGS et le service rtPS. Il permet de servir les applications temps réel qui génèrent périodiquement des paquets de taille variable, l'exemple donné dans le standard est celui d'une application voix sur IP avec suppression de silence.

Nous explicitons par la suite les différents mécanismes permettant aux flux d'exprimer leurs besoins.

1.2.4.2 Expression de besoin en bande passante

L'expression des besoins de bande passante sur la voie montante se fait pour chaque connexion. En premier lieu à l'établissement d'une nouvelle connexion, la SS indique à la BS une spécification du trafic correspondant à la nouvelle connexion ainsi que le type de service demandé, ces informations sont portées par le message de gestion DSA-REQ (*Dynamic Service Addition Request*). La BS prendra en compte ces informations afin d'organiser l'approvisionnement en bande passante. La spécification du trafic peut être considérée comme une

première expression des besoins de la connexion (en termes de débit, de délai maximum, de gigue tolérée entre autres). Par la suite, et suivant le type de service adopté, la connexion peut utiliser divers mécanismes lui permettant l'expression de ses besoins : les opportunités d'envoi unicast de requêtes, les opportunités d'envoi avec contention, le mécanisme de *piggybacking*, ou le *Poll Me Bit* dont nous détaillerons l'utilisation suivant le type de service.

- **UGS :** Une connexion UGS porte un trafic régulier. De par cette caractéristique, une connexion UGS n'a pas à exprimer de besoins en cours d'activité, elle se contente de la spécification initiale. Il lui est accordé, de façon périodique de la bande passante correspondant à la demande faite à l'établissement de la connexion. La BS se base sur l'information *Maximum Sustained Traffic* contenue dans la spécification du trafic afin d'établir le volume de cette allocation périodique. Une connexion UGS n'a aucun autre moyen de demander de la bande passante. Dans l'éventualité d'une perte d'UL-MAP ou d'une désynchronisation d'horloge provoquant le dépassement de la limite fixée pour la file d'attente de la connexion, la connexion UGS peut utiliser le drapeau *Slip Indicator* (SI bit) afin d'informer la BS de la situation. Une connexion UGS peut fixer le drapeau *Poll Me* (PM bit) à 1 si d'autres connexions (qui ne sont pas de type UGS) souhaitent être scrutées. Les drapeaux SI et PM font partie de la sous-entête *Grant Management* qui peut être insérée en amont des PDU d'une connexion UGS.
- **ertPS** Une connexion de type ertPS a, de la même façon que pour UGS, des opportunités pour envoyer des données sans qu'elle ait à le demander. La taille de l'allocation est déduite du champ *Maximum Sustained Traffic* contenue dans la spécification du trafic. Il est cependant possible pour une connexion ertPS de demander la modification de la taille de l'allocation afin de l'adapter au profil du trafic.
- **rtPS** La BS fournit à une connexion de type rtPS, de façon périodique, et selon la spécification du trafic indiquée à l'établissement de la connexion, l'occasion d'envoyer une requête de bande passante de type unicast. Une occasion de type unicast (en anglais : *unicast request opportunity*) est un intervalle de temps dédié à la connexion lui permettant d'exprimer ses besoins en bande passante. Il n'est pas permis à une connexion de type rtPS d'utiliser les opportunités d'envoi de requêtes avec contention (ces opportunités sont des intervalles accessibles par CSMA/CA aux connexions de moindre priorité).
- **nrtPS** La BS fournit à une connexion nrtPS des opportunités unicasts de façon régulière (le standard indique que la période de ces opportunités doit être de l'ordre de la seconde). Une connexion nrtPS peut utiliser les opportunités d'envoi de requêtes avec contention.
- **BE** La BS pourra offrir à une connexion de type BE d'envoyer des requêtes unicasts afin d'exprimer ses besoins en bande passante. La connexion pourra aussi utiliser les opportunités d'envoi de requêtes avec contention.

1.2.4.3 L'allocation de bande passante

À la réception des différentes requêtes, la BS doit effectuer un ordonnancement lui permettant d'établir l'allocation de la bande passante sur la voie montante puis de construire et d'envoyer le message UL-MAP décrivant cette allocation. Les premières versions de l'ébauche du standard IEEE 802.16 (802.16 04) spécifient deux modes d'expression de l'allocation de la bande passante : une allocation par connexion (GPC, *Grant Per Connection*) et une allocation par station (GPSS, *Grant Per Subscriber Station*). Il est essentiel de noter que, d'après le standard, l'allocation de la bande passante au niveau de la BS se fait par connexion (la BS se base sur les

spécifications des connexions ainsi que sur les requêtes envoyées par celles-ci pour ordonnancer les transmissions), l'expression de cette allocation dans l'UL-MAP est ensuite faite selon l'un des deux modes : en GPC, chacune des connexions se voit individuellement accorder dans l'UL-MAP la possibilité d'utiliser le médium soit pour envoyer des données soit pour envoyer des requêtes. En mode GPSS, la BS met en commun le total des allocations faites aux connexions d'une même station. La station reçoit une allocation qu'elle doit ensuite distribuer entre ses connexions.
Il est à noter que le mode GPC est obsolète. Pour des raisons de mise à l'échelle, la dernière version du standard l'a occulté et n'offre que la possibilité d'exprimer les allocations en mode GPSS.

1.2.5 Travaux autour de la QoS pour IEEE 802.16

Tout comme pour IEEE 802.11, les travaux de recherche sur 802.16 ont commencé bien avant la fin du processus de standardisation. Nous présentons ici quelques uns de ces travaux de recherche, essentiellement ceux qui se sont intéressés aux aspects autour de la Qualité de Service dans IEEE 802.16.

1.2.5.1 Travaux sur le schéma d'accès

Certains travaux se sont intéressés au schéma d'accès de 802.16 en général et plus particulièrement au schéma d'accès avec contention qu'emploie 802.16 pour certaines requêtes de bande passante. (Wang 05b) propose un modèle analytique permettant l'analyse de la performance de l'accès avec 802.16. (Kobliakov 06) propose un système d'accès avec contention pour les réseaux centralisés et l'applique à l'accès avec contention pour l'envoi des requêtes avec contention de 802.16. (Vinel 06) et (Ni 07) explorent différentes manières d'organiser l'accès avec contention pour l'envoi des requêtes sous 802.16. (Liu 08) se concentre sur le système que le standard 802.16e (802.16e 05) met en place pour les requêtes des flux utilisant ertPS, il propose une modification et en fait une analyse.

1.2.5.2 Propositions d'architectures de QoS

Quelques travaux ont mis en place des architectures pour 802.16 couvrant divers mécanismes de QoS. (Alavi 05) et (Delicado 06) proposent une architecture de qualité de service générique adaptée au standard 802.16. (Cicconetti 06) analyse la performance des mécanismes de Qualité de Service que met en place le standard IEEE 802.16 dans divers scénarios d'utilisation. Il étudie notamment les scénarios de service aux zones résidentielles ou à une PME.

1.2.5.3 Travaux sur l'ordonnancement

D'autres travaux se sont intéressés aux mécanismes de qualité de service de 802.16 en proposant des algorithmes d'ordonnancement. (Xergias 05) présente un algorithme d'ordonnancement pour WiMAX. Dans (Marchese 07b) et (Marchese 07a), les auteurs proposent un système d'allocation de bande passante couplé à un contrôle d'admission basé sur un modèle mathématique du service fourni. (Ghazal 08) propose l'utilisation d'un ordonnaceur basé sur *Deficit Round Robin* pour parvenir à satisfaire les contraintes QoS des flux et une équité entre les flux.

1.2.5.4 Intégration dans un réseau hybride

À un autre niveau, (Chen 05) propose une méthode pour décider dynamiquement de la durée des sous-trames montantes et descendantes en mode TDD réduisant ainsi le gaspillage de bande passante. (Mukul 06) propose un système d'ordonnancement et un système de requête adaptatif permettant d'assurer un meilleur service au flux rtPS. (Delicado 08) propose et analyse une modification du système de requêtes afin de réduire la période montante où l'accès se fait avec contention.

1.2.5.5 Travaux sur le *Cross Layering*

Les études sur les aspects *Cross Layering* se sont intéressés aux couches supérieures et aux couches inférieures à la couche MAC. (Alexander 06) propose une manière d'intégrer les mécanismes QoS de 802.16 dans une architecture avec une couche session basée sur SIP afin de permettre un meilleur transport de la voix sur 802.16. (Yahiya 08) propose un système intégrant des informations du niveau physique pour assurer la satisfaction de contraintes de délai des flux contraints.

1.2.5.6 Intégration dans un réseau hybride

D'autres travaux se sont intéressés aux réseaux hybrides sans fil, l'intégration des mécanismes permettant de fournir une qualité de service aux applications utilisant ce genre de réseau. (Nie 05) donne un algorithme de prise de décision pour le *handoff* entre réseaux coexistants sans fil employant des technologies différentes. (Soundararajan 07) propose une solution pour réduire les délais que subit les applications voix sur IP en traversant un réseau hybride 802.11 - 802.16. (Takizawa 07) se concentre plutôt sur la ressource radio qui devient difficile à gérer dans un environnement où divers réseaux sans fil coexistent. (Ali-Yahiya 08) propose une architecture générale pour l'interconnexion de réseaux sans fil basés sur le standard IEEE 802.21 et qui permet une mobilité transparente pour les utilisateurs.

Un modèle du schéma d'accès de IEEE 802.11e EDCA

2

"It can scarcely be denied that the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience."
Albert Einstein

Sommaire

L'objectif de ce chapitre est de présenter un modèle formel du comportement d'une catégorie d'accès EDCA. Ce modèle se situe, d'une part, dans le cadre d'un comportement sous saturation (la catégorie d'accès modélisée a en permanence des données à transmettre) et, d'autre part, dans la considération explicite du phénomène de collision virtuelle qu'introduit EDCA (ceci n'est fait, à notre connaissance, par aucun autre modèle). Ce chapitre est structuré en quatre parties :

- La première partie consiste en une rapide présentation des principaux modèles qui nous ont influencés ;
- la deuxième partie concerne la présentation du modèle que nous proposons,
- la troisième partie développe une méthode de réduction de ce modèle,
- la quatrième partie se situe dans le cadre d'une réflexion sur la problématique de la gestion de la collision virtuelle telle que spécifiée par le standard ; nous proposons une modification de cette gestion et nous introduisons cette modification au modèle ; nous effectuons ensuite une analyse des conséquences de cette modification.

2.1 Sur des modèles analytiques existants

Dans le but premier de modéliser le comportement d'EDCA et d'en évaluer la performance, plusieurs travaux se sont intéressés à la proposition de modèles de comportement d'EDCA. Les principaux modèles sont des modèles en chaîne de Markov représentant le comportement d'une catégorie d'accès EDCA vis à vis de son environnement. Ces modèles sont essentiellement basés sur le modèle de DCF proposé par Giuseppe Bianchi (Bianchi 00). Les plus marquants sont celui de Zhu et Chlamtac (Zhu 05) et celui de Kong et al. (Kong 04). Aucun de ces modèles d'EDCA ne prend en compte, de façon explicite, le phénomène de la collision virtuelle qu'introduit EDCA. Une collision virtuelle a lieu lorsqu'au moins deux fonctions d'accès au sein d'une même QSTA tentent d'accéder au médium au même moment ; dans ce cas, la fonction d'accès de plus haute priorité accède au médium, là où les autres se comportent comme si une collision réelle avait eu lieu. Nous exposons maintenant les deux principaux modèles sur lesquels s'est basé notre travail avant de détailler notre proposition.

2.1.1 Le modèle de DCF présenté par Bianchi

Bianchi (Bianchi 00) représente à l'aide d'une chaîne de Markov, le comportement, sous saturation, de l'accès d'une station utilisant DCF comme fonction d'accès au médium. Le modèle est à deux dimensions :

- La première dimension représente les étapes de transmissions : la première transmission et les éventuelles retransmissions dues à une collision.
- La deuxième dimension représente le comportement du backoff : le choix de la valeur du compteur de backoff et le décompte de ce compteur.

À partir de ce modèle, il calcule, en assumant des conditions de canal idéal (c'est à dire que la seule raison pour qu'un paquet transmis ne soit pas correctement reçu est une collision réelle), la probabilité stationnaire d'accès d'une station à un intervalle de temps (*timeslot*) donné. Ensuite, l'évaluation de la performance du schéma d'accès est effectuée en considérant les différents événements pouvant avoir lieu pendant cet intervalle de temps (un accès avec succès ou une collision).

2.1.2 De DCF vers EDCA

La principale action pour réaliser le passage est d'introduire des variables représentant la différence entre les catégories d'accès : différence d'espace inter-trame (IFS-*inter frame space*), différence d'intervalle de fenêtre de contention entre autres. Certains de ces modèles ont gardé l'aspect bidimensionel du modèle originel (Zhu 05), d'autres (Kong 04) ont introduit une troisième dimension permettant de mieux représenter l'état de la file vis à vis de l'état du médium. Nous nous intéresserons au modèle de Kong et. al. (Kong 04) plutôt qu'au modèle de Zhu et Chlamtak (Zhu 05) à cause de cette troisième dimension qui nous permettra de mieux représenter la collision virtuelle.

2.1.2.1 Modèle à 3 dimensions : le modèle de Kong et al.

La troisième dimension introduite par Kong et al. représente les différents états du médium et le comportement de la file vis à vis de ces états : occupé par la file en question, par une autre file ou si le médium est libre. Cette complexité supplémentaire permet au modèle d'assurer une meilleure représentation du comportement d'une catégorie d'accès en introduisant des aspects qui étaient ignorés par les autres modèles.

Cependant le modèle de Kong et al. ne représente pas explicitement le phénomène de la collision virtuelle. Malgré la prise en compte de ce type de collision dans les expressions analytiques qui servent à l'évaluation de performance, la modélisation *per se* de la collision virtuelle est incorrecte : pour le modèle, le comportement de la file face à tout type de collision est représenté de la même manière. D'autres aspects sont ignorés par le modèle de Kong ou sont mal représentés. Nous expliciterons dans le paragraphe 2.2.5 les erreurs ainsi que les modifications et corrections apportées au modèle.

2.2 Notre proposition : un modèle du comportement d'une AC EDCA intégrant la collision virtuelle

Ce modèle corrige par rapport à celui de Kong et al. des erreurs de conceptions (détaillées dans le paragraphe 2.2.5) et introduit de façon explicite la notion de collision virtuelle. Une des spécificités du modèle est sa présentation : en blocs, facilitant ainsi la compréhension de son fonctionnement ainsi que son utilisation.

2.2.1 Vue comportementale d'une AC_i

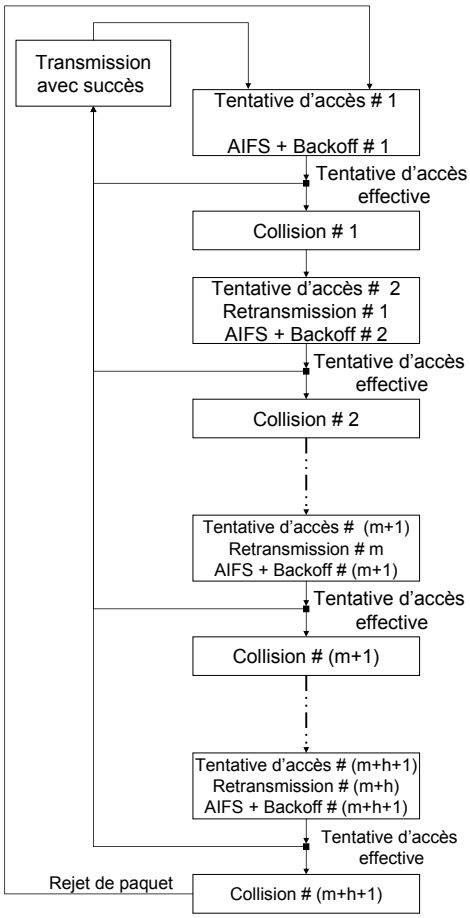

Fig. 2.1 – *Comportement d'une AC_i*

Nous donnons dans la figure 2.1 une vue fonctionnelle d'une catégorie d'accès AC_i ($i \in (0,1,2,3)$ représentant la priorité de l' AC par ordre décroissant). La transmission d'un paquet

consiste en une série de tentatives d'accès. Chaque tentative d'accès se décompose dans les faits en trois processus : le premier (AIFS) représente le test d'occupation du médium ; le deuxième (Backoff) représente l'évitement des collisions (*Collision Avoidance*) ; le troisième est la tentative effective de transmission : la file décide d'envoyer le paquet sur le médium ce qui aura pour conséquence :

- soit une transmission avec succès au terme de laquelle la procédure d'envoi d'un nouveau paquet sera entamée,
- soit une collision (virtuelle ou réelle) suite à laquelle le paquet sera retransmis ou rejeté si le nombre de retransmissions limite est atteint.

2.2.2 Un modèle en blocs : bases de la modélisation

Préalablement à la présentation du modèle, il est nécessaire de donner un certain nombre de définitions qui seront indispensables à la compréhension de la suite. Nous allons donc dans la suite de ce paragraphe définir :

- Certaines des variables qui régissent le modèle.
- Les différents blocs qui constituent le modèle général.
- La grammaire d'un état de la chaîne de Markov et la signification des différentes dimensions.
- Certaines des probabilités que nous allons retrouver dans la définition du modèle.

2.2.2.1 Définitions

AC_i est la catégorie d'accès modélisée, cette catégorie d'accès évolue dans un **environnement local** : elle subit l'effet des autres catégories d'accès de sa station notamment la collision virtuelle ; et dans un **environnement global** : elle subit l'effet des catégories d'accès des autres stations notamment la collision réelle.

$A[AC_i]$ est la valeur de AIFS correspondant à la catégorie d'accès modélisée. Cette valeur détermine le nombre de *timeslot* consécutifs libres du médium que la fonction d'accès doit observer avant de mettre en œuvre l'évitement de collision basé sur la valeur du compteur de Backoff ($B\_C[AC_i]$ que nous définissons plus bas).

N est une moyenne (en termes de *timeslot*) de la durée d'occupation du médium suite à l'interruption d'une phase d'observation de l'état "médium libre".

$\lceil T_s \rceil$ est le plus petit entier supérieur à la durée d'une transmission avec succès en termes de *timeslot*.

$\lceil T_c \rceil$ est le plus petit entier supérieur à la durée d'une collision en termes de *timeslot*.

m est le nombre de retransmissions au bout desquelles la valeur de la fenêtre de contention atteint la valeur maximale possible : la valeur de la fenêtre de contention pour la $m^{ième}$ retransmission est de $CW_{max}[AC_i]$.

h est le nombre de retransmissions supplémentaires (suite aux m retransmissions) aboutissant à la limite de retransmissions (*Retransmission Threshold*) : ainsi $m + h$ est la limite de retransmissions pour AC_i ; durant ces h retransmissions, la taille de la fenêtre de contention est toujours $CW_{max}[AC_i]$.

$W_j[AC_i]$ est la taille de la fenêtre de contention de la catégorie d'accès ayant subi, pour l'envoi du paquet en cours, j collisions. $W_j[AC_i]$ est compris entre $CW_{min}[AC_i]$ et $CW_{max}[AC_i]$.

La valeur j sera une des dimensions du modèle. $W_j[AC_i]$ aura pour valeurs :

$$\begin{cases} W_0[AC_i] = CW_{min}[AC_i], & j = 0 \\ W_{j+1}[AC_i] = min(2 * W_j[AC_i] + 1, CW_{max}[AC_i]), & 0 \leq j \leq m-1 \\ W_j[AC_i] = CW_{max}, & m \leq j \leq m+h \end{cases}$$

$B\_C[AC_i]$ est la valeur du compteur de backoff de la catégorie d'accès. Cette valeur est choisie de façon aléatoire dans un intervalle de contention $[0, W_j[AC_i]]$ au début de la procédure de AIFS. Si cette valeur est de 0, la transmission effective du paquet sera effectuée dés la fin du dernier *timeslot* de AIFS. Si cette valeur est différente de 0, le compteur sera décrémenté à chaque *timeslot* libre observé jusqu'à 0, suite à quoi une transmission effective du paquet sera tentée.

2.2.2.2 Les blocs constituants

Nous appelons "modèle général d'une catégorie d'accès" le modèle final en chaîne de Markov représentant le comportement de AC_i. Ce modèle général est composé de blocs dont certains sont présents en plusieurs instances. Les blocs composant le modèle général sont :

le bloc AIFS est un bloc instantiable qui représente soit le début d'envoi d'un nouveau paquet soit le début d'une nouvelle tentative d'envoi suite à une collision. Notons que des périodes AIFS existent à deux niveaux dans le fonctionnement de EDCA :
- au début de chaque tentative d'envoi d'un paquet (la première tentative et toutes les tentatives suivants l'occurrence d'une collision) ;
- durant chaque étape de décrémentation du compteur de Backoff, suite à l'occurence d'une occupation du médium.

Le bloc AIFS ne concerne que le premier cas, le deuxième étant pris en compte par le bloc Backoff. Le bloc consiste donc en le déroulement de la période AIFS qui précède le choix de la valeur du compteur de backoff. Si, durant cette procédure, le médium devient occupé, la décrémentation du compteur AIFS est arrêtée pendant N *timeslots* puis reprise du début. À la fin de la procédure AIFS, si le médium est toujours libre, le choix de $B\_C[AC_i]$ dans l'intervalle $[0, W_j[AC_i]]$ est fait. Si $B\_C[AC_i] = 0$, AC_i tentera directement un accès au réseau, sinon la procédure de Backoff prendra son cours jusqu'à ce que le compteur parvienne à 0.

le bloc Backoff est un bloc instantiable qui représente les différentes phases qu'implique la décrémentation du compteur de Backoff (lorsque le choix de sa valeur au niveau du bloc AIFS était différent de 0) avant toute tentative d'envoi du paquet. Pour que le compteur de Backoff soit décrémenté, il faut que le médium soit libre. Si durant cette procédure, le médium devient occupé, la décrémentation de backoff est gelée pendant le temps d'occupation du médium suivi d'une procédure AIFS classique avant que la décrémentation reprenne là où elle s'était arrêtée. Lorsque le compteur parvient à 0 et si le médium est toujours libre, un envoi du paquet sera tenté.

le bloc de transmission avec succès n'est présent qu'une seule fois dans le modèle général, il représente le comportement d'une catégorie d'accès lors de la transmission avec succès d'un paquet de cette catégorie d'accès. Dans ce cas, AC_i occupera le médium pour cette transmission pendant $\lceil T_s \rceil$ *timeslots*. AC_i entamera ensuite la tentative de transmission d'un nouveau paquet.

le bloc de collision est un bloc instantiable, il représente le comportement d'une catégorie d'accès lorsque la dite catégorie d'accès subit une collision virtuelle ou réelle. Si AC_i subit une collision réelle, elle occupera le médium pendant $\lceil T_c \rceil$ *timeslots*. Si AC_i perd une collision virtuelle (disons face à AC_k, plus prioritaire), deux possibilités existent :
- Soit AC_k tente un accès au médium et ne subit pas de collision réelle, auquel cas le médium sera occupé par la station pendant $\lceil T_s \rceil$ *timeslots* ;
- soit AC_k tente un accès au médium et subit une collision réelle, auquel cas le médium sera occupé par la station pendant $\lceil T_c \rceil$ *timeslots*.

2.2.2.3 Grammaire de l'état

Un état au sein de la chaîne de Markov doit être représenté de façon unique et sans ambiguïté. Il doit spécifier plusieurs informations :
- Le nombre de tentatives de transmission qu'a subi le paquet en cours, ceci sera la première dimension de la chaîne de Markov : j .
- La valeur, à un moment donné, du compteur de backoff, cette valeur sera la deuxième dimension de la chaîne de Markov : k.
- L'état temporel du médium, qui correspond à la troisième dimension introduite par Kong et al (Kong 04) : d. Elle indique par sa valeur les différents états de AC_i vis à vis du médium, ces états seront explicités plus bas.

Un état de la chaîne de Markov est donc représenté par un triplet (j, k, d) représentant successivement les trois dimensions. Les trois variables auront, suivant les différents cas, un ensemble de valeurs dont nous détaillons la signification ci-après :
- j : Lorsque $0 \leq j \leq m + h$, j représente le nombre de tentatives de transmissions du paquet : $j = 0$ signifie que le paquet est en cours de première tentative d'envoi, $j = 1$ signifie que le paquet a subi une première collision et est en cours de la deuxième tentative d'envoi et ainsi de suite. La valeur de j permettra d'instancier les différents blocs instantiables (AIFS, Backoff et Collision). Pour les cas où la valeur de j n'est pas signifiante, c'est à dire pour les états où le paquet est en cours de transmission avec succès, nous avons $j = -1$.
- k : Lorsque $0 \leq k \leq W_j[AC_i]$, k représente la valeur du compteur de backoff au début de chaque *timeslot* durant la $j^{ième}$ étape de backoff. Les valeurs possibles de k dépendent donc de j (j influant directement sur la taille de la fenêtre de contention comme on l'a vu précédemment). Comme pour j, nous avons choisi d'attribuer à k des valeurs négatives lorsque l'on sort du sens premier donné à k (donc dans tous les états où AC_i n'a pas de procédure de backoff en cours). Ces valeurs seront détaillées après la spécification des valeurs de d.
- d : La valeur de d représente une combinaison entre l'état actuel du médium et un compteur de *timeslots* attaché à cet état du médium. $d = 0$ signifie que le médium est libre et que la période AIFS de non occupation du médium est écoulée (c'est donc un état où AC_i est en backoff avec un médium libre). $d > 0$ signifie une de deux choses :
  - le médium est occupé par la catégorie d'accès modélisée ou par une autre catégorie d'accès du réseau
  - le médium est libre et la catégorie d'accès modélisée est en période AIFS.

Notons que k n'a de signification que lors du décomptage du compteur de Backoff. k n'a donc pas de signification dans les états de la phase AIFS qui précède la tentative d'envoi et dans les états de transmission correcte ou de collision. Par contre, dans ces deux types d'états, on peut avoir des valeurs de j et d identiques. Donc, pour éviter des confusions et garder à

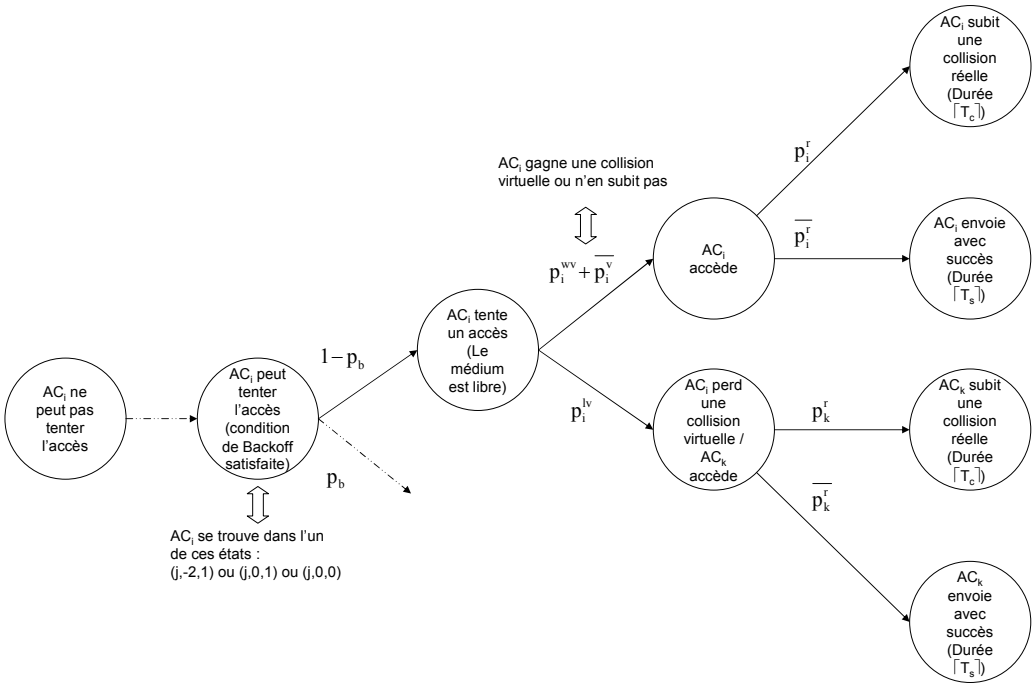

FIG. 2.2 – *Graphe explicitant les différentes probabilités définies*

chacun de ces états son unicité, nous avons choisi de leurs associer des valeurs de k négatives et différentes ($k = -1$ pour la transmission correcte et la collision, $k = -2$ pour la phase AIFS)

2.2.2.4 Les probabilités de transition

Nous définissons dans ce paragraphes les différentes probabilités qui seront utilisées pour les transitions entre les états de la chaîne de Markov. Nous présentons en premier lieu les probabilités basiques utilisées, nous donnerons ensuite d'autres probabilités construites à partie de celles-là. Le graphe de la figure 2.2 visualise les différentes probabilités que nous allons définir et aide à la compréhension de l'enchaînement de situations.

- p_b est la probabilité que le médium devienne occupé (c'est à dire la probabilité que le médium passe de l'état libre à l'état occupé) ; $(1 - p_b)$ est celle que le médium reste libre.
- Nous définissons les probabilités reliées à l'accès de AC_i au médium : $\overline{p_i^v}$ est la probabilité que AC_i, souhaitant accéder au médium, ne subisse pas de collision virtuelle (c'est à dire la probabilité qu'il n'y ait aucune autre AC dans la station de AC_i souhaitant accéder au médium au même *timeslot* que AC_i) ; ${p_i}^{wv}$ est la probabilité que AC_i entre en collision virtuelle avec une autre AC de la station et gagne (c'est à dire la probabilité qu'il y ait au moins une autre AC dans la station de AC_i, moins prioritaire que AC_i, et aucune plus prioritaire, souhaitant accéder au médium au même *timeslot* que AC_i) ; ${p_i}^{lv}$ est la probabilité que AC_i entre en collision virtuelle avec une autre AC de la station et perde (c'est à dire la probabilité qu'il y ait au moins une autre AC dans la station

de AC_i, plus prioritaire que AC_i, souhaitant accéder au médium au même *timeslot* que AC_i). Notons que $\overline{p_i^v} + {p_i}^{wv} + {p_i}^{lv} = 1$.
- p_i^r est la probabilité que AC_i subisse une collision réelle sachant qu'elle accède au médium, . $\overline{p_i^r}$ est la probabilité que AC_i ne subisse pas de collision réelle sachant qu'elle accède au médium. Notons que $p_i^r + \overline{p_i^r} = 1$. Le même raisonnement s'applique, en cas de collision virtuelle perdue de AC_i, à la file gagnant la collision virtuelle (que nous appellerons AC_k) : p_k^r et $\overline{p_k^r}$ seront respectivement la probabilité que AC_k subisse ou non une collision réelle lorsqu'elle accède au médium.
- Le choix du compteur de backoff étant fait de façon aléatoire et uniforme dans l'intervalle de contention $[0, W_j[AC_i]]$, le choix d'une valeur pour le compteur de backoff est donc fait avec une probabilité $\frac{1}{W_j[AC_i]+1}$.

En nous basant sur ces probabilités basiques, il nous est possible de définir d'autres probabilités plus compliquées qui seront utilisées :
- $p_i^{(2)}$ est la probabilité d'une tentative de transmission qui échoue pour AC_i, résultant en une occupation du médium durant $\lceil T_c \rceil$ *timeslots*. Ceci peut être :
  - une collision réelle de AC_i lorsqu'elle accède au médium ;
  - ou une collision réelle de AC_k, accédant au médium suite à une collision virtuelle entre AC_i et AC_k, perdue par AC_i.

  Nous avons donc $p_i^{(2)} = (\overline{p_i^v} + {p_i}^{wv})p_i^r + {p_i}^{lv}p_k^r$.
- $p_i^{(3)}$ est la probabilité d'une tentative de transmission qui échoue pour AC_i, résultant en une occupation du médium durant $\lceil T_s \rceil$ *timeslots*. Ceci est donc résultat d'une collision virtuelle que perd AC_i, laissant place à AC_k qui assurera une transmission avec succès. Nous avons donc $p_i^{(3)} = {p_i}^{lv}\overline{p_k^r}$.
- p_i est la probabilité d'une collision subie par AC_i : qu'elle soit réelle ou virtuelle : $p_i = p_i^{(2)} + p_i^{(3)}$.

2.2.3 Détail des blocs

Nous détaillerons en premier lieu chacun des blocs constituant le modèle général. Nous indiquerons ensuite la manière de connecter et instancier les blocs emmenant au modèle général. Dans chacun des blocs représentés par la suite, les états d'entrée et de sortie du bloc sont représentés en trait gras ; les autres états du bloc sont représentés en trait normal. Les états ne faisant pas partie du bloc en question (c'est à dire les états de sortie ou les états d'entrée des autres blocs connectés au bloc en question) sont représentés en pointillés. Toutes les transitions sont étiquetées de leurs probabilités de transition. Il est à noter que le temps de séjour dans chaque état est d'1 *timeslot* ; il est donc inutile d'indiquer cette information dans la présentation des blocs qui suit. Nous avons intégré des étiquettes textuelles explicatives lorsque ceci s'avère nécessaire.

2.2.3.1 Le bloc AIFS

La figure 2.3 présente le bloc AIFS du modèle général. La figure et ses états sont explicites au vu des définitions données précédemment. Les états :$[(j,-2,1);(j,-2,2);\dots(j,-2,A)]$ sont les états où s'effectue la procédure AIFS *per se*. Nous entrons dans ce bloc soit suite à une collision à l'étape $j-1$ (si $j \neq 0$) ou suite à un *drop* (rejet d'un paquet à cause du dépassement de la limite de retransmissions) ou un envoi avec succès du paquet précédent (si $j = 0$). La fonction d'accès attend que le médium ait été libre pendant A *timeslots* consécutifs avant de pouvoir choisir le compteur de Backoff. Si le médium devient occupé pendant cette

période (les transitions internes vers $(j, -2, N + A)$), la fonction d'accès doit attendre que le médium redevienne libre (nous modélisons ceci par une attente de N *timeslots*) avant de pouvoir re-enclencher la procédure AIFS. Il est spécifié dans le standard (802.11e 05) qu'à la fin du dernier *timeslot* de AIFS, si le médium est libre, la compteur de backoff subit une première décrémentation. Ainsi, lorsque le compteur de Backoff choisi à l'origine est à 0, il est procédé à une tentative de transmission immédiatement après le dernier *timeslot* AIFS, ceci explique les transitions entre l'état $(j, 0, 0)$ et les états de résultat de tentative de transmission : $(-1, -1, \lceil T_s \rceil)$ pour une transmission avec succès, $(j, -1, \lceil T_c \rceil)$ pour une collision résultant en une occupation de $\lceil T_c \rceil$ *timeslots* du médium et $(j, -1, \lceil T_s \rceil)$ pour une collision résultant en une occupation de $\lceil T_s \rceil$ *timeslots* du médium (notons qu'ici réside une des modifications apportée au modèle de Kong et al.). Si le compteur de backoff choisi est strictement supérieur à 0, il est décrémenté et la fonction d'accès passe en procédure de backoff (nous passons ainsi de l'état $(j, -2, 1)$ vers l'un des états $[(j, 0, 0), (j, 1, 0) \dots (j, W_j - 1, 0)]$).

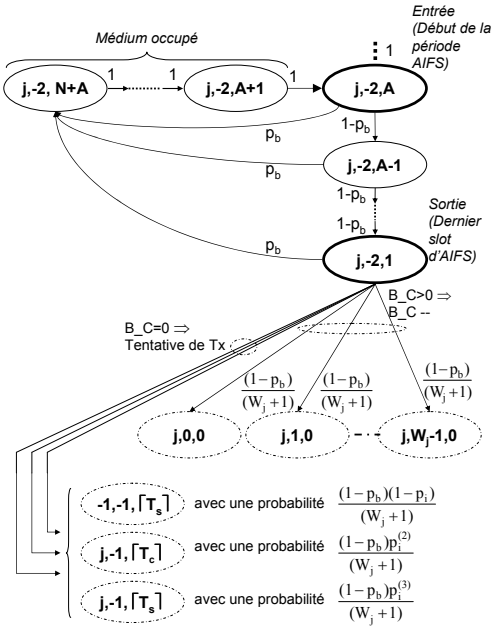

FIG. 2.3 – *Bloc AIFS :* $0 \leq j \leq m + h$

2.2.3.2 Le bloc Backoff

La figure 2.4 présente le bloc Backoff du modèle général. Les états d'entrée du modèle sont $[(j, 0, 0), (j, 1, 0) \dots (j, W_j - 1, 0)]$. La transition entre ces états représente la décrémentation du compteur de backoff tant que le médium est libre. Si le médium devient occupé, la décrémentation est gelée pendant la période d'occupation du médium puis pendant A *timeslots* de médium libre représentant la période AIFS (représenté par l'ensemble des états schématisés au dessus des états de décrémentation du compteur de backoff). La décrémentation reprend dés que le dernier *timeslot* de AIFS libre passe. Si le compteur de backoff est à 0, une tentative de transmission sur le médium est entreprise, aboutissant, comme dans le bloc AIFS, à un parmi trois états : $(-1, -1, \lceil T_s \rceil)$ pour une transmission avec succès, $(j, -1, \lceil T_c \rceil)$ pour une collision résultant en une occupation de $\lceil T_c \rceil$ *timeslots* du médium et $(j, -1, \lceil T_s \rceil)$ pour une collision résultant en une occupation de $\lceil T_s \rceil$ *timeslots* du médium .

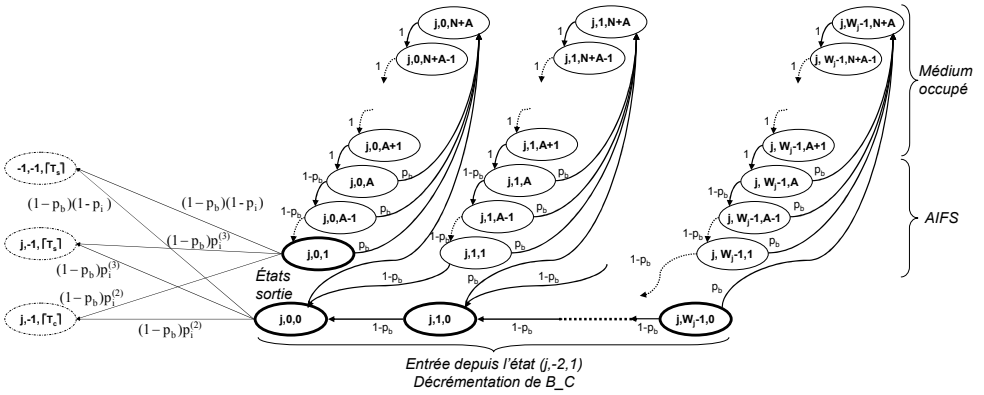

FIG. 2.4 – *Bloc Backoff :* $0 \leq j \leq m+h$

2.2.3.3 Les blocs "résultat d'une tentative de transmission"

La figure 2.5 présente les deux blocs : collision et transmission avec succès du modèle général formant ainsi la partie "résultats d'une tentative de transmission" que nous avons présentée précédemment.

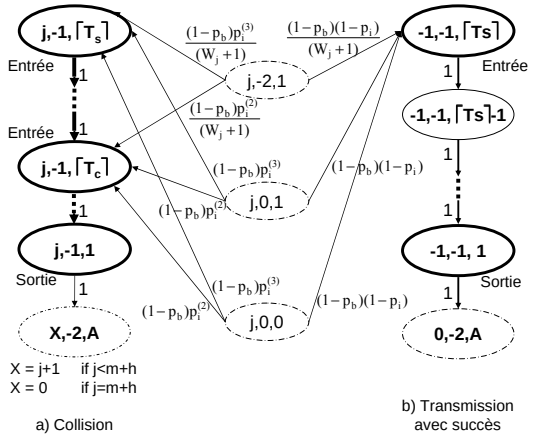

FIG. 2.5 – *Résultats d'une tentative effective de transmission :* $0 \leq j \leq m+h$

Collision La figure 2.5-a présente la partie collision des résultats de tentative de transmission. Cette partie est instantiable par j. On y arrive par l'un de deux états : $(j, -1, \lceil T_s \rceil)$ si la tentative de transmission a abouti à une collision provoquant une occupation pendant $\lceil T_s \rceil$ *timeslots* du médium (voir 2.2.2) ; $(j, -1, \lceil T_c \rceil)$ si la tentative de transmission a abouti à une collision provoquant une occupation pendant $\lceil T_c \rceil$ *timeslots* du médium. Nous considérons ici que $\lceil T_s \rceil > \lceil T_c \rceil$ ce qui est généralement le cas. Le bloc consiste en l'attente pendant un certain nombre de *timeslots* d'occupation du médium avant de procéder à une de deux choses :

- Si la limite sur le nombre de retransmissions est atteinte ($j = m+h$), le paquet est rejeté (*Drop*) et un nouveau paquet sera sujet à transmission (transition de l'état $(m+h, -1, 1)$ à $(0, -2, A)$)
- sinon ($j \neq m+h$), le paquet aura droit à une nouvelle tentative de transmission (transition de l'état $(j, -1, 1)$ à $(j+1, -2, A)$).

Transmission avec succès La figure 2.5-b présente la partie transmission avec succès des résultats de tentative de transmission. Ce bloc est le seul présent de façon unique dans le modèle. On y arrive par l'état $(-1, -1, \lceil T_s \rceil)$ représentant le début d'une transition avec succès, on y découle le temps de la transmission avant de passer à la tentative de transmission d'un nouveau paquet (transition de l'état $(-1, -1, 1)$ à $(0, -2, A)$)

2.2.4 Construction du modèle général

Comme nous l'avons dit précédemment, le modèle général se construit en instanciant par j ($0 \leq j \leq m + h$) chacun des blocs instanciables : AIFS, Backoff et Collision et en les connectant afin de parvenir à un modèle général respectant le schéma que nous présentions figure 2.1 : la boite "Tentative d'accès #1" du schéma 2.1 sera représenté par un bloc AIFS et un bloc Backoff avec $j = 0$; "Collision #1" sera représentée par un bloc Collision avec $j = 0$ et ainsi de suite. La boite "Transmission avec succès" du schéma 2.1 sera évidemment représentée par le bloc de chaîne de Markov Transmission avec succès.

2.2.5 Comparaison qualitative au modèle de Kong et al.

Le but de ce paragraphe est de montrer les différentes erreurs de conceptions trouvées dans le modèle de Kong et al. (Kong 04), en plus de la mauvaise représentation de la collision virtuelle, qui nous ont menées à intégrer plusieurs modifications dans notre modèle.

- Le modèle de Kong et al. ne considère la période AIFS que lors de la première tentative de transmission ; suite à une collision Kong et al. fait le choix d'une valeur pour le compteur de backoff sans passer par une période AIFS ce qui est contraire au standard. Nous intégrons donc cette correction en combinant à chaque période de backoff une période AIFS correspondante.
- Il est stipulé dans le standard que lorsque la valeur du compteur choisie à la fin de la période AIFS est de 0, il est procédé à une tentative de transmission directement après le dernier *timeslot* de la période AIFS (si le médium est libre). Ceci n'est pas pris en considération dans le modèle de Kong et al.
- La fonction d'accès doit vérifier que le médium est libre juste avant d'envoyer (procédure de CSMA normale). Kong et al. n'intègre pas cette vérification, il en découle une importante erreur dans la conception du modèle, erreur corrigée par celui que nous avons présenté.

2.2.6 Utilisation du modèle

Le modèle ainsi construit est un modèle paramétrable. Les paramètres sont toutes les variables intervenant dans le fonctionnement du modèle. Les valeurs de certaines de ces variables sont déterminées par choix (ou peuvent être sujets à variation dans le cadre d'une étude de l'effet de ces variables) : le paramètre A d'AIFS, les tailles de fenêtre de contention W_j, la limite de retransmissions (les variables m et h), la moyenne d'occupation du médium N entre autres. D'autres dépendent du scénario dans lequel est placée la catégorie d'accès modélisée et de ce que l'on appellera les environnements local et global : les différentes probabilités de collision, la probabilité d'occupation du médium etc.
Ces dernières variables étant dépendantes du scénario, nous proposons deux moyens d'établir leurs valeurs et l'évolution qu'elles subissent :

- La méthode analytique qui consiste à dériver des équations analytiques représentant les différentes variables. Il sera ensuite question d'une résolution de système d'équation non linéaire. Cette méthode pourra être utilisée pour une analyse de performance de la fonction d'accès ou de l'effet des différentes variables.
- La méthode dite par injection consiste à récupérer les valeurs de ces variables, par mesure ou par simulation, et de les injecter dans le modèle. La seule inconnue restante devient la matrice des probabilités en régime stationnaire des états.

2.2.6.1 Usage analytique du modèle

La méthode analytique consiste en l'expression des différentes métriques de performance que l'on souhaite mesurer en fonction des probabilités en régime stationnaire des états du modèle. Certaines des variables modifiant le comportement du modèle doivent être exprimées de façon analytique. Nous donnons ici les résultats nécessaires à l'utilisation analytique du modèle. Les calculs, lourds et dont le seul intérêt réside dans leurs résultats sont donnés en annexe A.

La matrice des probabilités en régime stationnaire Afin de parvenir à la matrice générale des probabilités, nous écrivons la probabilité en régime stationnaire de chaque état de la chaîne de Markov en fonction de l'un d'eux. Nous ferons ensuite usage de l'équation d'harmonie des états (c'est à dire que la somme des probabilités de tous les états vaut 1) pour calculer la probabilité en régime stationnaire de chaque état. Nous notons la probabilité en régime stationnaire d'un état (j,k,d) : $b_{j,k,d}$. La probabilité en fonction de laquelle nous souhaitons écrire toutes les autres est $b_{0,-2,1}$, ce choix peut sembler aléatoire, nous avons en fait choisi le premier état à partir duquel une tentative de transmission peut être entamée. Nous procédons à un calcul par bloc du modèle général avant de généraliser le calcul pour parvenir à l'ensemble de la matrice. Nous écrirons en premier lieu les probabilités, par bloc et par étage de backoff, en fonction de $b_{j,-2,1}$ avant d'écrire celle-là en fonction de $b_{0,-2,1}$ et de généraliser ainsi notre calcul à toute la matrice. Ces calculs sont donnés en annexe A. Nous donnons ici le résultat final.

Le bloc backoff

$$b_{j,k,0} = \frac{W_j - k}{W_j + 1}(1 - p_b)p_i^j b_{0,-2,1}, \quad \forall j, \forall k$$

$$b_{j,k,d} = \frac{W_j - k}{W_j + 1}\frac{p_b}{(1 - p_b)^{d-1}}p_i^j b_{0,-2,1}, \quad \forall j, \quad \forall k, \quad 1 \leq d \leq A$$

$$b_{j,k,d} = \frac{W_j - k}{W_j + 1}\frac{p_b}{(1 - p_b)^{A-1}}p_i^j b_{0,-2,1}, \quad \forall j, \quad \forall k, \quad A \leq d \leq N + A$$

Le bloc transmission avec succès

$$b_{-1,-1,d} = (1 - p_b)(1 - p_i^{m+h+1})b_{0,-2,1}, \quad 1 \leq d \leq \lceil T_s \rceil$$

Le bloc collision

$$\forall j, 0 \leq j \leq m + h$$

$$b_{j,-1,d} = (1 - p_b)p_i^{(3)}p_i^j b_{0,-2,1}, \quad \lceil T_c \rceil + 1 \leq d \leq \lceil T_s \rceil$$

$$b_{j,-1,d} = (1 - p_b)p_i p_i^j b_{0,-2,1}, \quad 1 \leq d \leq \lceil T_c \rceil$$

Le bloc AIFS

$$\forall j, 0 \leq j \leq m+h$$
$$b_{j,-2,d} = \frac{(1-(1-p_b)^A)}{(1-p_b)^{A-1}} p_i^j b_{0,-2,1}, \quad A+1 \leq d \leq N+A$$
$$b_{j,-2,d} = \frac{1}{(1-p_b)^{d-1}} p_i^j b_{0,-2,1}, \quad 1 \leq d \leq A$$

Le système ainsi obtenu n'est pas solvable.

Conséquence de l'équation d'harmonie Nous utilisons pour résoudre le système l'équation d'harmonie de la chaîne de Markov. Nous obtenons donc une expression de $b_{0,-2,1}$ en fonction des variables de la chaîne de Markov :

$$\begin{aligned} b_{0,-2,1} = \Bigg[& \frac{1+Np_b}{2(1-p_b)^{A-1}} \sum_{j=0}^{m+h} p_i^j W_j \\ & + \frac{(1-(1-p_b)^A)(1+Np_b)}{p_b(1-p_b)^{A-1}} \frac{1-p_i^{m+h+1}}{1-p_i} \\ & + (1-p_b)\left(\lceil T_c \rceil p_i + (\lceil T_s \rceil - \lceil T_c \rceil) p_i^{(3)}\right) \frac{1-p_i^{m+h+1}}{1-p_i} \\ & + \lceil T_s \rceil (1-p_b)(1-p_i^{m+h+1}) \Bigg]^{-1} \end{aligned}$$

Définition analytique des probabilités Il reste, afin de pouvoir résoudre le système, d'exprimer les probabilités d'occupation et de collision de façon analytique. Nous différencions trois comportement d'une AC_i qui nous permettrons d'établir ces définitions : *une tentative d'accès* qui peut résulter en une collision virtuelle ou en un accès, *un accès* qui peut résulter en une collision réelle ou à une transmission avec succès.

une tentative d'accès : Soit α_i la probabilité que AC_i tente d'accéder au médium à un *timeslot* donné (la catégorie d'accès se trouve dans un état où elle peut tenter l'accès et le médium est libre), nous pouvons l'exprimer ainsi :

$$\alpha_i = (1-p_b) \sum_{j=0}^{m+h} \left(\frac{1}{W_j+1} b_{j,-2,1} + b_{j,0,1} + b_{j,0,0} \right)$$

α_i est la somme des probabilités en régime stationnaire des états où un accès peut être tenté. L'accès n'est tenté que si le médium est libre. À partir de cette définition, nous pouvons écrire les différentes probabilités de collision virtuelle :

- $\overline{p_i^v} = \prod_{x \in [0..3] x \neq i} (1-\alpha_x)$; AC_i ne rentre en collision virtuelle que si au moins une autre AC essaye d'accéder au médium en même temps qu'elle.
- ${p_i}^{lv} = 0$ si $i = 0$; la file de plus haute priorité (AC_VO) ne perd jamais de collisions virtuelles.
- ${p_i}^{lv} = 1 - \prod_{x<i}(1-\alpha_x)$ si $i \neq 0$; AC_i perd la collision virtuelle si au moins une file de plus haute priorité essaye d'accéder au médium en même temps qu'elle.
- ${p_i}^{wv} = 1 - ({p_i}^{lv} + \overline{p_i^v})$; AC_i a gagné une collision virtuelle dans tous les autres cas.

L'accès : Nous définissons τ_i (et respectivement Γ_s) comme étant la probabilité que AC_i (respectivement la station s) accède effectivement au médium pendant un *timeslot* donné (la catégorie d'accès se trouve donc dans un état où elle peut tenter l'accès, le médium est libre, et

elle ne perd pas de collision virtuelle ou elle n'en subit pas). Nous avons donc :

$$\tau_i = (\overline{p_i^v} + p_i^{wv})\alpha_i \text{ et}$$
$$\Gamma_s = \sum_{i=0}^{i=3} \tau_i$$

Nous avons défini précédemment p_i^r (ou p_k^r) comme étant la probabilité que AC_i (ou AC_k) subisse une collision réelle lorsqu'elle accède au médium. Techniquement, cette probabilité correspond, dans le cas de notre modèle, à la probabilité que la station s où évolue AC_i subisse une collision réelle. En effet, lorsqu'une station accède au médium, elle est sujette à la même probabilité de collision réelle quelle que soit l'AC qui accède effectivement. Cette probabilité ne dépend que de l'activité des AC des autres stations. Autrement dit, lorsqu'une AC quelconque de la station s accède au médium, elle ne subira de collisions réelles que lorsqu'au moins une autre station du réseau tente d'accéder au médium au même instant. Nous pouvons donc écrire :

$$p_i^r = p_k^r = 1 - (1 - \Gamma_1) \dots (1 - \Gamma_{s-1})(1 - \Gamma_{s+1}) \dots (1 - \Gamma_M)$$

M étant le nombre de stations du réseau.
Nous pouvons aussi définir p_b (qui est donc la probabilité que le médium devienne occupé). C'est donc la probabilité qu'au moins une station du réseau accède au médium pour un *timeslot* donné. Nous avons donc

$$p_b = 1 - (1 - \Gamma_1)(1 - \Gamma_2) \dots (1 - \Gamma_M)$$

L'ensemble de ces définitions nous permettra de définir les probabilités $p_i^{(2)}$ et $p_i^{(3)}$ (et donc p_i) utilisées dans le modèle.

2.2.6.2 Utilisation par injection d'information

Le système ainsi obtenu, suivant le scénario, consiste en un système non linéaire de N équations à N inconnues. Il est donc solvable par les méthodes classiques d'analyses numérique. Cependant l'utilisation du modèle à ce niveau se limitera à un usage hors ligne à des fins d'analyse de performances. Notre propos ne se situe pas à ce niveau, nous souhaitons avoir un modèle aisément utilisable et aussi simple que possible. Pour ce faire, un premier moyen serait l'injection d'information. Le modèle ayant pour finalité une utilisation en ligne (*Online*), il sera donc possible d'avoir un certain nombre d'informations par mesure réelle sur le médium. Des informations telles que l'occupation du médium ou les probabilité de collisions peuvent donc être injectées directement dans les calculs de métriques ce qui réduit la complexité d'utilisation du modèle tout en le rendant plus performant (car il base sa vision de l'environnement de AC_i sur de valeurs réelles mesurées et non sur une expression analytique).
Un moyen supplémentaire de simplification de l'utilisation et de la prise en main du modèle serait la réduction du modèle proposé dans le paragraphe 2.3

2.3 Réduction du modèle

Nous avons procédé, afin de simplifier la prise en main du modèle et son utilisation, à une réduction du modèle. Le modèle obtenu après réduction consiste en trois états, ceux que nous

considérons importants du point de vue de l'utilisateur. Ces trois états sont : l'état de début de transmission, l'état de transmission avec succès et l'état de rejet du paquet. Le modèle obtenu est statistiquement équivalent au modèle général présenté ci-avant. Nous avons dérivé, à partir du modèle réduit, des métriques de performance qui peuvent être utilisées, par exemple, par le processus décisionnel d'un contrôle d'admission. Nous présentons en premier lieu, les règles de Beizer (Beizer 71) que nous avons utilisées pour la réduction du modèle. Nous présentons par la suite quelques unes des étapes de réduction avant de décrire le modèle obtenu et les métriques de performance dérivées. Dans les figures qui suivent, une transition entre un état a et un état b sera étiquetée par la probabilité de transition P_{ab} et le temps de séjour conditionnel dans l'état a avant transition vers l'état b, T_{ab}.

2.3.1 Les règles de Beizer

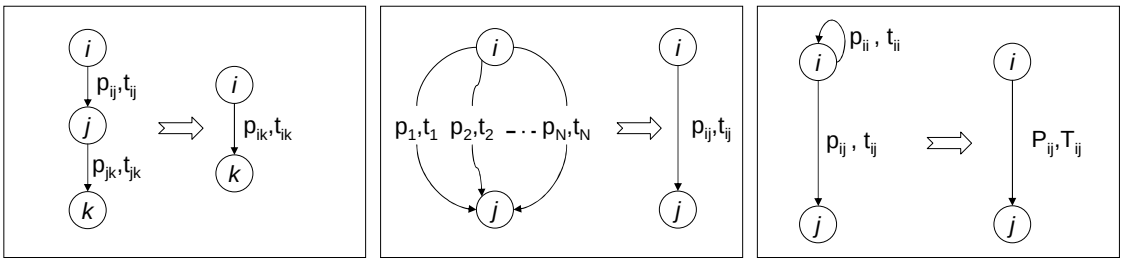

FIG. 2.6 – *Les règles de réduction de Beizer*

Dans (Beizer 71), Beizer détaille un ensemble de règles utilisées pour remplacer un ensemble de transitions et d'états d'un graphe d'états probabiliste temporisé par des liens qui leur sont statistiquement équivalents. Les règles correspondent aux trois situations que l'on peut rencontrer : les liens en série, les liens parallèles et les boucles.

2.3.1.1 La règle "série"

La règle consiste à remplacer une série linéaire de liens par un lien statistiquement équivalent, réduisant ainsi le nœud intermédiaire. La figure 2.6-a illustre cette transformation. Il résulte de la réduction du nœud j un lien avec les caractéristiques suivantes :

$$p_{ik} = p_{ij} \times p_{jk} \text{ et } t_{ik} = t_{ij} + t_{jk}$$

2.3.1.2 La règle "parallèle"

La règle consiste à remplacer un ensemble de transitions reliant deux nœuds par un lien statistiquement équivalent. La figure 2.6-b illustre cette transformation. Il résulte de la réduction des liens reliant le nœud i au nœud j un lien avec les caractéristiques suivantes :

$$p_{ij} = \sum_{k=1}^{N} p_k \text{ et } t_{ij} = \frac{\sum_{k=1}^{N} p_k \times t_k}{\sum_{k=1}^{N} p_k}$$

2.3.1.3 La règle "boucle"

La règle consiste à intégrer l'effet d'une transition "boucle" (c'est à dire une transition ayant pour point de sortie et d'entrée le même nœud) dans les autres liens sortant de ce nœud. La

figure 2.6-c illustre cette transformation. Il résulte de la réduction du lien en boucle sur le nœud i une modification des liens sortants de la façon suivante :

$$P_{ij} = \frac{p_{ij}}{1 - p_{ii}} \text{ et } T_{ij} = t_{ij} + \frac{t_{ii} \times p_{ii}}{1 - p_{ii}}$$

2.3.2 Première phase de réduction : par bloc

Nous procédons, encore une fois au traitement des réductions par bloc afin d'en faciliter la compréhension. Chacun des blocs est réduit à un ou deux états (suivant la stratégie à adopter pour la suite), ensuite l'ensemble des blocs réduits est connecté pour procéder à sa réduction. Nous détaillons, en premier lieu l'abstraction des chacun des blocs AIFS et Backoff avant de mettre en place la connexion entre les blocs et de réduire le modèle complet.

2.3.2.1 Réduction du bloc AIFS

La réduction du bloc AIFS est effectuée au moyen de trois interventions successives consistant en des applications des règles de réduction présentées précédemment.

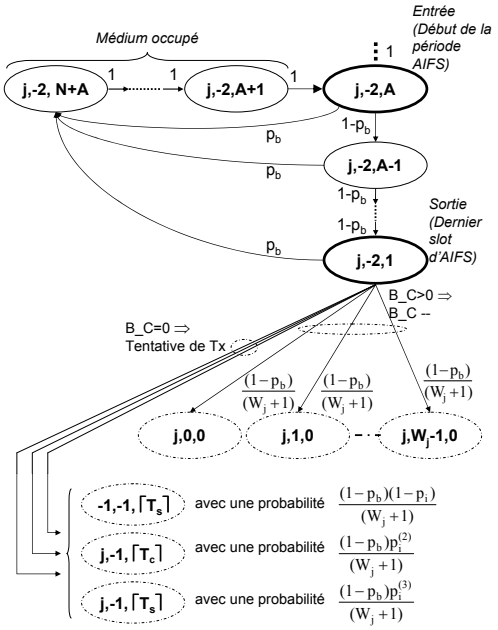

FIG. 2.7 – *Rappel du bloc AIFS :* $0 \leq j \leq m + h$

1. *La première intervention* consiste en une série d'applications : application de règles série entre les états $(j, -2, N + A)$ et $(j, -2, A)$ ce qui donne une transition entre ces deux états de durée N et avec une probabilité de 1 ; application de règles série entre les états $(j, -2, A)$ et $(j, -2, 1)$, ce qui donne une transition de durée $(A - 1)$ avec une probabilité $(1 - p_b)^{(A-1)}$; application conjointe des règles série et parallèle entre les états $(j, -2, A)$ et $(j, -2, N + A)$ (avant d'atteindre l'état $(j, -2, 1)$), ce qui donne une transition avec la probabilité

$$P = p_b + (1 - p_b)p_b + (1 - p_b)^2 p_b + \cdots + (1 - p_b)^{A-2} p_b$$
$$P = 1 - (1 - p_b)^{A-1}$$

et avec une durée

$$T = \frac{p_b + 2(1-p_b)p_b + 3(1-p_b)^2 p_b + \cdots + (A-1)(1-p_b)^{A-2} p_b}{P}$$
$$T = \frac{p_b \sum_{l=1}^{A-1} (l(1-p_b)^{(l-1)})}{P}$$

Suite à cette première intervention, on obtient le graphe représenté sur la figure 2.8.

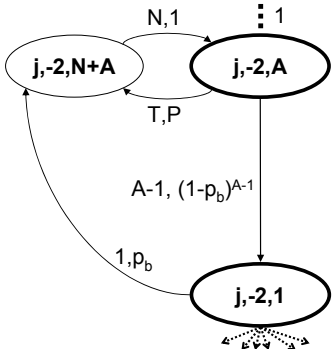

FIG. 2.8 – *Première étape de réduction du bloc AIFS*

2. *La deuxième intervention*, dont le but est de faire abstraction de l'état $(j, -2, N + A)$, consiste en deux applications des règles série, d'une part entre les états $(j, -2, 1)$ et $(j, -2, A)$ (en passant par $(j, -2, N + A)$) et, d'autre part, autour de l'état $(j, -2, A)$ (là encore en passant par $(j, -2, N + A)$). Le résultat de cette deuxième intervention est représenté par le graphe de la figure 2.9.

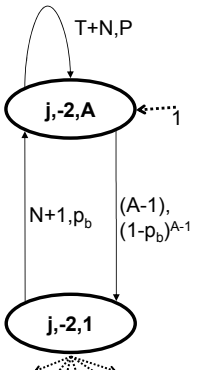

FIG. 2.9 – *Deuxième étape de réduction*

3. *La troisième intervention* concerne la réduction de la boucle autour de l'état $(j, -2, A)$. On obtient le graphe final de la procédure AIFS représenté par le graphe de la figure 2.10 où : $T_{AIFS} = (A-1) + (T+N)\frac{P}{1-P}$.

2.3.2.2 Réduction du bloc Backoff

La réduction du bloc Backoff s'effectue via trois grandes interventions successives impliquant l'application des différentes règles de réduction. Nous rappelons dans la figure 2.11 le croquis du bloc Backoff (nous faisons remarquer dans ce croquis les états d'entrée dans ce bloc à partir de l'état $(j, -2, 1)$, état de sortie du bloc AIFS).

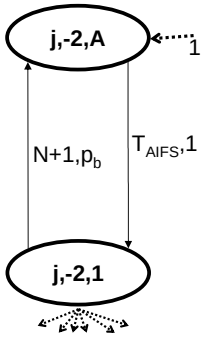

FIG. 2.10 – *Bloc AIFS réduit*

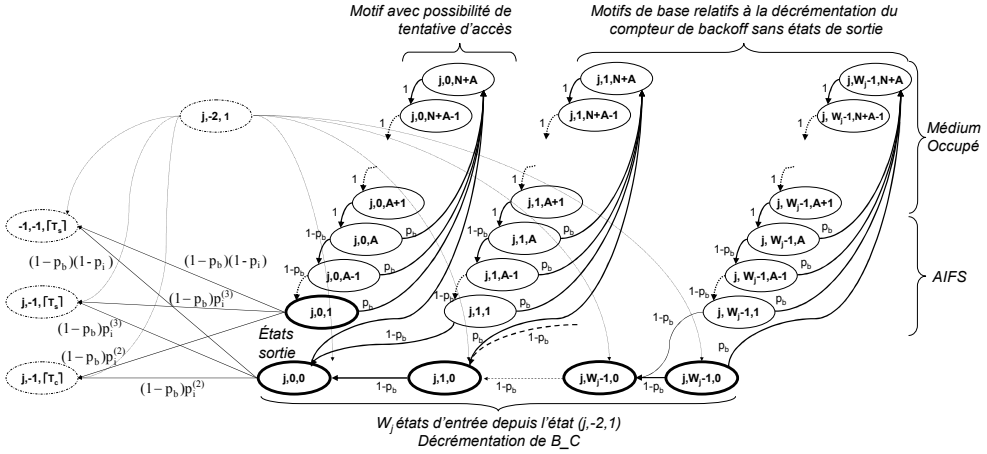

FIG. 2.11 – *Le bloc Backoff initial*

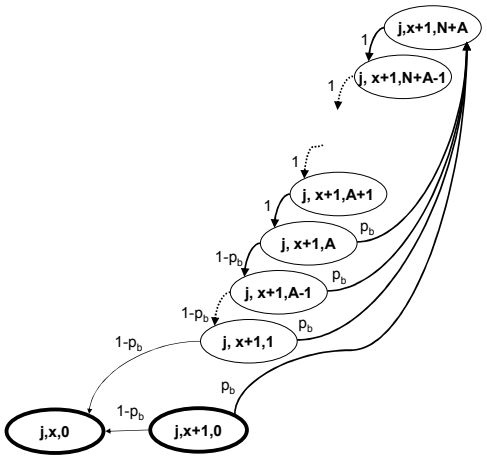

FIG. 2.12 – *Motif de base, ($x \in [0, W_j - 2]$)*

1. *La première grande intervention* concerne en premier lieu la réduction d'un motif de base du processus de décrémentation de Backoff. La figure 2.12 représente un tel motif de base faisant décrémenter la valeur du compteur de backoff (nous passons donc de

l'état $(j, x+1, 0)$ à l'état $(j, x, 0)$. Cette réduction est obtenue de manière similaire à la première intervention de la réduction du bloc AIFS, ce qui donne le graphe de la figure 2.13. On a

$$P' = 1 - (1 - p_b)^A$$

et

$$T' = \frac{p_b + 2(1-p_b)p_b + 3(1-p_b)^2 p_b + \cdots + A(1-p_b)^{A-1} p_b}{P'}$$

$$T' = \frac{p_b \sum_{l=1}^{A} (l(1-p_b)^{(l-1)})}{P'}$$

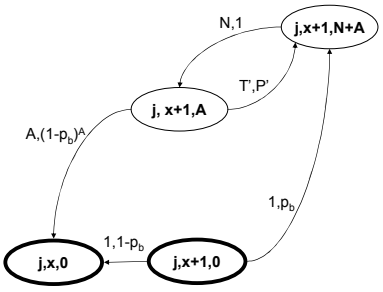

FIG. 2.13 – *Motif de base-première réduction, ($x \in [0, W_j - 2]$)*

Notons que la différence dans les expressions de P' et T' (par rapport aux expressions de P et T du bloc AIFS) résulte du fait qu'ici on doit considérer le départ de l'état $(j, x+1, 1)$, ce qui n'est pas le cas de l'état $(j, -2, 1)$ du bloc AIFS. Cette première grande intervention se termine ensuite en deux phases, tout d'abord la réduction du chemin $(j, x+1, 0)$, $(j, x+1, N+A)$, $(j, x+1, A)$, $(j, x+1, 0)$ en une nouvelle transition entre l'état $(j, x+1, 0)$ et $(j, x, 0)$. Cette nouvelle transition s'obtient au moyen d'interventions similaires aux deuxième et troisième interventions de la réduction du bloc AIFS, elle vient s'ajouter à la transition existante entre $(j, x+1, 0)$ et $(j, x, 0)$. Ces deux transitions, en appliquant la règle parallèle, se réduisent ensuite à une seule transition donnant le graphe de la figure 2.14 où $X = (N + A) + (T' + N)\frac{P'}{1-P'}$.

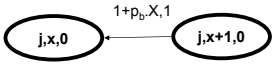

FIG. 2.14 – *Motif de base réduit, ($x \in [0, W_j - 2]$)*

2. *La deuxième grande intervention* consiste en la réduction du graphe représentant le départ de l'état de sortie $(j, -2, 1)$ du bloc AIFS jusqu'à parvenir au début du motif contenant les tentatives d'accès au médium (état $(j, 0, 0)$). Compte tenu du résultat de la première grande intervention, on a donc le graphe de la figure 2.15 que l'on va réduire. Nous avons une succession de réductions série-parallèle qui permettent d'aboutir au graphe représenté sur la figure 2.16.
3. *La troisième grande intervention* concerne plusieurs réductions dont, tout d'abord, la première réduction du premier motif du backoff depuis l'état $(j, 0, 0)$ (d'où on peut avoir

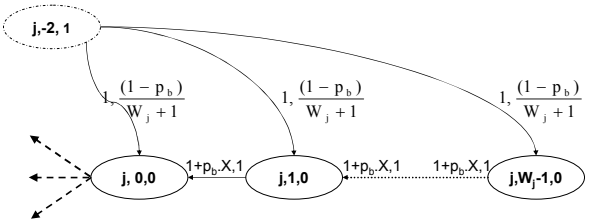

FIG. 2.15 – *Bloc de backoff : les états d'entrée*

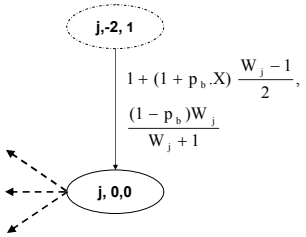

FIG. 2.16 – *Bloc de backoff - Réduction des états de décompte*

des tentatives d'accès si le médium n'est pas occupé) jusqu'à l'état $(j, 0, 1)$ où on peut avoir des tentatives d'accès (si le médium n'est pas occupé). Cette première réduction concerne (comme pour les motifs relatifs à la décrémentation du compteur de Backoff) les réductions relatives aux état de médium occupé et de l'écoulement de AIFS (avec pour simple différence le fait qu'au dernier *timeslot* d'AIFS, un accès peut être tenté si le médium est libre). Cette première réduction donne le graphe de la figure 2.17. Une

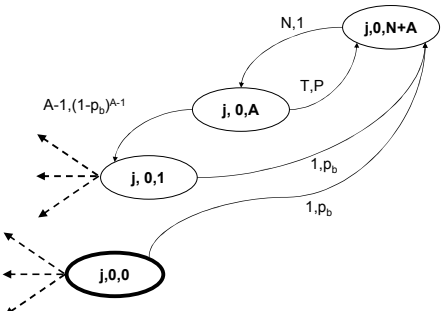

FIG. 2.17 – *Bloc de backoff - Réduction du motif avec* $k = 0$

deuxième réduction consiste à obtenir la transition qui traduit le passage direct de l'état $(j, 0, 0)$ à l'état $(j, 0, 1)$ (application des règles série, parallèle et boucle). On obtient le graphe de la figure 2.18.

2.3.2.3 Réduction du bloc "Tentative d'accès"

Nous représentons dans la figure 2.19 ce que devient la connexion des blocs AIFS et Backoff avec le bloc "Tentative d'accès" suite aux réductions précédemment exposées. Nous procédons à une réduction de ce bloc (figure 2.20) afin de n'y conserver que le résultat d'une tentative d'accès, la réduction est obtenue en utilisant les règles de réduction série et boucle). Nous

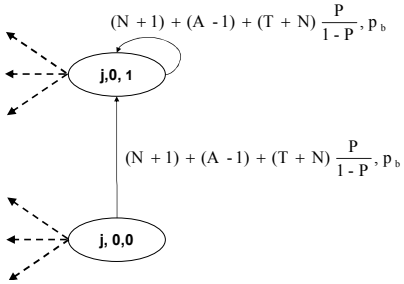

FIG. 2.18 – *Bloc de backoff - Motif avec $k = 0$ réduit*

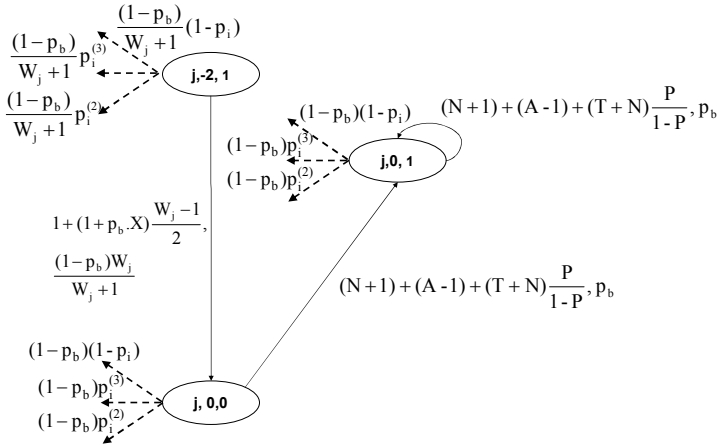

FIG. 2.19 – *Le bloc tentative d'accès*

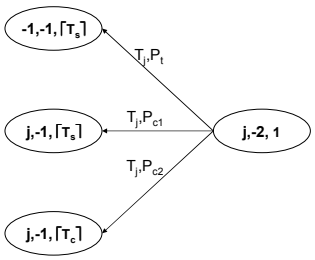

FIG. 2.20 – *Le bloc tentative d'accès - Première réduction*

avons dans la figure 2.20 les probabilités suivantes :

$$
\begin{aligned}
P_t &= (1-p_b)(1-p_i)\frac{1}{W_j+1} \\
&\quad + (1-p_b)(1-p_i)\frac{W_j(1-p_b)}{W_j+1} \\
&\quad + (1-p_b)(1-p_i)\frac{W_j p_b}{W_j+1} \\
P_t &= (1-p_b)(1-p_i)
\end{aligned}
$$

$$P_{c1} = (1 - p_b)p_i^{(2)}$$
$$P_{c2} = (1 - p_b)p_i^{(3)}$$

et le temps (qui représente le temps d'accès par étage de backoff) :

$$\begin{aligned} T_j = &\frac{1}{W_j + 1} + \Big[[1 + \frac{W_j - 1}{2}(1 + p_b X)] + 1\Big] \frac{W_j(1 - p_b)}{W_j + 1} \\ &+ \Big[[1 + \frac{W_j - 1}{2}(1 + p_b X)] \\ &\quad + [(N + 1) + (A - 1) + (T + N)\frac{P}{1 - P}] \\ &\quad + [(N + 1) + (A - 1) + (T + N)\frac{P}{1 - P}]\frac{p_b}{1 - p_b} + 1\Big] \frac{W_j p_b}{W_j + 1} \end{aligned}$$

Ces résultats nous amènent à noter que les probabilités P_t, P_{c1} et P_{c2} sont indépendantes de la valeur de j, quelque soit l'étage de backoff (donc quelque soit le nombre de tentative d'accès que le paquet a déjà vécu) la valeur de ces probabilités restera la même. Par contre, le temps T_{Tj} dépend directement de l'étage de backoff vu qu'il dépend de la taille de la fenêtre de contention (il croit avec la croissance de la taille de la fenêtre de contention).

2.3.3 Modèle de la tentative d'accès j

Suite aux réductions des blocs de base que nous venons de présenter et en intégrant le bloc "résultat d'une tentative d'accès" (figure 2.5), nous pouvons maintenant donner le modèle d'une tentative d'accès complète : depuis le début de la période AIFS de l'étage de backoff en question jusqu'à, soit le début de la transmission avec succès, soit le début de la période AIFS de l'étage de backoff suivant en cas de collision. Ce modèle est représenté dans la figure 2.21. Ce graphe peut ensuite être réduit donnant le graphe de la figure 2.22 dans lequel nous

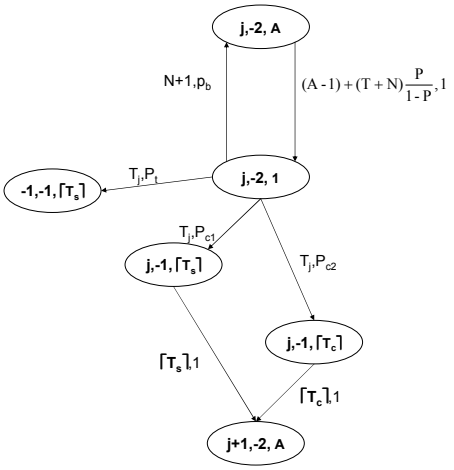

FIG. 2.21 – *le modèle d'une tentative d'accès*

avons :

$$
\begin{aligned}
T_N &= N + 1 \\
T_A &= (A-1) + (T+N)\frac{P}{1-P} \\
P_c &= P_{c1} + P_{c2} \\
Tc_j &= \frac{(T_j + \lceil T_s \rceil)P_{c1} + (T_j + \lceil T_c \rceil)P_{c2}}{P_c} \\
T_{Tj} &= T_j
\end{aligned}
$$

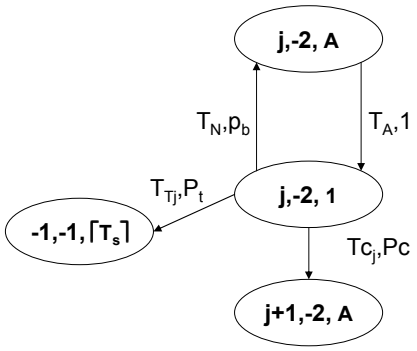

FIG. 2.22 – *le modèle d'une tentative d'accès réduit*

2.3.4 Modélisation de la vue comportementale globale réduite

2.3.4.1 Modèle intermédiaire

Le modèle intermédiaire que nous présentons dans la figure 2.23 représente essentiellement le comportement *Binary Exponential Backoff* : composante principale de l'évitement de collision dans CSMA/CA. Nous présentons ce modèle intermédiaire pour deux raisons : en premier, le modèle détaille les différentes étapes de collision que pourrait subir un paquet en cours de sa transmission. Ceci permettra une éventuelle étude à ce niveau. La seconde raison, qui est plutôt une raison de présentation et de compréhension, est que le modèle intermédiaire correspond au schéma que nous présentions dans la figure 2.1, ceci permettra au lecteur de remettre chacun des modèles ou des blocs réduits dans le contexte du comportement d'une catégorie d'accès EDCA.

Dans la figure 2.23, les états $(j, -2, A)$ (avec $1 \leq j \leq m + h$) représentent les états où une période AIFS est décomptée suite à une collision, avant la reprise d'une tentative de transmission. L'état $(0, -2, A)$ représente le début de la première tentative de transmission d'un paquet. Les états $(j, -2, 1)$ (avec $0 \leq j \leq m + h$) représentent les différentes étapes de Backoff . Celles-ci sont suivies soit par une collision (transition vers l'état $(j + 1, -2, A)$) ou par une transmission avec succès (représentée par l'état $(-1, -1, \lceil T_s \rceil)$). Si l'état de Backoff $(m + h, -2, 1)$ est suivi d'une collision, le paquet est rejeté. Nous avons introduit un nouvel état $(m + h, -1, 0)$ représentant cette situation où le paquet est rejeté. La transition entre les états $(-1, -1, \lceil T_s \rceil)$ ou $(m + h, -1, 0)$ et l'état $(0, -2, A)$ représente la décision prise par la fonction d'accès d'entamer la transmission d'un nouveau paquet suite respectivement à

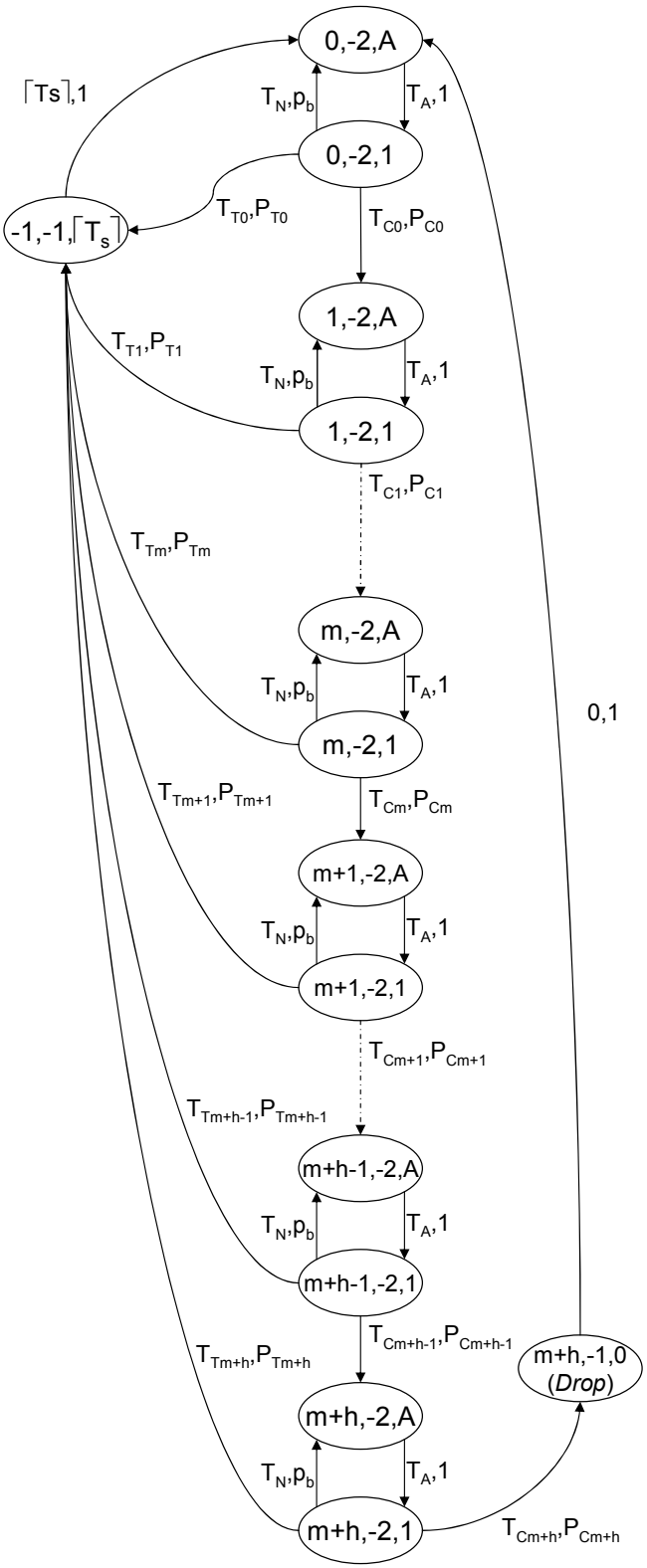

FIG. 2.23 – *Modèle intermédiaire*

une transmission avec succès ou à un rejet de paquet. Les différents temps de séjours et probabilités de transition utilisés dans ce graphe sont :

- T_A représente la durée de la première période AIFS par tentative de transmission, elle prend en compte les éventuelles durées d'occupation du médium par d'autres catégories d'accès.
- T_{Tj} représente la durée moyenne d'une période de Backoff précédant une transmission avec succès.
- T_{cj} représente la durée moyenne d'une période de Backoff précédant une collision auquel nous rajoutons les durée de collisions.
- P_{Tj} représente la probabilité qu'une tentative de transmission aboutisse à une transmission avec succès.
- P_{cj} représente la probabilité que la tentative de transmission aboutisse à une collision.

2.3.5 Le modèle réduit

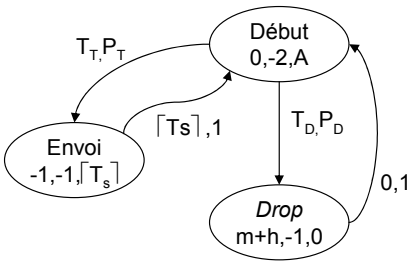

FIG. 2.24 – *Modèle réduit de EDCA*

La figure 2.24 représente le modèle réduit final, obtenu suite à l'application des mêmes procédés de réduction au modèle intermédiaire. Le comportement d'une catégorie d'accès EDCA est réduit aux trois états intéressants du point de vue de l'utilisateur du modèle : l'état de tentative d'accès $(0, -2, A)$ étiqueté l'état 1, l'état de transmission avec succès $(-1, -1, \lceil T_s \rceil)$ étiqueté 2 et l'état de rejet d'un paquet $(m + h, -1, 0)$ étiqueté 3.
Les probabilités P_D et P_T respectivement représentent la probabilité qu'un paquet soit rejeté (P_D) ou transmis avec succès (P_T). Nous avons : $P_D = p_i^{m+h+1}$, un paquet est rejeté s'il a subi une collision à chacune des (m+h+1) tentative de transmission.
Nous avons aussi $P_T = 1 - p_i^{m+h+1}$ qui est donc, de façon évidente, le complément de P_D.
T_D et T_T représentent respectivement la durée qui sépare le début de la première tentative de transmission d'un paquet et le moment où il est rejeté (T_D) ou transmis avec succès (T_T).

$$T_D = \sum_{j=0}^{m+h} (T_A + \frac{(T_A + T_N)p_b}{1 - p_b} + T_{cj})$$

T_D est la somme de toutes les durées de collisions ($m + h + 1$ collisions en tout), auquel nous rajoutons les périodes AIFS, les périodes de Backoff et les périodes où le médium est occupé par une autre catégorie d'accès.

$$T_T = \frac{1 - p_i}{1 - p_i^{m+h+1}} \sum_{j=0}^{m+h} \left[p_i^j (T_A + \frac{(T_A + T_N)p_b}{1 - p_b} + T_{Tj}) + \sum_{l=0}^{j-1} (T_A + \frac{(T_A + T_N)p_b}{1 - p_b} + T_{cl}) \right]$$

T_T intègre pour chacune des valeurs possibles de j ($0 \leq j \leq m + h$) deux termes : la durée d'une transmission avec succès à la $j^{ième}$ tentative et les durées de collisions qui l'ont précédée

(tout en rajoutant les périodes AIFS, les périodes de Backoff et les périodes où le médium est occupé par une autre catégorie d'accès).

2.3.5.1 Régime stationnaire

Les probabilités en régime stationnaire du modèle réduit sont, pour chacun des états 1, 2 et 3 : ($\Pi_1 = \frac{1}{2}$; $\Pi_2 = \Pi_1 P_T$; $\Pi_3 = \Pi_1 P_D$).

2.3.6 Métriques dérivées du modèle réduit

À partir de ce modèle réduit, nous pouvons aisément obtenir certaines métriques de performance essentielles du point de vue de l'utilisateur.

Le débit maximal : le *throughput* Comme nous l'avons déjà dit, le modèle décrit le comportement d'une catégorie d'accès saturée dans un environnement décrit par les différentes probabilités d'occupation du médium et de collisions (p_b, $p_i^{(2)}$ et $p_i^{(3)}$). Nous pouvons dériver une moyenne du débit maximal que peut atteindre la catégorie d'accès dans ces conditions : c'est le rapport entre le temps de transmission utile et la somme des séjours dans chacun des états.

$$Throughput_i = \frac{\Pi_2 Payload}{\Pi_1(P_T T_T + P_D T_D) + \Pi_2 T_s + \Pi_3 \times 0}$$

Payload étant le nombre de *timeslots* nécessaire à la transmission de la partie utile du paquet (notons que le temps de séjour dans l'état 3 est nul, cet état est virtuel et ne sert qu'à signifier le rejet d'un paquet). Nous avons donc :

$$Throughput_i = \frac{P_T Payload}{P_T T_T + P_D T_D + P_T \lceil T_s \rceil}$$

Le délai d'accès moyen Le délai d'accès moyen d'un paquet est le temps qui sépare la prise en considération d'un paquet pour une première tentative de transmission et le début de sa transmission avec succès. Ceci correspond donc à T_T.

La probabilité de rejet d'un paquet Cette probabilité correspond à P_D. C'est la probabilité pour un paquet de la catégorie d'accès d'être au final rejeté à cause du nombre de collisions qu'il subit.

2.4 La problématique de la gestion des collisions virtuelles : proposition d'une stratégie

Le modèle que nous avons développé, à la fois dans sa version générale et dans sa version réduite, peut servir dans divers cas d'applications. Une première application est l'évaluation de performances du schéma d'accès EDCA. Dans ce paragraphe, nous nous plaçons dans ce cas en considérant la problématique particulière de la gestion des collisions virtuelles. Nous proposons une nouvelle stratégie pour cette gestion et faisons la comparaison entre la gestion décrite dans le standard et cette nouvelle stratégie. Cette stratégie induit des modifications, dans le modèle, que nous expliquons.

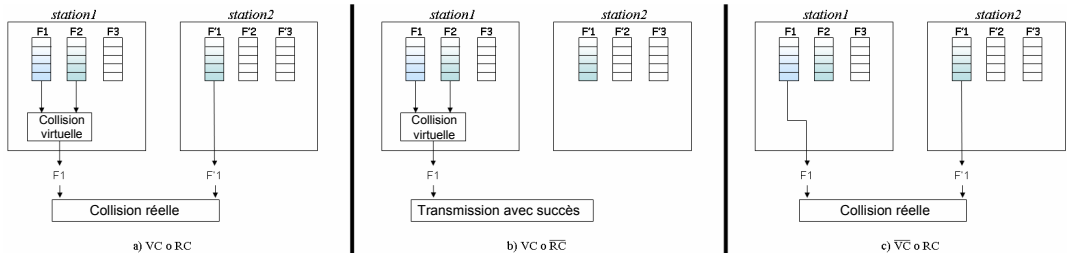

FIG. 2.25 – *Différents scénarios de collisions : a)*$VC \circ RC$; *b)*$VC \circ \overline{RC}$; *c)*$\overline{VC} \circ RC$

2.4.1 Contexte

EDCA a été conçu pour introduire de la différenciation à l'accès au médium entre des catégories d'accès différentes. En effet, les paramètres protocolaires des fonctions d'accès EDCAFs sont supposés favoriser statistiquement les AC_VO par rapport aux AC_VI, les AC_VI par rapport aux AC_BE et celles-ci par rapport aux AC_BK. Puisque les EDCAFs associées aux ACs d'un même type utilisent les mêmes paramètres protocolaires, EDCA est supposé être équitable entre ACs de même priorité, leur garantir donc, dans des conditions équivalentes, la même probabilité d'accès au médium. Nous montrons dans la suite que la gestion de la collision virtuelle introduite par EDCA introduit de l'iniquité entre des EDCAF de même priorité. Cette notion prend toute son importance en cas de différence de charge entre les stations. Or dans les scénarios usuels des réseaux sans fil, le trafic est souvent asymétrique. Le trafic descendant est beaucoup plus important que le trafic montant par station : le point d'accès sera donc beaucoup plus chargé que les autres stations. Nous proposons une modification de la gestion de la collision virtuelle que nous analyserons.

2.4.2 Description du problème

2.4.2.1 Définitions

Afin d'illustrer ce problème d'iniquité, nous définissons les différents scénarios qui peuvent se produire lors d'un accès d'une AC pour transmission. Les situations sont représentées dans la figure 2.25 :

- $VC \circ RC$: Cas où une collision virtuelle (VC) a lieu au sein d'une station suivie d'une collision réelle (RC) (figure 2.25-a).
- $VC \circ \overline{RC}$: Cas où une collision virtuelle (VC) a lieu au sein d'une station non-suivie d'une collision réelle ($\overline{RC}$) (figure 2.25-b).
- $\overline{VC} \circ RC$: Cas où une collision réelle (RC) a lieu au sein d'une station qui n'était pas précédée de collision virtuelle ($\overline{VC}$) (figure 2.25-c).
- $\overline{VC} \circ \overline{RC}$: Cas où aucune collision n'a lieu, une catégorie d'accès transmet avec succès au médium (non représenté sur la figure 2.25).

2.4.2.2 Le problème de l'iniquité de EDCA

Le but premier de la gestion de collision est de protéger et d'améliorer l'utilisation totale du médium ce qui, dans certains scénarios de collision virtuelle, n'est pas toujours vrai. Soit une station avec les catégories d'accès AC_VO et AC_VI actives. Notons que les paramètres d'accès des deux AC sont proches : elles ont la même valeur pour AIFS et des intervalles

de contention proches ce qui implique que les collisions virtuelles entre les deux AC seront fréquentes. Lorsque les deux fonctions d'accès rentrent en collision virtuelle, la fenêtre de contention de la fonction attachée à AC_VI sera doublée, celle de l'AC_VO pourra accéder au médium. Si cet accès se déroule sans collision réelle (scénario $VC \circ \overline{RC}$), il semble sans intérêt de pénaliser la fonction AC_VI. Cette pénalisation aura plusieurs effets pervers. Le premier de ceux-là est l'éventuelle inversion de priorité. Lorsque la fenêtre de contention de la fonction attachée à AC_VI se voit doublée, elle pourra se trouver en situation moins prioritaire que des fonctions attachées à des AC normalement moins prioritaire. Ceci est rare si les valeurs par défaut (802.11e 05) sont utilisées. L'autre problème est celui de l'iniquité. EDCA est censé donner la même chance d'accès aux ACs de la même priorité d'une BSS. Cependant, comme illustré ci-après, la chance d'accès devient dépendante du contexte local. Nous définissons le contexte local d'une catégorie d'accès comme étant les catégories d'accès actives au sein de la station et leurs profils d'arrivée. Soit deux stations au sein d'une même BSS, la première *station1* a les catégories d'accès AC_VO et AC_VI actives. La seconde *station2* a uniquement AC_VI active. Les deux catégories AC_VI ont le même profil d'arrivée mais sont placées dans des contextes locaux différents. Pour être *équitable*, EDCA devrait donner une chance d'accès égale aux deux catégories AC_VI. Cependant, la catégorie d'accès AC_VI de la *station1* subissant la collision virtuelle aura une fenêtre de contention en moyenne supérieure à celle de la *station2* et par conséquent le débit utile de la première sera inférieur à celui de l'autre.

2.4.3 Propositions de modification

2.4.3.1 Solution théorique : la catégorie d'accès omnisciente

Pour pouvoir assurer une équité parfaite entre les catégories d'accès d'une même priorité, il faudrait que les paramètres d'accès EDCA de toutes les fonctions d'accès d'une priorité aient toutes les mêmes valeurs à tout moment. Il faudrait donc appliquer la politique de gestion de collisions suivante : si une collision a lieu, qu'elle soit réelle ou virtuelle, toutes les fonctions d'accès y participant et toutes les fonctions de même priorité doivent subir le même traitement : le doublement de leurs fenêtres de contention. Si une fonction d'accès parvient à envoyer avec succès un paquet sur le médium, toutes les files de la même priorité doivent en bénéficier en réinitialisant leurs fenêtres de contention. Cette vue nous fait considérer un système affecté globalement par des événements locaux assurant ainsi l'équité recherchée. Ce mode de fonctionnement n'est évidemment pas réalisable, il faudrait que toutes les stations soient omniscientes et au courant de ce qui se passe au niveau des autres stations. Il faudrait donc en dériver le meilleur mode de fonctionnement réalisable.

2.4.3.2 Modification

Afin de résoudre ce problème, nous proposons la solution *"Pénalisation_VC_Conditionnelle"*. Cette modification est basée sur le raisonnement suivant : la gestion de la collision virtuelle définie par 802.11e se justifie dans un cadre où les collisions réelles sont fréquentes. Cependant, dans un cadre de trafic asymétrique, un scénario avec beaucoup de collisions virtuelles (au niveau du point d'accès) et peu de collisions réelles devient fréquent. Dans ce cadre, il serait intéressant de ne pas augmenter la taille de la fenêtre de contention en cas de $VC \circ \overline{RC}$, nous réduirons ainsi l'iniquité sans pour autant risquer de détériorer l'utilisation du médium (l'absence de collision réelle est synonyme de non surcharge ponctuelle du réseau). En re-

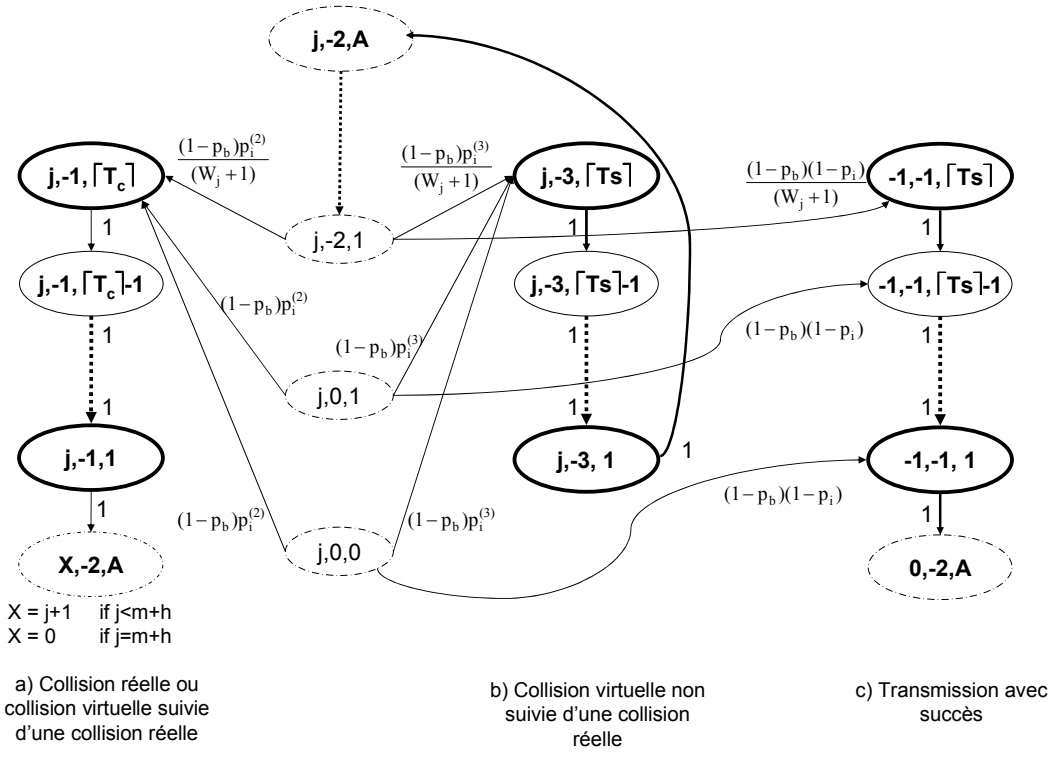

FIG. 2.26 – *Version modifiée du bloc "tentative d'accès" intégrant la modification proposée*

vanche, si une collision réelle se produit après l'occurrence de la collision virtuelle (scénario $VC \circ RC$), toutes les files qui ont subi au préalable la collision virtuelle sont sanctionnées afin d'éviter une nouvelle collision réelle.

Formellement, pour une catégorie d'accès AC_i donnée, EDCA devra fonctionner ainsi :

1. si AC_i perd une collision virtuelle sans que celle-ci soit suivie d'une collision réelle (scénario $VC \circ \overline{RC}$), AC_i n'est pas pénalisée.
2. si AC_i perd une collision virtuelle et que celle-ci est suivie d'une collision réelle (scénario $VC \circ RC$), AC_i est pénalisée.
3. dans tous les autres cas, le comportement classique d'EDCA est respecté.

La pénalisation d'une fonction d'accès se traduit par le doublement de la fenêtre de contention dans la limite de l'intervalle de la fenêtre de contention. Nous étudierons cette proposition et la comparerons au comportement classique d'EDCA en utilisant le modèle.

2.4.4 Analyse

2.4.4.1 Le modèle EDCA modifié

Afin d'évaluer la performance de la proposition que nous faisons dans ce chapitre nous avons eu à introduire cette modification dans le modèle général que nous avons développé. Cette modification intervient dans le bloc "tentative d'accès". Á ce niveau, la modification du comportement se fera lorsqu'une collision virtuelle est non suivie de collision réelle, la catégorie d'accès subissant la collision virtuelle devra rester dans le même étage dans lequel elle se trouve. Nous avons, pour cela, introduit un ensemble d'états représentant cette situation. Nous introduisons une nouvelle valeur de k, les états avec $k = -3$ (voir la section 2.2.2.3

pour une explication générale des valeurs négatives de k) sont les états où l'AC a subi une collision virtuelle mais que celle-ci n'était pas suivie d'une collision réelle (scénario $VC \circ \overline{RC}$). Nous définissons donc les états $(j, -3, d)$ avec $0 \leq j \leq m+h$ et $0 \leq d \leq \lceil T_s \rceil$. Le nouveau bloc "tentative d'accès" est représenté dans la figure 2.26. Cette modification implique une modification des équations des probabilités en régime stationnaire que nous résumons ainsi :
Soit $Y = \frac{p_i^{(2)}}{1-p_i^{(3)}}$

Le bloc backoff

$$b_{j,k,0} = \frac{W_j - k}{W_j + 1}(1-p_b)Y^j b_{0,-2,1}, \quad \forall j, \forall k$$
$$b_{j,k,d} = \frac{W_j - k}{W_j + 1}\frac{p_b}{(1-p_b)^{d-1}}Y^j b_{0,-2,1}, \quad \forall j, \quad \forall k, \quad 1 \leq d \leq A$$
$$b_{j,k,d} = \frac{W_j - k}{W_j + 1}\frac{p_b}{(1-p_b)^{A-1}}Y^j b_{0,-2,1}, \quad \forall j, \quad \forall k, \quad A \leq d \leq N + A$$

Le bloc transmission avec succès

$$b_{-1,-1,d} = (1-p_b)(1-p_i^{(3)})(1-Y^{m+h+1})b_{0,-2,1}, \quad 1 \leq d \leq \lceil T_s \rceil$$

Le bloc collision

$$\forall j, 0 \leq j \leq m+h$$
$$b_{j,-1,d} = (1-p_b)p_i^{(2)}Y^j b_{0,-2,1}, \quad 1 \leq d \leq \lceil T_c \rceil$$
$$b_{j,-3,d} = (1-p_b)p_i^{(3)}Y^j b_{0,-2,1}, \quad 1 \leq d \leq \lceil T_s \rceil$$

Le bloc AIFS

$$\forall j, 0 \leq j \leq m+h$$
$$b_{j,-2,d} = \frac{(1-(1-p_b)^A)}{(1-p_b)^{A-1}}Y^j b_{0,-2,1}, \quad A+1 \leq d \leq N+A$$
$$b_{j,-2,d} = \frac{1}{(1-p_b)^{d-1}}Y^j b_{0,-2,1}, \quad 1 \leq d \leq A$$

Conséquence de l'équation d'harmonie

$$b_{0,-2,1} = \Bigg[\frac{1+Np_b}{2(1-p_b)^{A-1}}\sum_{j=0}^{m+h} Y^j W_j + \frac{(1-(1-p_b)^A)(1+Np_b)}{p_b(1-p_b)^{A-1}}\frac{1-Y^{m+h+1}}{1-Y} + (1-p_b)\left(\lceil T_c \rceil p_i^{(2)} + \lceil T_s \rceil p_i^{(3)}\right)\frac{1-Y^{m+h+1}}{1-Y} + \lceil T_s \rceil (1-p_b)(1-Y^{m+h+1})\Bigg]^{-1}$$

2.4.4.2 Analyse

Nous mettons en place divers scénarios pour analyser la modification en comparant sa performance à celle de EDCA. Nous faisons varier, dans les scénarios, les probabilités de collisions réelles ou de collisions virtuelles. Le modèle étant celui d'une catégorie d'accès saturée, varier le taux de collisions virtuelles revient à varier une des variables du modèle : la taille $\lceil T_s \rceil$. En effet, en réduisant le temps que passe une station à envoyer avec succès, nous augmentons

son taux d'accès (à cause de la caractéristique saturée du modèle). Ainsi, faire varier la valeur de $\lceil T_s \rceil$ au sein d'une station ayant au moins deux catégories d'accès actives fera varier le taux de collisions virtuelles (si le taux d'accès de chacune des catégories d'accès augmente, le taux de collisions virtuelles augmentera). Afin de varier le taux de collisions réelles, il suffit de faire varier le nombre de stations actives sur le réseau. Nous commençons par présenter les scénarios et les résultats avant de les analyser.

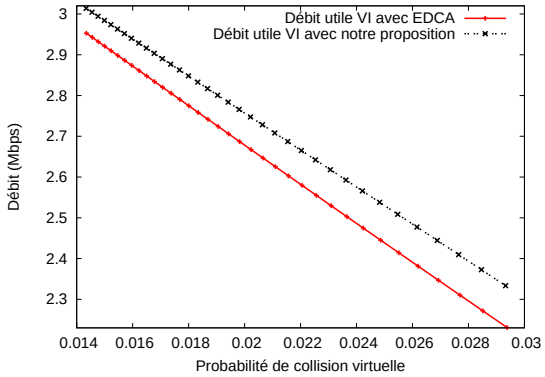

FIG. 2.27 – *Scénario 1, Débit utile de AC_VI*

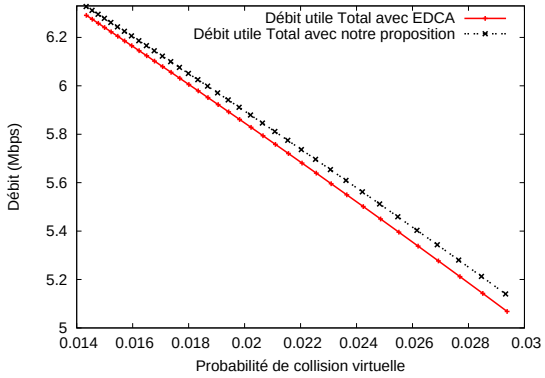

FIG. 2.28 – *Scénario 1, Débit utile total*

Scénario 1 Pour le scénario 1, le réseau ne contient qu'une seule station ayant des transmissions actives, la station a les catégories d'accès AC_VO et AC_VI actives. Nous varions dans ce scénario le taux de collisions virtuelles. Le taux de collisions réelles est nul vu qu'il n'y a qu'une seule station qui transmet. Nous comparons la performance de notre proposition à celle de EDCA en comparant le débit utile de la catégorie d'accès AC_VI adoptant l'un ou l'autre comportement (figure 2.27). Nous regardons aussi le débit total de la station avec EDCA et avec notre proposition (figure 2.28).

Scénario 2 Dans le scénario 2, nous mettons en place le scénario présenté ci-avant (paragraphe 2.4.2) pour exposer le problème de l'iniquité. Deux stations sont actives dans le réseau, la première avec les catégories d'accès AC_VO et AC_VI actives (c'est une station de type 1). La seconde station a uniquement la catégorie d'accès AC_VI active (c'est une station de type 2). L'équité dans ce cadre signifie que : même si l'AC_VI de la station de type 1 est "voisine"

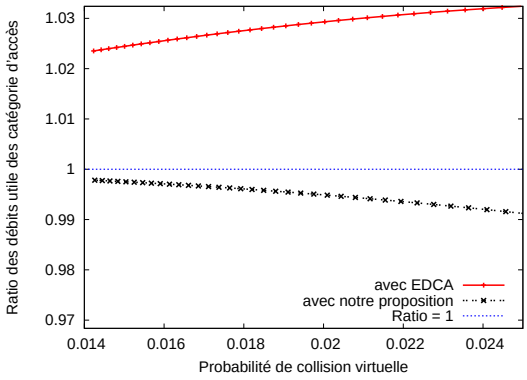

FIG. 2.29 – *Scénario 2*

d'une AC_VO, l'indépendance des fonctions d'accès implique que les deux filles AC_VI devraient au final avoir le même débit utile. Nous varions dans ce scénario le taux de collisions virtuelles comme nous l'avions expliqué plus haut. Varier ainsi le taux de collisions virtuelles n'aura que très peu d'effet sur le taux de collisions réelles vu qu'il n'y a que deux stations actives dans ce scénario. Nous comparons la performance de notre proposition et celle de EDCA en analysant le ratio entre le débit utile des deux catégories d'accès AC_VI. Plus ce ratio s'approche de 1 et plus le comportement peut être considéré comme équitable. La figure 2.29 présente ce résultat en fonction du taux de collisions virtuelles.

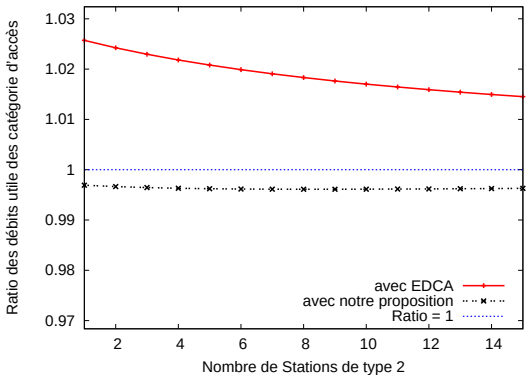

FIG. 2.30 – *Scénario 3*

Scénario 3 Dans le scénario 3 nous avons une seule station de type 1 et un nombre N variant de stations de type 2. Ce scénario nous permettra d'analyser l'effet de la variation de la collision réelle sur le comportement de notre proposition. La variation du taux de collisions virtuelles sera faible dans ce cadre par rapport à celle du taux de collisions réelles. La figure 2.30 donne le ratio du débit d'une AC_VI ne subissant pas de collisions virtuelles au débit de l'AC_VI de la station de type 1 subissant des collisions virtuelles (il est à noter que le débit de toutes les files AC_VI ne subissant pas de collisions virtuelles est le même). Les Figure 2.31 et 2.32 comparent le débit utile total de toutes les catégories d'accès du réseau adoptant l'une ou l'autre approche (nous divisons les résultats en deux phases, une avec un faible nombre de stations, l'autre avec un nombre de stations plus élevé).

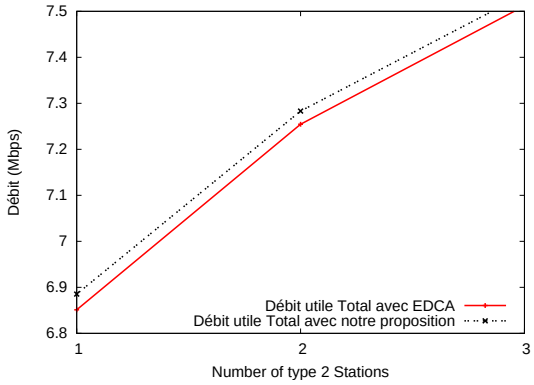

FIG. 2.31 – *Scénario 3 - débit total, nombre de stations faible*

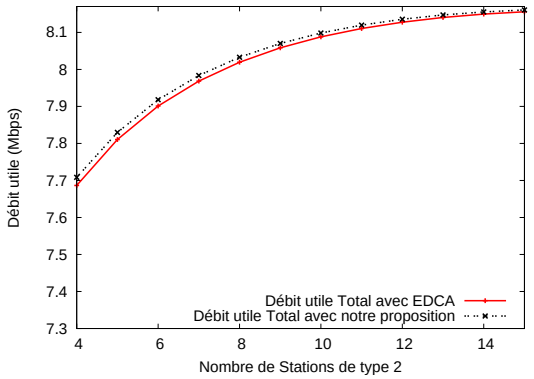

FIG. 2.32 – *Scénario 3 - débit total, nombre de stations élevé*

Analyse Le premier scénario montre que notre modification améliore le débit utile de la catégorie d'accès AC_VI (figure 2.27) dans un contexte où elle n'est sujette qu'à des collisions virtuelles ; dans ce cadre, puisque la fenêtre de contention de la catégorie AC_VI n'est jamais doublée avec notre proposition, AC_VI aura un meilleur débit. Nous montrons aussi dans la figure (figure 2.28) que notre approche améliore le débit utile total de la station quelque soit le taux de collisions virtuelles. L'augmentation du taux de collisions virtuelles est accompagnée en général d'une diminution du débit en général. Ceci est dû à la manière qu'on emploie pour faire croître le taux de collisions virtuelles : en diminuant le temps que la fonction passe lors d'un envoi avec succès (donc techniquement ça correspond à une diminution de la taille des paquets envoyés). Ainsi, chaque accès avec succès portera moins d'information utile, le débit utile total sera donc moindre.
Le second scénario montre (figure 2.29) que lorsque le taux de collisions virtuelles augmente, l'iniquité augmente avec EDCA et avec notre proposition (chacune dans un sens, les courbes s'éloignant de 1). La collision virtuelle étant un événement local, son effet reste local et donc rajoutera à l'iniquité. Il est cependant intéressant de noter que notre proposition inverse et réduit l'iniquité.
Pour le troisième scénario, on peut clairement voir (figure 2.30) que l'augmentation du taux de collisions réelles réduira l'iniquité d'EDCA. En effet, même si les AC_VI des stations de type 2 ne subissent pas de collisions virtuelles, elles se trouvent autant pénalisées que l'AC_VI de la station de type 1 par l'occurrence plus fréquente des collisions réelles. Notre approche réduira tout de même l'iniquité par rapport à EDCA. Le ratio calculé reste stable avec notre proposition : la collision réelle affectant autant les deux types de AC_VI. La figure

2.31 montre que notre approche améliore le débit utile total lorsque le nombre de stations actives (et par conséquent le taux de collisions réelles) est faible. Lorsque le nombre de stations augmente, et par conséquent lorsque le taux de collisions réelles augmente, le comportement de notre approche tendra vers celui d'EDCA.
Ainsi, de façon générale, notre proposition permet de réduire l'iniquité d'EDCA, même si l'équité n'est pas parfaite (une telle équité nécessite la mise en place de notre solution théorique). Elle parvient dans des situations à faible taux de collisions réelles à augmenter l'utilisation du médium, et tend vers le comportement d'EDCA dans les cas où les collisions réelles sont plus fréquentes.

Conclusion

Le modèle que nous avons développé dans ce chapitre est un modèle exhaustif en termes de représentation des mécanismes intervenant dans la méthode d'accès au médium.
En outre, ce modèle considère de manière explicite la collision virtuelle. Nous avons également présenté une abstraction de ce modèle, le modèle réduit, très utile pour un utilisateur.
Ce modèle peut cependant être critiqué à plusieurs niveaux :

- Le modèle ne prend pas en compte l'effet du BEB sur la probabilité de collision, celle-ci devant, dans la majorité des cas, être réduite grâce à ce mécanisme.
- La probabilité d'occupation est aussi fixée alors qu'elle doit varier en fonction de l'état des différentes catégories d'accès.
- Le modèle suppose que la catégorie d'accès est saturée.

D'une manière générale, les deux premiers points posent le problème de l'évaluation des valeurs des probabilités qui interviennent dans le modèle. Nous avons opté pour l'injection par mesures de ces probabilités, les simulations effectuées ont montré l'intérêt de cette solution (Masri 07a).
La critique sur la saturation concerne un grand nombre des modèles proposés. Pour l'utilisation que nous en faisons, le modèle des catégories d'accès saturées répond à nos besoins. Il est cependant intéressant de réfléchir aux moyens de passer d'un modèle saturé à un modèle non saturé, ceci passe par l'intégration de plusieurs mécanismes au modèle :

- La prise en compte du mécanisme de *post-backoff* : suite à un envoi avec succès (ou à un *Drop* - perte) et si la file d'attente de la catégorie d'accès est vide, la fonction d'accès doit réaliser un backoff supplémentaire permettant d'éviter les effets de capture ;
- La prise en compte de profils d'arrivée du trafic et de la taille de la file d'attente modifiant ainsi le comportement de la fonction d'accès. Cette prise en compte peut passer par l'intégration, en parallèle avec le modèle d'un modèle des arrivées.

Nous exposerons dans le chapitre suivant un algorithme de contrôle d'admission utilisant pour son fonctionnement la version réduite de notre modèle EDCA.

Un contrôle d'admission hybride pour IEEE 802.11e EDCA

3

You dream about going up there... But that is a big mistake !
Sebastian

Sommaire

Ce chapitre concerne la présentation et l'évaluation d'un algorithme de contrôle d'admission que nous avons développé pour EDCA. Il existe plusieurs types d' algorithmes de contrôle d'admission pour EDCA (Gao 05) :

- des algorithmes basés sur des mesures, c'est à dire qui prennent en compte, pour leurs prises de décision d'admission d'un nouveau flux, des mesures continues réalisées sur l'état du réseau (par exemple du débit des flux actifs, des tailles des files et des taux de collisions...) ;
- des algorithmes basés sur des modèles, qui utilisent des métriques obtenues à partir de modèles analytiques du fonctionnement pour prendre des décisions.

Plus récemment est apparu un troisième type d'algorithmes qui combinent les propriétés des deux types précédents, c'est à dire qui intègrent une partie mesure et une partie modèle, et que l'on nomme algorithmes hybrides.
Notre proposition concerne un algorithme hybride. Son évaluation est faite par comparaison avec le travail de Pong et Moors (Pong 03) sur le contrôle d'admission. Pong et Moors ont proposé l'utilisation du modèle DCF de Bianchi (Bianchi 00), c'est à dire qu'ils se basent sur le débit nominal que peut atteindre une station DCF sur le réseau. Ce modèle s'adapte directement au cas d'une seule catégorie d'accès par station ; dans le cas où on a x catégories d'accès par station, la modélisation effectuée considère que l'on a x stations ayant chacune une seule catégorie d'accès ce qui représente une vue trop simplifiée et qui ignore la présence de collisions virtuelles par exemple.
L'utilisation d'un modèle au sein d'un algorithme de contrôle d'admission nécessite que le modèle soit aussi proche que possible de la réalité. L'utilisation du modèle de Bianchi ne suffit pas à représenter ni tous les mécanismes d'accès des différentes catégories d'accès d'une station, ni l'occupation effective du médium par les différentes catégories d'accès des différentes stations au sein d'un réseau.
L'objectif de l'algorithme que nous présentons, et qui est basé sur le modèle exhaustif du comportement de EDCA présenté dans le chapitre 2, est d'appréhender de manière plus précise la problématique du contrôle d'admission dans un contexte où une station contient plusieurs catégories d'accès.
Ce chapitre est structuré suivant quatre parties :

- la première partie concerne la présentation des mécanismes fondamentaux de l'algorithme,
- la deuxième partie est relative à l'analyse et l'évaluation de l'algorithme,
- la troisième partie concerne une proposition de modification visant à corriger une vue trop optimiste des mécanismes fondamentaux,
- la quatrième partie développe un cadre d'implémentation de l'algorithme.

3.1 Présentation de l'algorithme

Le fonctionnement général de l'algorithme de contrôle d'admission est le suivant : l'algorithme fonctionnera au niveau d'une entité centrale qui aura pour charge de protéger les flux actifs, cette entité peut être le point d'accès par exemple. À l'arrivée d'un nouveau flux, un calcul sera effectué afin de savoir si ce nouveau flux peut être admis ou non, l'algorithme se pose à ce niveau deux questions distinctes : est ce que le flux aura ce qu'il a demandé s'il est admis ? et est-ce qu'en l'admettant, la qualité perçue par les flux déjà admis ne sera pas dégradée ? Ceci est ce que nous appellerons par la suite le processus de décision. Ce processus sera formalisé. Afin de pouvoir répondre à ces questions, l'algorithme construit une vue de l'état du réseau (débit maximal de chaque AC active du réseau). Cette vue sera déduite, dans notre proposition, en couplant deux aspects :

1. Le modèle réduit que nous avons présenté précédemment est un premier élément. Ce modèle permettra à l'algorithme de calculer le débit maximal que peut atteindre une catégorie d'accès EDCA lorsqu'elle est placée dans un certain contexte réseau.
2. Le deuxième aspect est la connaissance qu'aura l'algorithme de contrôle d'admission de l'état du réseau. Le contexte réseau que le modèle utilise est représenté par des mesures faites de l'état du réseau ainsi que des estimations effectuées en se basant sur ces mesures. Les mesures concerneront la probabilité d'occupation du médium ainsi que des probabilités de collisions des files.

Ces deux aspects seront détaillés par la suite.

3.1.1 L'algorithme

Algorithme 1 : Contrôle d'admission utilisant le modèle réduit

```
1  for each Update_Period do
2      Update_Busy_Probabilities;
3      Update_Collision_Probabilities;
4      if New_Flows ≠ ∅ then
5          F_i = Get_New_Flow;
6          Update_Parameters(F_i);
7          Calculate_Achievable_Throughput( Admitted_Flows ∪ F_i);
8          if Check_Throughput (Admitted_Flows ∪ F_i) then
9              Admit(F_i);
10             Admitted_Flows = Admitted_Flows ∪ F_i;
11         else
12             Reject (F_i);
```

L'algorithme 1 présente le fonctionnement général du contrôle d'admission. L'idée est donc d'utiliser la métrique de débit du modèle réduit dans laquelle est injectée la condition du réseau à travers les différentes probabilités (p_b, $p_i^{(2)}$ et $p_i^{(3)}$).
À chaque période de mise à jour (*Update_Period*), qui peut par exemple être la durée d'une supertrame 802.11, le point d'accès doit mettre à jour sa vision de l'état du réseau :

- en recalculant la probabilité d'occupation du médium (la méthode *Update_Busy_Probabilities*),
- en prenant connaissance des différentes probabilités de collisions des catégories d'accès actives du réseau (la méthode *Update_Collision_Probabilities*).

Ceci peut être fait par *piggybacking* en utilisant un champs des paquets transmis ou en utilisant un protocole de signalisation simple. Si une requête pour un nouveau flux est arrivée entre temps, le point d'accès calculera le débit atteignable pour chacun des flux actifs ainsi que pour le nouveau flux (en utilisant la métrique *Achievable Throughput* dérivée du modèle). Si le débit théorique que peut atteindre un flux utilisant une catégorie d'accès (ou un ensemble de flux utilisant une catégorie d'accès) est supérieur à ce que demande le flux (ou l'ensemble des flux), le nouveau flux demandant l'accès sera admis, sinon il sera rejeté.
Nous définissons deux ensembles qui seront utilisés par l'algorithme : un ensemble répertoriant les flux actifs (*Admitted_Flows*) et un autre répertoriant les flux demandant l'accès (*New_Flows*). Il est essentiel, pour le bon fonctionnement du modèle, de garder cette liste des flux actifs, l'admission d'un nouveau flux dépend aussi de la protection accordée à ces flux. Dans le cas où la spécification d'un flux actif viendrait à changer (une modification du profil physique de la station émettrice par exemple), le flux est obligé de refaire une requête d'admission avant de pouvoir continuer à utiliser le réseau.

3.1.2 Mesures et estimations

Dans un premier lieu, nous avons conçu l'algorithme de contrôle d'admission pour utiliser directement les mesures fournies par les différentes stations ou réalisées par le point d'accès dans le calcul du débit maximal possible. Cependant, ceci n'est évidemment pas le meilleur moyen pour avoir une vision correcte du réseau. En effet, la décision que nous prenons doit être faite, sur la base d'informations aussi correctes que possible sur l'état du médium. Les mesures faites avant l'admission représentent l'état du médium avant l'admission. La décision à prendre doit être prise en prenant en compte l'état du réseau si le flux est accepté. La prise de mesure dans le futur n'étant pas possible, nous proposons un mécanisme simple d'estimation de l'état futur du réseau. Nous proposons donc d'estimer ce que pourraient être les probabilités d'occupation et de collision réelle si jamais le flux est accepté. Nous nous basons pour réaliser ces estimations :

- sur les mesures réalisées de l'état actuel du réseau,
- sur les formules analytiques de ces probabilités que nous avons détaillées dans le paragraphe 2.2.6.1.

Les estimations que nous faisons seront utilisées en place des mesures réelles dans le calcul du débit maximal atteignable servant le processus de décision.

3.1.2.1 Rappel

Soit F_i le flux demandant l'admission, F_i utilise la catégorie d'accès AC_i de la station s. Nous avons défini τ_i comme étant la probabilité que AC_i accède effectivement au médium pendant un *timeslot* donné. Nous appelons aussi Γ_s la probabilité que la station s accède effectivement au médium pendant un *timeslot* donné. Parmi les catégories d'accès d'une même station, une seule peut accéder au réseau à un *timeslot* donné (les autres sont soit inactives, soit en période de Backoff, soit elles ont perdu une collision virtuelle). Nous avons donc, pour une station donnée :

$$\Gamma_s = \sum_{i=0}^{3} \tau_i$$

Nous avions défini p_b comme étant la probabilité que le médium devienne occupé, si nous négligeons les autres raisons d'occupation du médium (c'est à dire, si nous considérons qu'un

médium n'est occupé que si une station l'occupe), nous pouvons écrire :

$$p_b = 1 - (1 - \Gamma_1)(1 - \Gamma_2) \dots (1 - \Gamma_M)$$

M étant le nombre de stations.
Nous définissons $p_{AC_i\_CR}$ comme la probabilité que AC_i accède au médium et qu'elle subisse une collision réelle. La différence avec p_i^r définie dans le paragraphe 2.2.6.1 est que p_i^r considère acquis l'accès de AC_i, dans les faits $p_{AC_i\_CR} = \tau_i p_i^r$. Sur la figure 2.2 du chapitre 2, $p_{AC_i\_CR}$ représente la probabilité de passer de l'état *"AC_i ne peut pas tenter l'accès"* à l'état *"AC_i subit une collision réelle"*. Nous avons donc :

$$p_{AC_i\_CR} = \tau_i(1 - (1 - \Gamma_1) \dots (1 - \Gamma_{s-1})(1 - \Gamma_{s+1}) \dots (1 - \Gamma_M))$$

3.1.2.2 Les estimations

Probabilité de collision réelle Nous souhaitons estimer l'effet qu'aura l'introduction de F_i dans le réseau sur la probabilité de collision réelle de AC_i. Nous supposons que seule l'activité d'accès de la station où évolue F_i sera affectée par l'introduction de F_i. Soit $p_{AC_i\_CR\_new}$ la probabilité estimée (si F_i est admis) d'accès de AC_i suivi d'une collision réelle, τ_{i_new} est l'estimation de la probabilité d'accès de AC_i si F_i est admis. $p_{AC_i\_CR\_old}$ et τ_{i_old} sont la probabilité d'accès suivi d'une collision réelle et la probabilité d'accès mesurées. Nous avons :

$$p_{AC_i\_CR\_new} - p_{AC_i\_CR\_old} = (\tau_{i\_new} - \tau_{i\_old})(1 - \prod_{j \neq s}(1 - \Gamma_j))$$

$$p_{AC_i\_CR\_new} = (\tau_{i\_new} - \tau_{i\_old})(1 - \prod_{j \neq s}(1 - \Gamma_j)) + p_{AC_i\_CR\_old}$$

Soit Δ_τ la différence introduite par l'éventuelle admission de F_i à la probabilité d'accès de la catégorie d'accès. Nous avons :

$$p_{AC_i\_CR\_new} = (\Delta_\tau)(1 - \prod_{j \neq s}(1 - \Gamma_j)) + p_{AC_i\_CR\_old}$$

L'estimation que nous faisons ici aura un effet direct sur la valeur de $p_i^{(2)}$.

Probabilité d'occupation De la même manière, nous définissons p_{b_new} comme étant l'estimation de la probabilité d'occupation du médium si F_i était accepté, p_{b_old} est la mesure faite de la probabilité d'occupation. Nous avons donc :

$$\begin{aligned} \frac{1 - p_{b\_new}}{1 - p_{b\_old}} &= \frac{(1 - \Gamma_1)(1 - \Gamma_2) \dots (1 - \Gamma_{i\_new}) \dots (1 - \Gamma_M)}{(1 - \Gamma_1)(1 - \Gamma_2) \dots (1 - \Gamma_{i\_old}) \dots (1 - \Gamma_M)} \\ &= \frac{(1 - \Gamma_{i\_new})}{(1 - \Gamma_{i\_old})} \end{aligned}$$

Nous adoptons le même raisonnement que pour l'estimation de p_{ir} :

$$\begin{aligned} \frac{1 - p_{b\_new}}{1 - p_{b\_old}} &= \frac{(1 - \Gamma_{i\_new})}{(1 - \Gamma_{i\_old})} \\ &= \frac{(1 - \Gamma_{i\_old} - \Delta_\tau)}{(1 - \Gamma_{i\_old})} \\ 1 - p_{b\_new} &= (1 - p_{b\_old}) \frac{(1 - \Gamma_i - \Delta_\tau)}{(1 - \Gamma_i)} \\ p_{b\_new} &= 1 - (1 - p_{b\_old}) \Big(1 - \frac{\Delta_\tau}{(1 - \Gamma_i)}\Big) \end{aligned}$$

Variation de l'accès La seule inconnue pour les deux estimations est Δ_τ. Δ_τ représente Le taux d'accès additionnel causé par l'introduction du nouveau flux à une catégorie d'accès, ces accès peuvent être des tentatives de transmission ou de retransmissions. Nous utilisons ce qui suit pour l'estimer :

$$\Delta_\tau = (1 - p_{AC_i\_CR\_old})\delta$$

δ étant le nombre d'accès introduits par le flux (c'est à dire le nombre de paquets qui devront être envoyés pendant une période *Update_Period*). Nous ne considérons dans cette formule qu'une seule possibilité de collision réelle par paquet à envoyer.

3.2 Analyse et évaluation

Nous présentons dans cette section une analyse par simulation de l'algorithme de contrôle d'admission. Nous avons implémenté notre algorithme ainsi que celui de Pong et Moors sous le simulateur à événements discrets ns-2 (ns2 05). Nous nous sommes basés pour ce faire sur le module de EDCA présenté par le groupe *Fachgebiet Telekommunikationsnetze* de la *Technischen Universität de Berlin* (TKN-TUB 07). Dans les simulations que nous présentons ici, le canal est considéré idéal, sans terminaux cachés, un nœud fait office de point d'accès où s'exécute le contrôle d'admission, les autres nœuds sont des stations utilisant le schéma d'accès EDCA pouvant jouer le rôle de sources ou de puits de trafic. Les stations fonctionnent à un débit physique de 11 Mbps.

Deux aspects de l'algorithme sont vérifiés, la qualité des décisions prises et sa performance vis à vis de l'algorithme de Pong et Moors. Nous avons choisi de comparer notre algorithme à celui de Pong et Moors car ils fonctionnent tous les deux sur les mêmes principes généraux. L'algorithme de Pong et Moors est celui qui fait référence dans le domaine des algorithmes hybrides de contrôle d'admission pour EDCA. Notre algorithme diffère de celui de Pong et Moors en deux points essentiels :

- Notre algorithme utilise un modèle d'EDCA, beaucoup plus élaboré que l'adaptation du modèle de Bianchi qu'utilise l'algorithme de Pong et Moors. Notre modèle intègre la représentation de l'occupation à proprement parler du médium, chose qu'ignore le modèle utilisé par Pong et Moors.
- Notre algorithme se base sur les mesures réelles pour faire une estimation de l'état futur du médium (voir le paragraphe 3.1.2.2). L'algorithme de Pong et Moors utilise les mesures directement.

3.2.1 Une catégorie d'accès par station

Des simulations ont été entreprises afin de vérifier la qualité des décisions d'admission prise par notre algorithme. Nous faisons cette analyse en comparant le débit atteint par chaque flux actif sur le réseau avec ou sans contrôle d'admission. Une vision du comportement de l'algorithme et du processus de décision est donnée par le graphe du débit atteignable.

3.2.1.1 Scénarios de Simulation

Le réseau est constitué de 8 stations, une d'elles fait office de point d'accès (où l'algorithme de contrôle d'admission fonctionne), les autres sont des QSTA. Au niveau de chaque QSTA, un flux CBR destiné au point d'accès est activé, l'activation des flux se fait périodiquement. Tous les flux ont les mêmes caractéristiques. Lors de l'activation d'un flux, la QSTA envoie une

Scénario	Taille de paquet (Octets)	Inter-arrivées(s)	Débit moyen du flux (Mbps)
Scénario 1	600	0.003	1.6
Scénario 2	800	0.004	1.6

TAB. 3.1 – *Spécification des scénarios*

demande d'admission avec la spécification du nouveau flux et attend une réponse de la part du contrôle d'admission avant d'accorder (ou de refuser) au nouveau flux l'accès au réseau. Plusieurs simulations ont été effectuées avec différentes spécifications de flux. Nous n'en présentons qu'un certain nombre, les conclusions des autres se rapportant aux conclusions de celles-là. Les figures 3.1 et 3.2 présentent les résultats des scénarios de simulation 1 et 2 respectivement exécutés avec les spécifications de trafic présentées dans la table 3.1.

Nous avons choisi de donner des résultats correspondant à des flux ayant la même requête de bande passante mais avec des tailles de paquet différentes afin d'analyser l'effet de cet aspect sur le fonctionnement du contrôle d'admission. Dans chaque graphe de résultats, différentes mesures de débits effectifs ou calculés sont tracées en fonction du temps de simulation, nous regroupons les résultats en trois groupes. Le premier groupe contient le graphe du débit atteignable calculé pour une des catégories d'accès par notre algorithme (toutes les catégories d'accès actives se trouvent dans la même situation, elles auront donc le même comportement à ce niveau), ce graphe est étiqueté 1 : il représente l'estimation que fait l'algorithme de contrôle d'admission, en se basant sur le modèle et les mesures, du débit que peut atteindre une catégorie d'accès dans les conditions du réseau si elle était saturée. Le second groupe, destiné à analyser la protection assurée par le contrôle d'admission aux flux actifs, est composé de deux graphes : le débit moyen de tous les flux actifs (c'est à dire le débit total divisé par le nombre de flux actifs) sans contrôle d'admission (étiqueté 2a) et avec l'algorithme que nous proposons (étiqueté 2b). Nous représentons aussi, dans un troisième groupe de graphes, l'utilisation du réseau (débit total) atteint avec l'algorithme de Pong et Moors (graphe étiqueté 3c) ou avec notre algorithme (graphe étiqueté 3d), ces derniers graphes nous permettront de comparer la performance de notre algorithme à celle de l'algorithme de Pong et Moors. L'arrivée d'un nouveau flux ainsi que la décision prise par l'algorithme que nous proposons sont indiquées par des flèches sous chaque graphe.

3.2.1.2 Résultats et analyse

Nous pouvons voir que notre algorithme intervient de façon à protéger les flux actifs des flux arrivant (en comparant les graphes 2a et 2b de chaque figure : les nouveaux flux sont rejetés lorsque l'estimation du débit atteignable devient inférieure à la demande du flux). Si nous comparons les résultats des simulations ayant la même bande passante mais avec une taille de paquet différente (comparer la figure 3.1 à la figure 3.2 par exemple), nous notons que le calcul du débit atteignable pour les flux avec des paquets de taille inférieure (donc un temps d'inter-arrivées inférieur) est moins optimiste. Ceci est dû essentiellement au fait que : des temps d'inter-arrivées plus petits signifient un nombre d'accès supérieur et donc un nombre de collisions supérieur, ceci rend l'estimation du débit inférieure. L'autre raison est la relation directe qui existe entre la taille du paquet et le débit calculé : la catégorie d'accès étant considérée saturée pour le calcul, ce qui importe pour l'estimation du débit est le nombre d'accès avec succès possibles, donc plus la taille du paquet envoyé au cours d'un accès est grande et plus l'estimation du débit au final est grande.

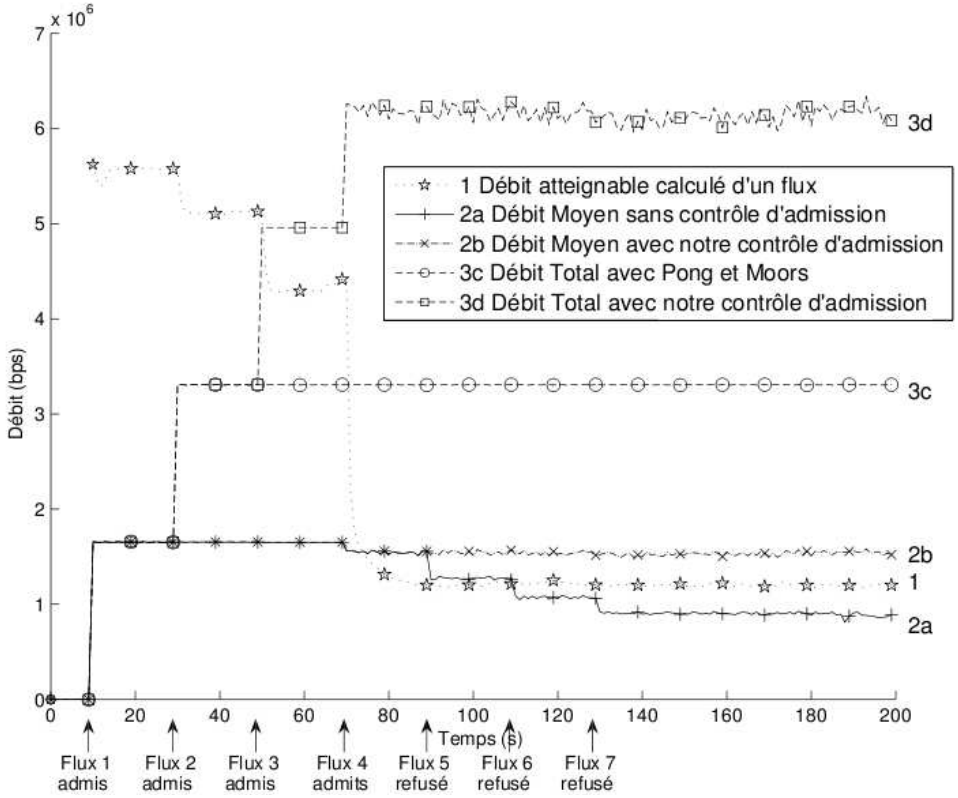

FIG. 3.1 – *Résultats du scénario 1*

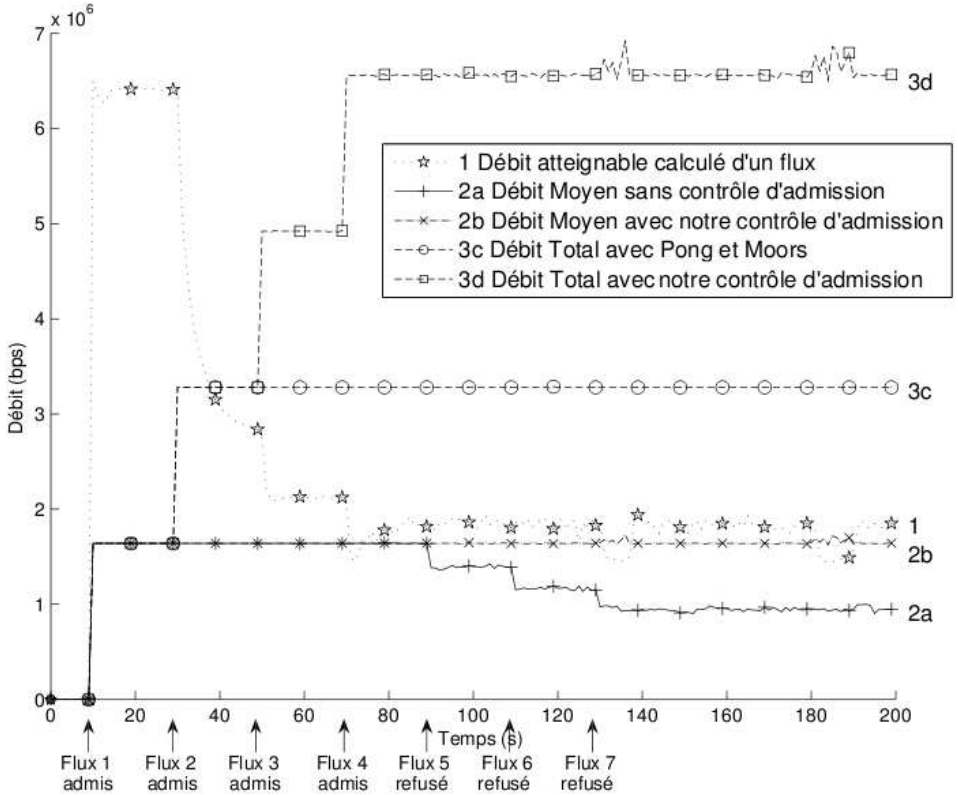

FIG. 3.2 – *Résultats du scénario 2*

AC	Taille de paquet (Octets)	Inter-arrivées(s)	Débit moyen du flux (Mbps)
AC_VO	400	0.032	100
AC_VI	700	0.0224	250

TAB. 3.2 – *Spécifications des flux*

3.2.1.3 Analyse de la comparaison à l'algorithme de Pong et Moors

Nous pouvons clairement voir que notre algorithme parvient à assurer une meilleure utilisation du réseau que celle assurée par l'algorithme de Pong et Moors. Cette meilleure admissibilité est acquise tout en assurant une protection des flux déjà actifs. L'algorithme de Pong et Moors a un comportement pessimiste : lors du calcul qu'il effectue du débit que peut atteindre une AC, il considère que toutes les autres AC actives sur le réseau sont saturées, l'algorithme que nous proposons n'applique la condition de saturation qu'à l'AC pour laquelle le débit atteignable est calculé. Ceci est essentiellement dû à la prise en compte, dans notre modèle (et par conséquent dans l'algorithme), de la nouvelle métrique d'occupation du médium p_b qui provient de la troisième dimension présente dans le modèle (cf. le paragraphe 2.2), cette métrique nous permet d'avoir une meilleure vision de l'état du réseau et de l'intégrer dans le processus de décision.

3.2.2 Plusieurs catégories d'accès par station

3.2.2.1 Scénarios de simulation

Plusieurs scénarios de simulations avec des spécifications de trafic différentes ont été conduites afin d'observer les performances de notre algorithme dans les cas où plusieurs catégories d'accès seraient actives à la fois au sein d'une même station. Nous utilisons la même configuration de réseau utilisée précédemment, à la seule différence que le débit physique utilisé est maintenant de 2 Mbps. Ceci a été fait afin de réduire les durées de simulation, les phénomènes étant pratiquement les mêmes avec des débits physiques différents. Deux flux par station seront activés, séparément et périodiquement. Le premier utilise la catégorie d'accès Voix (AC_VO), le deuxième la catégorie d'accès Vidéo (AC_VI). Les flux sont décrits dans la table 3.2.

3.2.2.2 Résultats et analyse

Comme pour les premiers scénarios, l'algorithme que nous proposons permet d'atteindre de meilleures utilisations du médium tout en assurant un degré de protection des flux actifs. L'algorithme de Pong et Moors est encore une fois trop pessimiste, un seul flux voix a été admis (dans la figure 3.3, les graphes 1b et 2c sont superposés), alors qu'il était possible d'en admettre plus sans pour autant dégrader la qualité perçue par les flux.

Nous avons cependant remarqué que dans **certains** cas de figure (ceci demeure rare), notre algorithme admettait un flux de trop (sur la figure 3.4, l'admission du quatrième flux VO dégrade le débit des flux VI). L'algorithme de Pong et Moors certes n'admet pas ce flux de trop, mais il n'a pas admis le troisième flux VO non plus. Nous analysons et proposons des solutions à ce problème dans le paragraphe suivant.

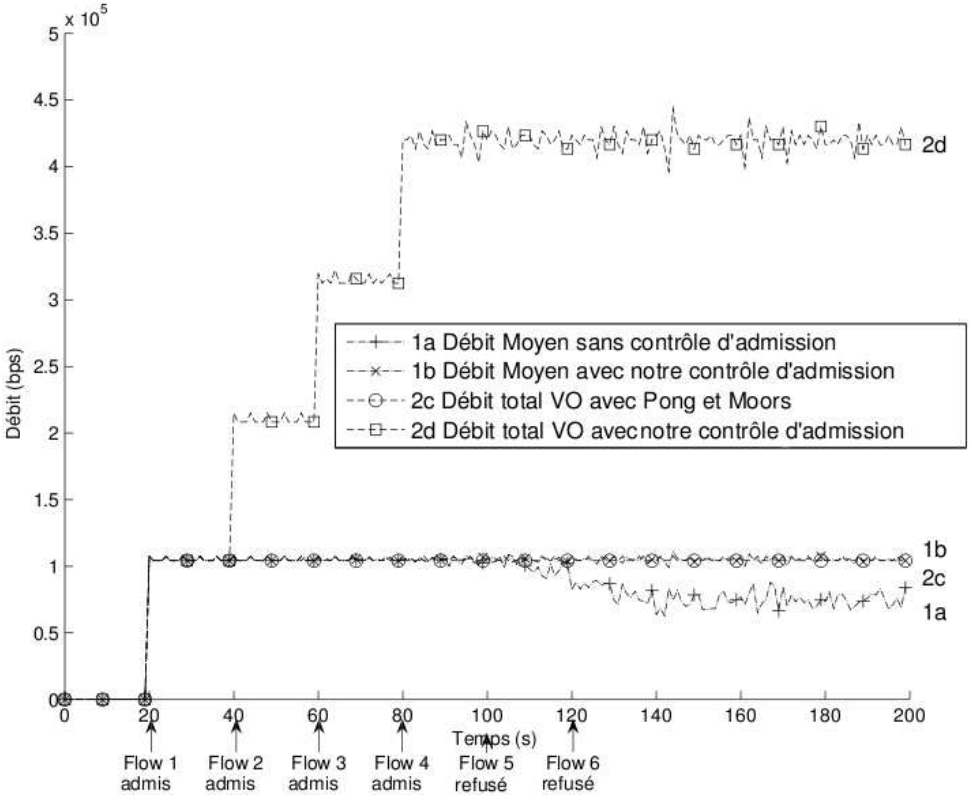

FIG. 3.3 – *Les flux utilisant AC_VO*

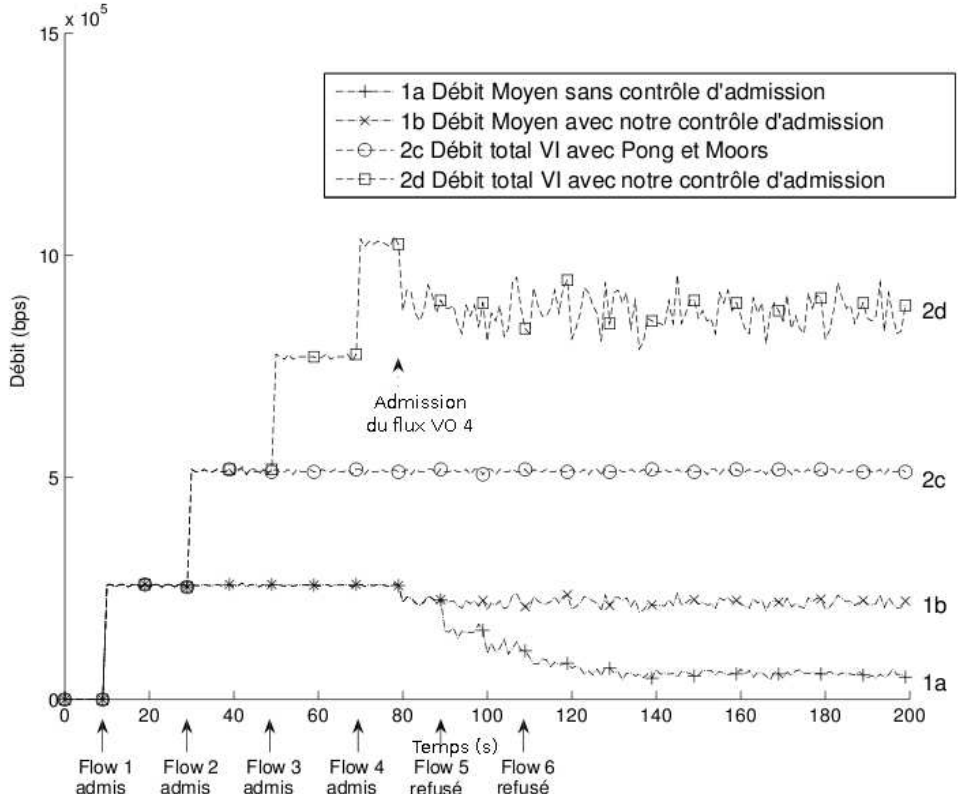

FIG. 3.4 – *Les flux utilisant AC_VI*

3.3 Amélioration de l'algorithme

Les estimations, même si elles vont dans le bon sens de l'évolution des probabilités estimées, n'atteignent pas un résultat exacte. Ceci s'explique par le fait que notre technique d'estimation ne considère l'effet de l'admission d'un flux que sur la probabilité de collision de son AC. Il est cependant clair que l'introduction du nouveau flux a un effet sur toutes les catégories d'accès du médium. Nous proposons donc deux moyens pour résoudre ce problème :

- Intégrer dans le processus d'estimation ou dans le processus de décision, des informations supplémentaires de l'état du réseau,
- corriger les estimations du processus d'estimation à l'aide d'un système de retour (*feedback*)

L'intuition derrière ces modifications est la suivante : nous souhaitons que le contrôle d'admission devienne plus réticent lorsque le réseau devient chargé. Les modifications que nous proposons dans la suite sont indépendantes et seront donc à appliquer de façon indépendante.

3.3.1 Information supplémentaire sur l'état du médium

3.3.1.1 Ajoutée au processus d'estimation

Lorsqu'une estimation de l'effet de l'introduction du nouveau flux F_i sur la probabilité d'occupation du médium p_b ou d'une probabilité de collision p_i est faite, nous considérions que l'effet sera en corrélation avec le nombre d'accès au médium qu'aura à faire ce nouveau flux. Il est cependant clair que l'effet sur le nombre de collisions ne sera que plus grand si le médium est déjà chargé avant l'arrivée de ce nouveau flux. Nous introduisons donc un facteur que nous appelons le facteur de réticence qui sera en corrélation avec l'état du médium : plus le médium est occupé et plus l'effet de l'introduction d'un nouveau flux sera grand. Ceci est fait en multipliant la valeur de l'estimation de p_b par un facteur $alpha$ de réticence calculé de la façon suivante :

$$\alpha = 1 + 0.2 * (1 - \frac{totalBandwidth - totalRequests}{totalBandwidth})$$

$totalBandwidth$ étant le débit utile théorique du réseau, $totalRequests$ est la somme de ce que demandent tous les flux actifs ou en attente d'admission. α sera plus grand si le médium est déjà chargé.

3.3.1.2 Ajoutée au processus de décision

Nous proposons de remplacer la simple comparaison faite entre le débit atteignable et le débit demandé (comparaison à partir de laquelle la décision est prise) par : $\alpha Calculated\_Achievable\_Throughput(F) > Requested\_Bandwidth(F)$ avec $0 \leq \alpha \leq 1$ calculé de la façon suivante : $\alpha = (1 - \beta) + \beta * \frac{totalBandwidth - totalRequests}{totalBandwidth}$ β spécifie le degré de réticence, il a été réglé à 0.2 pour les simulations. Plus la demande totale de bande passante est grande, et plus α sera grand, et plus il y aura de la réticence au niveau du processus de décision. Avec β aussi petit que 0.2, la décision sera tout de même basée essentiellement sur les calculs du modèle.

Scénario	Taille de paquet (Octets)	Inter-arrivées(s)	Débit moyen du flux (Mbps)
Scénario 1	600	0.002	2.4
Scénario 2	800	0.004	1.6
Scénario 3	600	0.004	1.2

TAB. 3.3 – *Spécifications des scénarios*

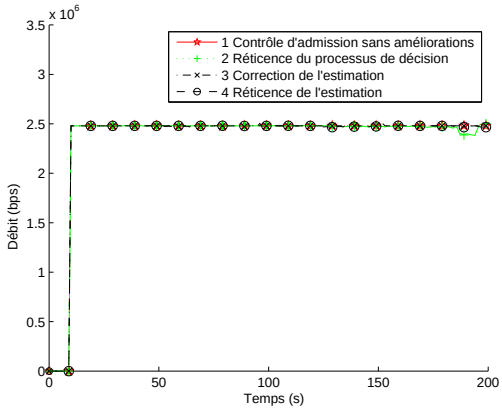

FIG. 3.5 – *Débit moyen des flux , scénario 1*

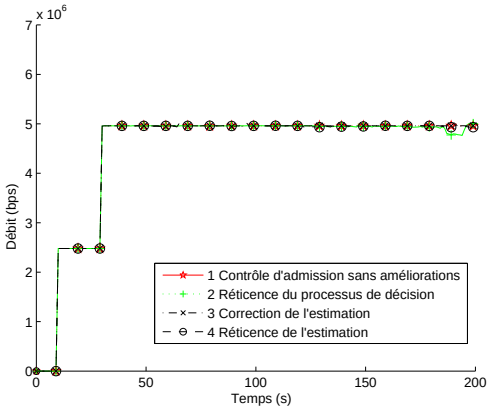

FIG. 3.6 – *Débit total des flux , scénario 1*

3.3.2 Correction d'estimation

Vu que le processus d'estimation ne parvient pas à assurer une estimation exacte (surtout en situation de charge), nous introduisons un simple processus de retour sans mémoire afin d'assurer la correction de l'estimation de la probabilité d'occupation du médium : nous rajoutons donc à chaque nouvelle estimation, la différence qu'il y a eu entre la dernière mesure et l'estimation qui lui correspondait. Soit p_{be_k} la $k^{ième}$ valeur estimée de p_b (utilisant le processus d'estimation proposé), p_{bm_k} la $k^{ième}$ valeur mesurée de p_b, $p_{b\_new\_k}$ la nouvelle estimation corrigée de p_b. La modification fonctionnera donc comme suit :

$$p_{b\_new\_k} = p_{be\_k} + (p_{bm\_k-1} - p_{be\_k-1})$$

3.3.3 Analyse des modifications

Le même procédé de simulation utilisé précédemment est repris ici. Les métriques utilisées pour comparer le fonctionnement de chacune des modifications et de l'algorithme sans modifications sont :

- Le débit total atteint pour un scénario donné par l'ensemble des flux en utilisant l'algorithme de contrôle d'admission avec ou sans les améliorations.
- Le débit moyen par flux atteint pour un scénario donné en utilisant l'algorithme de contrôle d'admission avec ou sans les améliorations.
- La fonction de distribution cumulative des délais de tous les paquets de données.

Plusieurs scénarios de simulation ont été testés, nous présentons dans ce qui suit quelques résultats significatifs. Dans les scénarios 1,2 et 3 (tableau 3.3) : le canal est considéré parfait, sans terminaux cachés, le débit physique est de 11 Mbps. Dans les stations, un nouveau flux est activé de façon périodique, demandant l'accès au réseau et attendant le résultat de la demande. Les résultats de ces simulations sont présentés dans les figures 3.5-3.13.

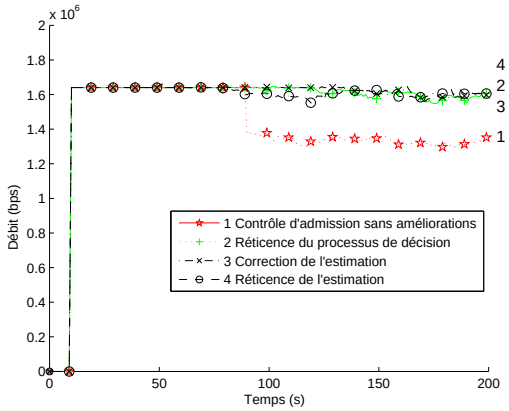

FIG. 3.7 – *Débit moyen des flux , scénario 3*

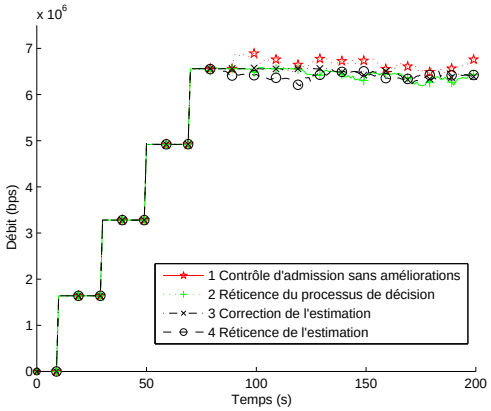

FIG. 3.8 – *Débit total des flux , scénario 3*

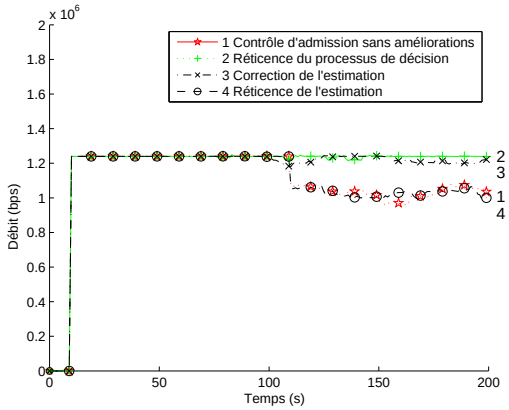

FIG. 3.9 – *Débit moyen des flux , scénario 3*

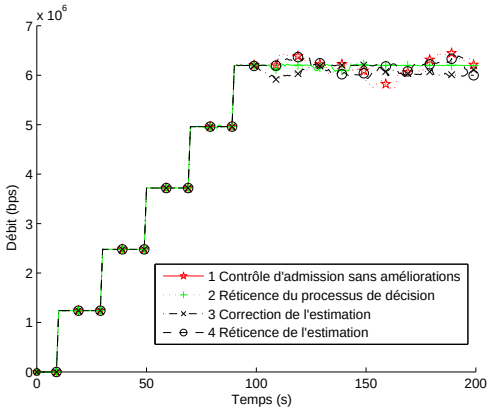

FIG. 3.10 – *Débit total des flux , scénario 3*

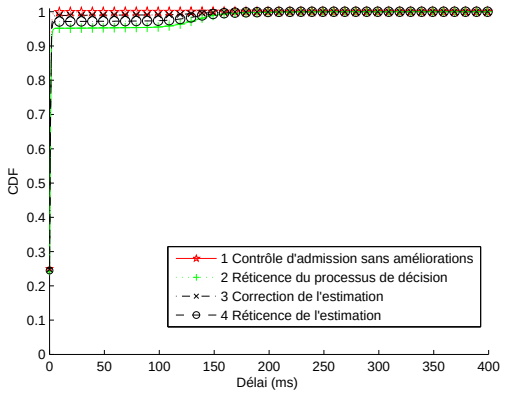

FIG. 3.11 – *FDC des délais, scénario 1*

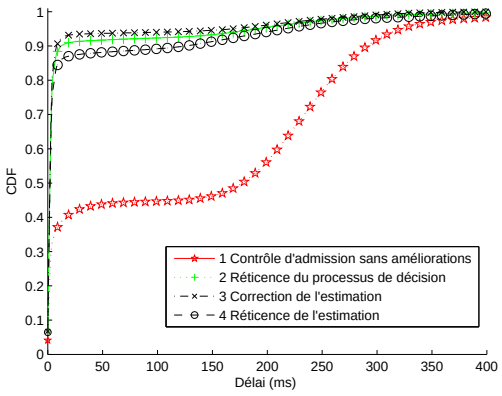

FIG. 3.12 – *FDC des délais, scénario 2*

3.3.3.1 Analyse

Nous pouvons clairement remarquer que deux des trois propositions assurent une correction du problème rencontré.

Scenario1 Les résultats sont présentés dans les figures 3.5, 3.6 et 3.11. Le scénario 1 est un cas où aucune mauvaise décision n'est prise par le contrôle d'admission. Nous pouvons voir que les modifications apportées n'ont pas eu d'effet négatif. L'algorithme ainsi que les

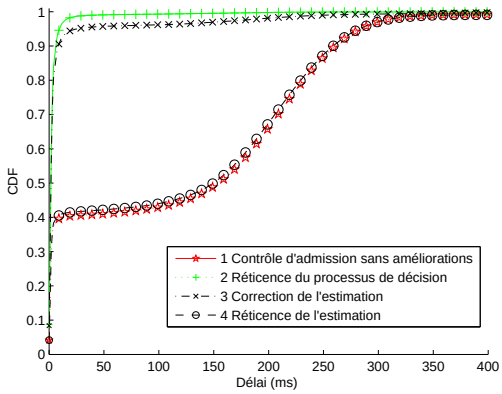

FIG. 3.13 – *FDC des délais, scénario 3*

modifications ont admis le bon nombre de flux : maximisant l'utilisation du médium sans avoir d'effet pervers sur la qualité perçue par les flux (fig. 3.5-3.6). Comme nous l'avons dit, le but des modifications apportées est de rendre les décisions d'admission plus réticentes afin d'éviter les mauvaises décisions. Ici aucune mauvaise décision n'est prise, la moyenne de débit offert à chaque flux respecte la demande du flux et les délais sont minimaux (fig. 3.11).

Scénario 2 Les résultats sont présentés dans les figures 3.7, 3.8, 3.12. Dans le scénario 2, l'algorithme non modifié admet un flux de trop. Les modifications corrigent ce problème. Il en résulte un meilleur débit moyen par flux (figure 3.7) (mieux dans le sens où il respecte la demande du flux) et une meilleure distribution des délais (figure 3.12) (avec les modifications, nous avons suivant le cas, 90 % des paquets avec un délai inférieur à 10 ms alors qu'avec l'algorithme non modifié ce n'est le cas que pour 37 % des paquets). Le flux de trop accepté cause un nombre de collisions au delà de ce qu'avait estimé l'algorithme, ce qui cause une dégradation du service fourni à l'ensemble des flux.

Scénario 3 Les résultats sont présentés dans les figures 3.9, 3.10, 3.13. La même analyse peut être faite dans ce cas avec comme principale différence que la première modification ne résout pas le problème. Ceci est principalement le cas lorsque des flux demandant une petite bande passante arrivent dans un contexte chargé : la modification apportée au processus d'estimation ne réussira pas à corriger le tir. Cette modification fonctionne de façon proportionnelle à la requête du nouveau flux, elle aura un effet moindre lorsque le flux a une demande moindre.

3.3.4 Conclusions

Parmi les trois modifications que nous proposons, deux permettent de corriger le problème observé avec l'algorithme seul sans en dégrader la performance. Nous préconisons l'utilisation de la modification **Correction d'estimation**. L'analyse par simulation a montré que cette modification permet de corriger le problème de sur-admission. Elle est, de plus, simple d'implémentation et plus intuitive. Suite à l'analyse par voie de simulation de l'algorithme, nous allons maintenant présenter une implémentation de l'algorithme que nous avons réalisée sur une plateforme d'expérimentation.

3.4 Implémentation de l'algorithme

Nous avons implémenté l'algorithme de contrôle d'admission que nous proposons dans le cadre de l'architecture du projet européen EuQoS (EuQoS 06, Racaru 08). Cette implémentation fut réalisée sur la plateforme d'expérimentation réseau LaasNetExp mise en place au sein du LAAS-CNRS (Owezarski 08). Nous présentons dans ce paragraphe le contexte de cette implémentation ainsi que les différentes étapes de la mise en place.

3.4.1 Contexte : l'architecture EuQoS

EuQoS (*End-to-end Quality of Service support over heterogeneous networks*) est un projet européen IST (*Information Society Technology*) dont le but est la recherche, l'intégration, le test et la validation de mécanismes de qualité de service de bout en bout. Le projet se place dans le contexte :

- des nouveaux besoins de qualité de service pour les applications telles que la VoIP (*Voice over IP*) la visioconférence ou la télémédecine ;
- de la structure actuelle et future de l'Internet essentiellement multi-domaines, multi-technologies et multi-services.

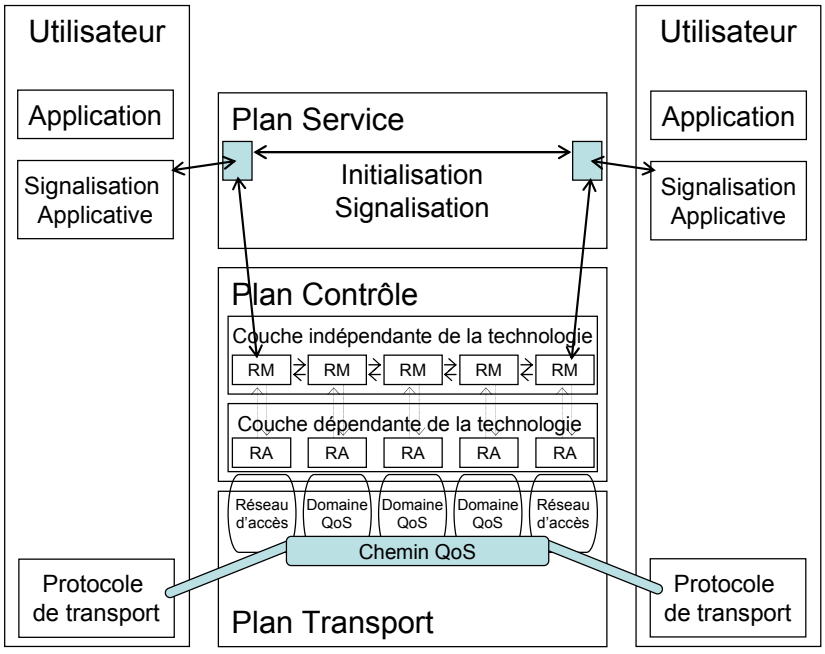

Fig. 3.14 – *Architecture EuQoS*

L'architecture de la solution EuQoS propose de gérer des chemins orientés QoS de bout en bout au travers de domaines distincts avec réseaux hétérogènes. L'architecture est découpée en trois plans horizontaux :

- un plan service permettant d'établir la session et d'initier la réservation du chemin avec la qualité de service demandée, il permet aussi de réaliser les vérifications sur les droits des utilisateurs.
- un plan de contrôle permettant la traduction des requêtes applicatives et la mise en œuvre des mécanismes nécessaires pour y répondre. Le plan contrôle est composé de

deux niveaux : un premier niveau, indépendant de la technologie, gère la convergence, il contient le RM (le *Resource Manager* qui a pour rôle de gérer les ressources sur le chemin de bout en bout) et le PCE (le *Path Computation Element* qui sert au calcul de chemins). Le deuxième niveau, qui lui dépend de la technologie sous-jacente, gère le contrôle d'admission et la configuration du réseau (réalisé par le RA-*Ressource Allocator*). C'est d'ailleurs au niveau du RA de WiFi que nous intervenons.

– un plan transport qui construit et gère les chemins orientés QoS établis entre les utilisateurs, à travers les différentes technologies sous-jacentes traversées.

3.4.2 Contexte : les composants matériels

Nous exposerons ici les différents composants matériels sur lesquels le contrôle d'admission sera *in fine* implémenté.

3.4.2.1 LaasNetExp

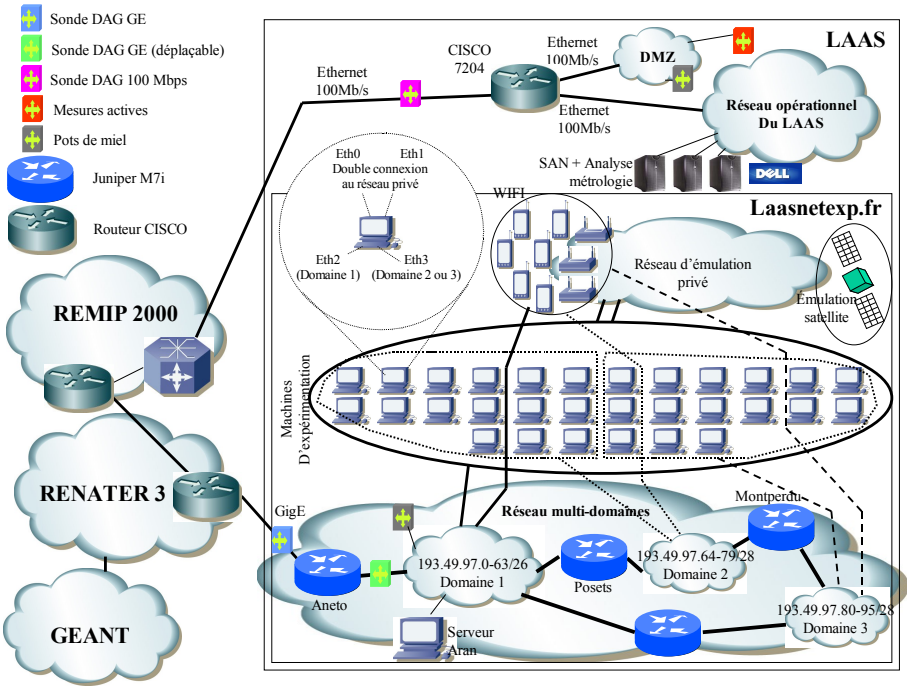

FIG. 3.15 – *Architecture de LaasNetExp*

Présentons en premier lieu la plateforme d'expérimentation LaasNetExp au niveau de laquelle est mise en place l'architecture EuQoS et qui nous servira de base. LaasNetExp est une plateforme conçue et installée au LAAS-CNRS pour l'expérimentation réseau en émulation et en environnement réel. La figure 3.15 détaille l'architecture de la plateforme. La conception de la plateforme s'est basée sur des choix à l'origine donnant à la plateforme les caractéristiques suivantes, que ce soit pour l'expérimentation en environnement réel ou en émulation :

– la reproductibilité des expérimentations effectuées ;
– la possibilité d'effectuer des mesures et de l'observation en temps réel ;
– la possibilité d'interconnecter LaasNetExp avec d'autres plateformes ;
– une expérimentation peut se faire de manière complètement isolée des autres expérimentations en cours et du trafic externe (par filtrage) ;

– elle donne la possibilité d'émuler des technologies réseaux spécifiques telles que les interconnexions avec des réseaux satellites.

La partie WiFi de la plateforme est composée de points d'accès CISCO AIR-AP1131AG-E-K9 (Cisco 09a) ainsi que d'ordinateurs portables équipés de cartes CISCO AIRONET AIR-CB21AG-E-K9 (Cisco 09b). Le point d'accès ainsi que les cartes réseaux sont compatibles WMM (*WiFi Multi Media*). WMM est un programme de certification établi par l'alliance WiFi permettant de certifier le respect du matériel à certains aspects de IEEE 802.11e (essentiellement la mise en place de la différentiation par EDCA).

3.4.3 Intégration du contrôle d'admission

L'implémentation que nous faisons de notre contrôle d'admission dans le cadre que nous venons de décrire est basée essentiellement sur trois classes effectuant l'ensemble des actions nécessaires. Il y a en premier la classe *WiFiCACAlgorithm* qui implémente le fonctionnement, à proprement parler, du contrôle d'admission dans le domaine WiFi ; cette classe est couplée à la classe *WiFiCalculator* réalisant l'ensemble des calculs nécessaires au bon fonctionnement du contrôle d'admission. La classe *WiFiCACAlgorithm* fait partie du système EuQoS. Une deuxième classe, *ACFlow*, permet la description que se fait le contrôle d'admission d'un flux unidirectionnel traversant un domaine WiFi. Une troisième classe essentielle dans notre cadre est la classe *NetworkSeeker*, cette classe permet au contrôle d'admission de récupérer des composants réseau les mesures nécessaires à son fonctionnement de façon périodique. En effet, dans notre contexte, le contrôle d'admission ne fonctionne pas au niveau du point d'accès WiFi mais au niveau d'une machine séparée implémentant le RA WiFi. Il est donc nécessaire de mettre en place l'échange d'information entre le RA et le point d'accès moyennant cette classe *NetworkSeeker*.

3.4.3.1 La classe *WiFiCACAlgorithm*

Cette classe fait partie de l'architecture EuQoS. Elle étend la classe *CACAlgorithm* qui définit la signature des méthodes que doit implémenter un contrôle d'admission. L'instance de cette classe fonctionnant au niveau du RA de chaque domaine WiFi est invoquée afin d'établir le contrôle d'admission de chaque nouveau flux accédant par le réseau concerné. À l'invocation de la méthode *checkResources* de l'instance, un objet *ACFlow* est créé traduisant la spécification du flux (TSPEC) et l'algorithme de contrôle d'admission que nous avons présenté est réalisé, donnant une réponse d'admission ou de refus du flux.

3.4.3.2 La classe *ACFlow*

Cette classe, comme nous l'avons vu, sert à représenter un flux du point de vue du contrôle d'admission. Elle contient donc, en plus de la spécification du trafic que doit porter ce flux, toutes les statistiques le concernant. Les statistiques vont du taux de collisions virtuelles, au taux de collisions réelles au taux d'accès. Cette classe implémente des méthodes permettant de mettre à jour ces statistiques et de réaliser certains calculs basés sur ces statistiques.

3.4.3.3 La classe *NetworkSeeker*

Cette classe permet d'invoquer le point d'accès de façon périodique afin de récupérer les statistiques nécessaires. Cette classe se connecte, par *Telnet*, au point d'accès et invoque les directives nécessaires à la lecture des statistiques. Ces statistiques sont lues sous la forme de

chaînes de caractères contenant l'ensemble des informations. Ces chaînes de caractères sont ensuite analysées, les informations nécessaires sont récupérées et placées dans les variables et les classes concernées.

3.4.3.4 Adaptation du contrôle d'admission au matériel

Nous exposons dans ce paragraphe les adaptations que nous avons eu à apporter au contrôle d'admission afin de le mettre en place dans le contexte matériel décrit. Ces adaptations sont surtout dues à l'impossibilité d'apporter des modifications aux pilotes des cartes WMM et du point d'accès. Nous avons donc eu à nous contenter des informations que le point d'accès nous fournit de façon native. La première des adaptations est due au manque de signalisation dans le réseau d'accès. Il a fallu donc se passer des statistiques de chaque flux sur la voie montante et de les remplacer par les statistiques sur la voix descendante. Ainsi, une AC d'une station était assimilée à une AC du point d'accès et les statistiques d'accès et de collisions utilisées étaient celles-là. L'autre adaptation est celle concernant la statistiques d'occupation. Le calcul de cette statistique s'est basé sur deux statistiques que fournissait le point d'accès CISCO : *txload* et *rxload*. Ces deux statistiques indiquent le degré d'occupation de la transmission et de la réception de l'interface radio. Il a donc fallu calculer à partir de ces deux variables le taux d'occupation du médium.

3.4.4 Bilan

L'implémentation du contrôle d'admission dans le contexte d'EuQoS a été réalisée et une première phase de test s'est déroulée avec succès. Il restera donc à intégrer l'implémentation sur la plateforme d'expérimentation et à réaliser les expérimentations dans un cadre réaliste et en évaluer les résultats.

Conclusion

L'algorithme de contrôle d'admission que nous proposons fait partie de la catégorie des algorithmes hybrides. Il utilise dans son processus de décision un modèle analytique du schéma d'accès EDCA (modèle que nous avons présenté dans le chapitre 2) ainsi que des mesures de l'état du réseau. À partir des mesures, notre algorithme met en place une estimation de ce que serait l'état du réseau dans l'éventualité d'admission du flux demandant l'accès. Un processus de retour *feedback* permet de corriger les estimations et d'améliorer ainsi la vue que possède notre algorithme de l'état du réseau. Les analyses ont montré que, comparée à des algorithmes de contrôle d'admission existants dans la littérature, notre proposition offre une meilleure admissibilité. L'ajout du processus d'estimation ainsi que la meilleure représentation que donne notre modèle du schéma d'accès nous permettent en général d'avoir une évaluation moins pessimiste et plus proche des possibilités du réseau.
Nous avons souhaité évaluer notre algorithme sur une plateforme d'expérimentation. L'implementation est terminée, il ne nous reste qu'à l'intégrer à l'architecture EuQoS et d'en vérifier le bon fonctionnement. Le contrôle d'admission que nous avons développé pour WiFi fut une partie importante du travail que nous avons réalisé. Il établit la gestion de la bande passante dans un domaine WiFi afin d'y améliorer la qualité de service.

4 Gestion agrégée de la bande passante pour IEEE 802.16 WiMAX

Dr. House : I was curious.
Since I'm not a cat,
that's not dangerous.
"House" saison 3, épisode 14

Sommaire

Ce chapitre se situe dans le contexte de la mise en œuvre des différentes classes de service de WiMAX : UGS, rtPS, nrtPS et BE.

L'objectif est de proposer une modalité de gestion de la bande passante sur la voie montante (*Uplink*) qui :

- a les propriétés de simplicité et de facilité de passage à l'échelle ;
- offre des garanties de Qualité de Service (QoS) exigées par les flux UGS et rtPS (flux contraints en termes de délai et de débit, que nous appelons aussi flux à contraintes temporelles) et aux flux nrtPS (flux contraints uniquement en terme de débit) ;

– protège les flux BE de la famine.

Nous ne nous intéresserons pas à la voie descendante, le problème se ramène à un problème de fourniture de QoS par un multiplexeur. Ce type de problème a été traité dans la littérature. Les propriétés de simplicité et de facilité de passage à l'échelle s'obtiennent sur la base de la notion d'agrégation qui s'exprime pendant la durée de vie des connexions à deux niveaux :

– au niveau des requêtes de bande passante qui ne sont plus par connexion mais par SS (notion de requête agrégée) ;
– au niveau de l'allocation de bande passante par la BS qui ne se fait pas par connexion mais par SS également.

Ce travail s'insère bien, en particulier, dans le cadre de l'utilisation de WiMAX en tant que lien *backhaul*[1] entre des réseaux locaux sans fil de type WiFi et le cœur du réseau Internet (voir figure 4.1). Cette utilisation est suggérée par certains documents du forum WiMAX (Forum 04). La SS peut alors faire office de point d'accès WiFi et les problèmes de mise à l'échelle doivent dans ce cas être présents à l'esprit de l'administrateur du réseau. En effet, une BS aura à gérer un nombre important de connexions différentes, ce qui fait que la complexité d'un système de gestion par flux applicatif est un des problèmes majeurs qui entravent la mise à l'échelle. Dans ce cadre, il paraît intéressant de simplifier le système de requêtes ainsi que le fonctionnement de la BS et de donner plus d'importance aux SS. En particulier, lorsque le nombre de connexions actives simultanément augmente, la SS pourra, avec l'allocation globale obtenue par la BS, servir correctement ses connexions.

Ce chapitre est structuré suivant trois parties :

– la première partie présente les différents aspects de notre proposition ;
– la deuxième partie concerne la modélisation ; plus précisément, nous montrons que notre système se comporte comme un serveur Latence-Débit (*LR - Latency Rate server*) et par conséquent est capable d'offrir des garanties sur le délai et le débit ;
– la troisième partie consiste en une évaluation de performances de notre proposition sur la base de simulations avec le simulateur ns-2 (ns2 05) appuyée d'une étude comparative avec des travaux ne considérant pas l'agrégation.

[1] Un lien *backhaul* relie l'épine dorsale de l'Internet et les réseaux de bordure

4.1 Notre proposition

4.1.1 Définition des moyens

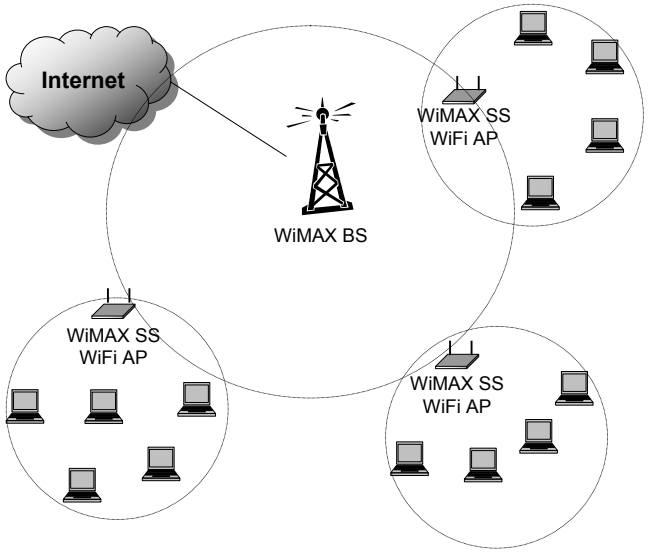

Fig. 4.1 – *Un des cas d'utilisation que nous visons*

Notre proposition est basée sur les deux éléments fondamentaux suivants :

- un algorithme de contrôle d'admission par connexion qui permet de garantir la QoS demandée par les flux UGS, rtPS et nrtPS et de protéger les flux BE contre la famine.
- La notion d'agrégation exprimée pendant la durée de vie des connexions :
  - au niveau de requêtes agrégées de bande passante par SS (c'est à dire résultant de toutes les connexions de la SS),
  - au niveau de l'allocation de bande passante de la BS aux SS sur la base des requêtes agrégées reçues des SS.

Plus précisément, la modalité de gestion de bande passante que nous proposons fonctionne comme suit : tous les flux montants (dans le sens SS vers BS) doivent se soumettre à un contrôle d'admission au niveau de la BS avant de pouvoir accéder au réseau. Le contrôle d'admission permettra à la BS d'assurer à l'ensemble des flux admis la qualité de service requise.

Dans le cadre de notre proposition, nous définissons la notion du contrat de niveau MAC de la SS (à ne pas confondre avec le SLA - *Service Level Agreement*) comme étant la somme des débits moyens requis de tous les flux admis de la SS à contraintes temporelles (ce sont les flux UGS et rtPS). Ce contrat est maintenu au niveau de la BS pour chaque SS et sera mis à jour à chaque admission/libération d'un flux.

Chaque SS devra, de façon périodique, communiquer à la BS les besoins de toutes ses connexions via une requête unique agrégée qui remplacera le système complexe de requêtes par connexion que préconise le standard 802.16. Cette nouvelle requête est ce que l'on appelle dans la suite du rapport la requête agrégée. Cette requête agrégée est envoyée à la BS par chaque SS de façon périodique (par exemple toutes les supertrames qui ont une durée que nous notons FS-*Frame Size*). Une caractéristique importante de notre système est de faire en sorte que la BS réponde positivement à une requête, qui respecte le contrat, dans un délai qui ne dépasse pas la période d'envoi de la requête : la requête, étant envoyée toutes les

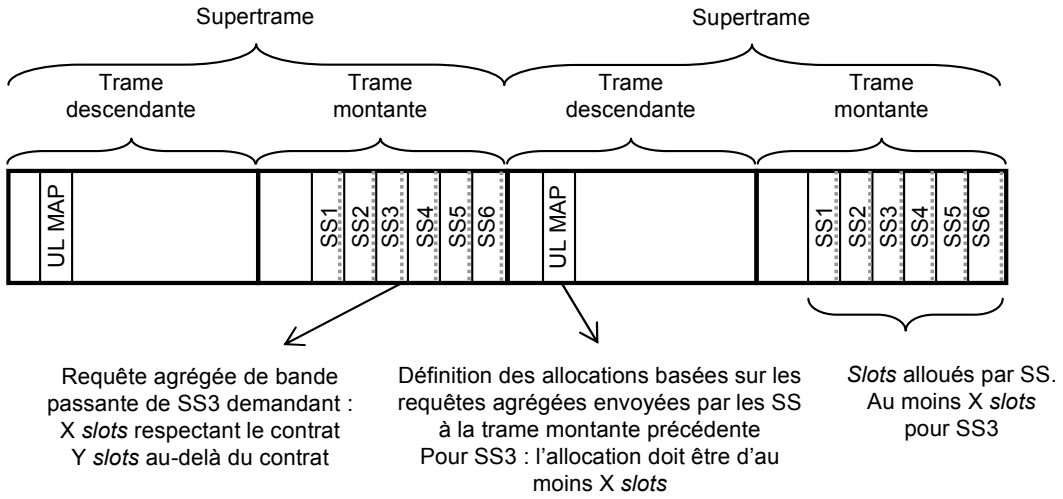

FIG. 4.2 – *Allocation de la bande passante*

supertrames, la BS est dans l'obligation d'allouer la bande passante demandée, qui respecte le contrat, dans la supertrame suivante (voir figure 4.2). Nous instaurons cette obligation afin de pouvoir donner des garanties strictes de l'ordre de la dizaine de *ms* sur les délais d'accès.

4.1.2 L'algorithme de contrôle d'admission

4.1.2.1 Présentation

L'algorithme de contrôle d'admission a pour rôle d'effectuer la réservation de *slots*[2] afin d'assurer aux flux UGS et rtPS les garanties sur les débits et les délais, de garantir un débit moyen aux flux nrtPS et de protéger les flux BE de la famine. L'algorithme de contrôle d'admission est exécuté de façon individuelle pour chaque nouvelle connexion, cependant la condition d'admission diffère selon le service demandé. Nous définissons une fonction de coût permettant d'évaluer le montant de *slots* à réserver à l'admission d'un nouveau flux. La fonction de coût prend, en entrée, un débit à assurer et une durée indiquant l'échéance au bout de laquelle l'allocation des ressources doit être effective (dans l'exemple de la figure 4.2 ainsi que dans la suite du chapitre, nous considérons une contrainte d'un FS pour les flux à contraintes temporelles). La fonction de coût donne, en sortie, le nombre de *slots* à réserver par supertrame permettant d'assurer le débit avec la contrainte donnée en entrée.

- Les flux UGS et rtPS sont considérés, dans la décision de contrôle d'admission, de façon individuelle et avec une contrainte d'allocation d'un FS. Le contrôle d'admission réserve, à chaque flux admis de ce type, toutes les supertrames, assez de *slots* pour servir ce flux. Le contrôle d'admission procède ainsi afin que le système puisse donner les garanties nécessaires de débit et de délai.
- Pour les flux nrtPS, le coût en *slots* n'est pas calculé de façon individuelle. Le contrôle d'admission calculera le coût de l'ensemble des flux nrtPS actifs auquel est rajouté le

[2]Il est important de noter que la notion de *slot* est fortement liée à la couche physique utilisée. Nous utilisons ici le terme générique de *slot* pour signifier l'unité d'allocation utilisée par le protocole d'ordonnancement dans la construction des messages UL-MAP : il peut donc s'agir d'un *minislot* pour les couches physiques à porteuse unique ou d'un couple symboles-porteuse pour les couches basées sur OFDM ou OFDMA.

flux demandant l'admission avec une contrainte d'allocation d'une seconde (ce choix n'est pas arbitraire, il provient du fait que le standard (802.16 04) préconise une période de scrutation des flux nrtPS de l'ordre de la seconde). Le coût des nouvelles connexions nrtPS est ainsi noyé dans celui de l'ensemble des connexions nrtPS actives. Le calcul du coût d'un nouveau flux en l'intégrant dans un agrégat de flux actifs (procédure utilisée pour nrtPS) permet de réduire le coût de ces nouveaux flux par rapport à une réservation par flux. La réservation par flux nous permet de garantir à la fois les débits et les délais (nécessaires pour les flux UGS et rtPS) ; la réservation par agrégat nous permet de garantir les débits mais relâche la contrainte sur le délai (qui de toute façon n'est pas nécessaire pour nrtPS) nous permettant ainsi d'augmenter l'admissibilité.

– Les connexions BE seront acceptées systématiquement. Il n'y a aucune garantie à offrir à un flux BE, et l'accès étant sans contention, l'admission d'un flux BE n'influera pas le service offert aux autres flux. Les flux Best Effort seront servis au mieux. Cependant, comme il est souvent recommandé, et afin d'éviter à ce genre de flux la famine, un pourcentage des *slots* montants leurs seront réservés d'office : ces *slots* sont considérés occupés par le contrôle d'admission lors des décisions d'admission des flux UGS, rtPS et nrtPS.

4.1.2.2 L'algorithme

L'algorithme 2 présente l'algorithme de contrôle d'admission utilisé dans l'architecture que nous proposons. Nous définissons dans ce qui suit les différents paramètres utilisés par l'algorithme :

– *TotalAvailableSlots* est une variable représentant le nombre de *slots* d'allocation, par supertrame, sur la voie montante qui n'ont pas été réservés par le contrôle d'admission. Cette variable est initialisée à un pourcentage du nombre de *slots* de la voie montante, le reste étant réservé aux flux *Best Effort*. Cette variable est mise à jour à chaque admission ou libération d'un flux.
– $SS_j.Rate$ représente le contrat de la station SS_j. Cette variable est initialisée à 0 et est mise à jour à chaque fois qu'une nouvelle connexion à contraintes temporelles de la dite SS est admise ou libérée ; la variable est incrémentée de la moyenne du débit spécifié par la spécification de cette connexion ($TSPEC_i$).
– *nrtPSRate* représente la somme des débits moyens de toutes les connexions nrtPS admises dans le réseau. Cette variable est initialisée à 0.
– *nrtPSSlots* est la variable représentant le nombre de *slots* réservés par l'ensemble des connexions nrtPS admises par le réseau par supertrame. Cette variable est initialisée à 0.

Nous définissons aussi la fonction $CalculateSlotCost(R, D)$ qui calcule le nombre de *slots* d'une trame montante qui devrait être réservé afin de servir un débit généré R (exprimé en $bits/s$) avec la contrainte temporelle d'allocation D. Cette valeur est calculée de la façon suivante :

$$SlotCost(R, D) = \lceil \frac{R * D}{StationTxRate * SlotSize} * \frac{1}{FrameCount} \rceil$$

où $FrameCount = \lceil \frac{D}{FS} \rceil$ est le nombre de supertrames que couvre la durée D, $StationTxRate$ est le débit de la transmission physique (exprimé en $bits/s$) de la station en question et $SlotSize$ est la durée d'un *slot* (dépendant de la couche physique utilisée, tel que précisé précédemment).

Techniquement, l'algorithme fonctionnera de la façon suivante :

Algorithme 2 : Admission($Flow_i, TSPEC_i, SS_j$)

```
1  if TSPEC_i.TypeOfService is UGS OR rtPS then
2      N = CalculateSlotCost(TSPEC_i.Rate, FS);
3      if N < TotalAvailableSlots then
4          Admit(Flow_i);
5          TotalAvailableSlots = TotalAvailableSlots − N;
6          SS_j.Rate = SS_j.Rate + TSPEC_i.Rate;
7      else
8          Reject(Flow_i);
9  else if TSPEC_i.TypeOfService is nrtPS then
10     N = CalculateSlotCost(nrtPSRate + TSPEC_i.Rate, 1s);
11     if (N − nrtPSSlots) < TotalAvailableSlots then
12         Admit(Flow_i);
13         TotalAvailableSlots = TotalAvailableSlots − (N − nrtPSSlots);
14         nrtPSSlots = N;
15         nrtPSRate = nrtPSRate + TSPEC_i.Rate;
16     else
17         Reject(Flow_i);
18 else if TSPEC_i.TypeOfServiceisBE then
19     Admit(Flow_i);
```

- Pour les flux UGS et rtPS : un calcul du nombre de *slots* à réserver sera effectué avec la contrainte d'allocation d'une seule supertrame (ligne 2 de l'algorithme 2). Le résultat est comparé au nombre de *slots* disponibles (ligne 3) ; s'il y a assez de *slots* disponibles, le flux sera admis (ligne 4), la réservation des *slots* sera effectuée (ligne 5) et le contrat de la SS sera mis à jour (ligne 6) ; sinon le flux sera refusé.
- Pour les flux nrtPS : le calcul de coût est fait sur l'ensemble des flux nrtPS actifs avec une contrainte d'allocation d'une seconde. Le résultat du calcul de la ligne 10 (placé dans la variable intermédiaire N) est donc le coût en *slots* de l'ensemble des flux nrtPS considérés (tous les flux nrtPS actifs et le flux nrtPS en attente d'admission). Si le nombre de *slots* supplémentaires ($N - nrtPSSlots$, N étant le coût total, $nrtPSSlots$ étant le nombre de *slots* déjà réservés par les flux nrtPS) est disponible (ligne 11) le nouveau flux est admis (ligne 12), la réservation des *slots* supplémentaires est effectuée (ligne 13) et les variables servant au calculs pour nrtPS sont mises à jour (lignes 14 et 15). S'il n'y a pas assez de *slots* disponibles, le flux est refusé. Il est à noter que le nombre de *slots* supplémentaires peut être nul (si $N = nrtPSSlots$). Le calcul des *slots* étant arrondi à la valeur entière supérieure, les besoins du nouveau flux nrtPS peuvent être déjà couverts par les admissions nrtPS précédentes.
- Les flux BE sont admis systématiquement (lignes 18 et 19).

4.1.3 Construction dans les SS des requêtes agrégées

Nous proposons, dans un contexte où le nombre de connexions simultanément actives est assez élevé (de l'ordre de la centaine), de remplacer le système complexe de requête de bande passante par une requête agrégée envoyée de façon régulière par la SS. Nous fonctionnons dans un mode GPSS (ce qui donnera une liberté absolue à la SS de redistribuer la bande passante allouée). La requête agrégée sera envoyée à chaque supertrame, à la fin de la période allouée à chaque SS. La requête agrégée fera office de photographie instantanée de l'état des files de la station. Elle y exprimera le besoin de toutes ses files. La requête agrégée est formée de deux champs, donnant à la BS l'information nécessaire à l'ordonnancement de la voie montante. Nous donnons dans le paragraphe suivant certaines définitions servant à l'expression des deux champs de la requête agrégée.

4.1.3.1 Définitions

Considérons une période d'activité des flux à contraintes temporelles de SS_j débutant à l'instant 0. Afin de définir la sémantique de chacun des champs de la requête agrégée transmise par la SS_j, nous utilisons les notations suivantes :

- $SS_j.Rate$ le contrat de la station SS_j, c'est donc la somme des débits moyens de tous les flux à contraintes temporelles (UGS et rtPS) de SS_j.
- $W_j(t)$ est le service reçu (entre l' instant 0 et l'instant t) par les flux contraints (UGS et rtPS) de SS_j.
- t^k_{xj} est l'instant de début de la période de transmission allouée à la SS_j sur la voie montante de la $k^{\text{ième}}$ supertrame WiMAX.
- t^k_{rj} est l'instant de construction et d'envoi de la requête agrégée de la SS_j sur la voie montante de la $k^{ième}$ supertrame WiMAX. Cet instant se trouve à la fin de la période de transmission allouée à SS_j.
- $Q_{UGS}(t^k_{rj})$, $Q_{rtPS}(t^k_{rj})$, $Q_{nrtPS}(t^k_{rj})$ et $Q_{BE}(t^k_{rj})$ représentent respectivement le nombre d'octets UGS, rtPS, nrtPS et BE en attente de transmission dans les files de la station au moment de la construction de la requête agrégée (t^k_{rj}).
- $R_{UGS}(t^k_{rj})$ et $R_{rtPS}(t^k_{rj})$ représentent le nombre d'octets UGS et rtPS, en attente de transmission dans les files de la station à l'instant t^k_{rj} , qui respectent le contrat de la station.
- $\varepsilon_{UGS}(t^k_{rj})$ est la différence entre l'état de la file UGS et les octets contractuels. Elle est calculée comme suit : $\varepsilon_{UGS}(t^k_{rj}) = Q_{UGS}(t^k_{rj}) - R_{UGS}(t^k_{rj})$. Cette valeur correspond à ce que le standard appelle le *Slip* d'une file. Les files UGS, dont les arrivées sont constantes, ne devraient donc voir arriver un tel décalage positif qu'en cas d'incident sur le réseau (une perte de message UL-MAP ou bien une désynchronisation d'horloges). Pour les files rtPS, un tel décalage est plus commun vu la variabilité des arrivées rtPS.

4.1.3.2 Les champs de la requête agrégée

Le premier champ, appelé *Contracted Bytes* (CB), est rempli avec le nombre d'octets présents dans les files de la SS respectant le contrat de cette SS. Le second champ, appelé *Additional Bytes* (AB), constitue la requête de l'ensemble des autres octets présents dans la file d'attente de la SS et qui ne sont pas couverts par le contrat. Ce champs AB, en exprimant tous les besoins supplémentaires de la SS, offre à la BS la possibilité d'allouer les *slots* résiduels en se basant sur une image fréquemment mise à jour de l'état des files des SS et d'améliorer en conséquence les performances des connexions. Prenons l'exemple de la SS_j et expliquons

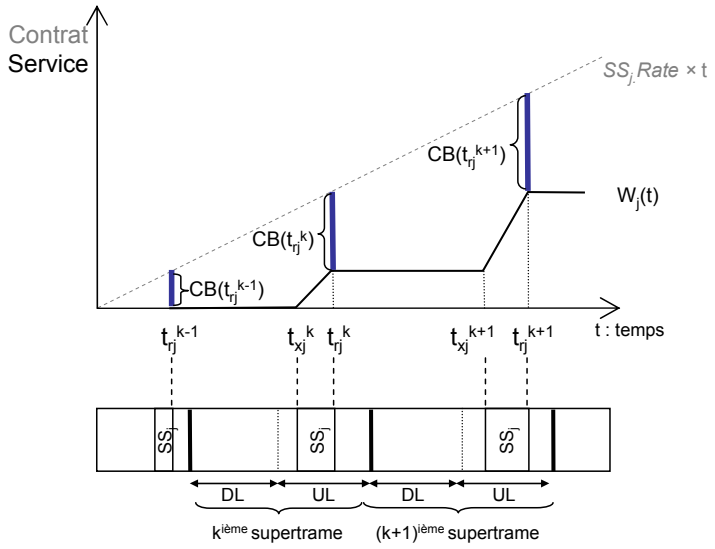

FIG. 4.3 – *Le champ* Contracted Bytes

les contenus des champs CB et AB aux instants de la construction et de l'envoi de la requête agrégée. La figure 4.3 nous permet de présenter le contenu de CB. Nous avons représenté, d'une part, l'évolution temporelle du contrat de SS_j ($SS_j.Rate \times t$) et, d'autre part, l'évolution du service reçu par SS_j pour les flux à contraintes temporelles ($W_j(t)$). Le service est obtenu, pendant la $k^{ième}$ supertrame, entre les instants t^k_{xj} et t^k_{rj} (durant la sous-trame montante-UL). À l'instant t^k_{rj} a lieu la construction et l'envoi de la requête agrégée dont la partie CB (au moins) devra être servie dans la $k+1^{ième}$ supertrame. Ainsi nous avons, à l'instant t^k_{rj} :

$$CB(t^k_{rj}) = SS_j.Rate \times t^k_{rj} - W(t^k_{rj})$$

Le champs AB, que nous ne représentons pas sur la figure, est simplement la différence entre le contenu des files d'attente de SS_j et la valeur du champs CB.
Techniquement, les champs de la requête agrégée seront construits de la façon suivante : le champ CB contient le nombre d'octets UGS et rtPS en file d'attente qui respectent le contrat de la SS en plus de la valeur de *Slip* de la file UGS. Nous servons ainsi UGS et pas rtPS pour respecter la définition de chacune des classes de service. Pour UGS : un dépassement du contrat est un événement exceptionnel qui doit être résolu rapidement ; pour rtPS un tel dépassement est normal vu la nature variante des flux utilisant le service rtPS : nous nous engageons à servir un débit moyen avec une garantie sur le délai, tout dépassement par rtPS sera donc relégué à la partie AB. Le champ AB contient la différence entre l'état des files de la SS et ce qui a été demandé dans le champ CB. Les champs seront donc calculés, au moment de la construction de la requête, de la manière suivante :

$$\begin{aligned} CB(t^k_{rj}) &= min\left[(SS_j.Rate \times t^k_{rj} - W(t^k_{rj})) + \varepsilon_{UGS}(t^k_{rj}), (Q_{UGS}(t^k_{rj}) + Q_{rtPS}(t^k_{rj}))\right] \\ &= R_{UGS}(t^k_{rj}) + R_{rtPS}(t^k_{rj}) + \varepsilon_{UGS}(t^k_{rj}) \\ \text{et } AB(t^k_{rj}) &= Q_{UGS}(t^k_{rj}) + Q_{rtPS}(t^k_{rj}) + Q_{nrtPS}(t^k_{rj}) + Q_{BE}(t^k_{rj}) - CB(t^k_{rj}) \end{aligned} \quad (4.1)$$

4.1.4 Répondre dans la BS aux requêtes agrégées : les allocations

Comme nous l'avons déjà spécifié, une des caractéristiques de notre architecture est que la BS se doit de répondre positivement à toute requête comprise dans le contrat. L'allocation doit avoir lieu dans la supertrame WiMAX qui suit celle portant la requête (figure 4.2). Ceci est rendu possible par le contrôle d'admission. L'algorithme d'ordonnancement de la voie montante fonctionnant au niveau de la BS doit en premier lieu allouer à chaque SS un nombre de *slots* suffisant à la transmission du nombre d'octets indiqué dans le champ CB de la requête agrégée de la SS (ceci est évidemment réalisé après un contrôle par la BS de la conformité de la requête au contrat). Les autres *slots* seront distribués entre les SS. S'il reste assez de *slots* pour servir l'ensemble des requêtes additionnelles (indiquées par les champs AB des requêtes), les *slots* seront distribués en conséquence. Sinon, les *slots* doivent être distribués proportionnellement aux contenus des champs AB des requêtes des SS. L'algorithme 3 décrit la procédure d'allocation. Nous appelons $AllocatedSlots(j)$ le nombre de *slots* alloués pendant la procédure d'allocation en cours à SS_j, le tableau est initialisé à 0. CB_j et AB_j sont la traduction en termes de *slots* des valeurs des champs CB et AB de la requête agrégée reçue de la part de SS_j dans la supertrame précédente. N est le nombre de SS dans le réseau. $AvailableSlots$ est le nombre de *slots* non encore alloués sur la voie montante concernée, elle est initialisée au nombre total de *slots* à allouer sur la voie montante.

Algorithme 3 : Procédure d'allocation et de construction de l'UL-MAP

```
1 for j in 1 to N do
2     AllocatedSlots(j) = AllocatedSlots(j) + CB_j;
3     AvailableSlots = AvailableSlots − CB_j;
4 if Σ_{j=1}^{N} AB_j < AvailableSlots then
5     for j in 1 to N do
6         AllocatedSlots(j) = AllocatedSlots(j) + AB_j;
7 else
8     for j in 1 to N do
9         AllocatedSlots(j) = AllocatedSlots(j) + (AB_j / Σ_{j=1}^{N} AB_j) * AvailableSlots;
```

4.1.5 Allocation de *slots* par la SS aux flux

Un autre algorithme est essentiel pour le système : c'est celui qui permet, au niveau de la SS, de distribuer les *slots* attribués par la BS entre les différentes connexions. Cet algorithme fonctionne comme suit : il distribue d'abord, aux connexions à contraintes temporelles (UGS et rtPS) , les *slots* nécessaires à la transmission des octets contractuels requis. Ensuite, les *slots* restants sont distribués, suivant la technique *Round Robin*, en premier lieu aux connexions UGS et rtPS le nécessitant, puis aux connexions nrtPS et enfin aux connexions BE. Ceci nous permet de servir, de façon prioritaire, les connexions à contraintes temporelles assurant ainsi une minimisation de leurs délais d'accès moyens. Les flux nrtPS seront correctement servis sur le long terme vu qu'une réservation de ressources a été effectuée au niveau du contrôle d'admission. Les flux BE ne subiront pas de famine, grâce à la réservation d'office que l'on a réalisée pour ce genre de trafic.

Nous estimons qu'il n'est pas nécessaire d'implémenter d'autres algorithmes d'ordon-

nancement (tel que celui qui organise l'ordre des connexions d'une SS au sein de la période allouée à une SS). Ces algorithmes ajouteront de la complexité avec peu d'effet sur les délais. En effet, pour notre exemple, chaque SS aura droit à une partie de la voie montante (qui sera souvent de l'ordre de quelques ms) ; organiser l'accès des connexions dans ce petit intervalle n'aura que très peu d'effet à notre avis.

4.1.6 Ajout de flexibilité : les politiques d'anticipation

Il peut être intéressant, pour les flux ayant des contraintes temporelles, d'anticiper sur leurs besoins, c'est à dire qu'au moment de l'envoi de la requête agrégée (instant t), d'exprimer un besoin de service qui serait normalement dû à un point chronologique ultérieur (instant $t + \Delta$), c'est donc un besoin supérieur à celui que l'on demande normalement. Une politique d'anticipation permet d'influencer directement les délais moyens des flux : en effet, plus la SS est capable d'anticiper sur les besoins futurs, meilleurs seront les délais d'accès en moyenne. Cette politique est d'autant plus intéressante que la durée de la supertrame est élevée (le délai minimal d'accès étant corrélé à cette durée). Le contrat de la SS doit évidemment intégrer le "surplus" induit par une politique d'anticipation. Dans ce document, nous présentons cette politique. Son évaluation sera faite dans le cadre de travaux futurs.
Évidemment, une politique d'anticipation aura un effet négatif sur l'admissibilité des flux mais elle améliore la garantie sur le délai. Nous présentons les politiques d'anticipation que nous proposons dans un cadre général où on parle aussi de la proposition présentée dans les paragraphes précédents et que nous appelons maintenant **la politique sans anticipation**.
Nous considérons trois politiques :

- **la politique sans anticipation** (où $\Delta = 0$), la requête couvre ce qui est dû jusqu'à l'instant d'envoi de la requête,
- **la politique avec anticipation simple** (où $\Delta = FS$), l'anticipation couvrira une supertrame supplémentaire,
- **la politique avec anticipation double** (où $\Delta = 2FS$), l'anticipation couvrira deux supertrames supplémentaires,

Techniquement, le choix d'une politique se manifestera par :

- une modification de la notion de contrat, celui-ci devant prendre l'anticipation en compte ;
- une modification de la fonction de calcul de coût au niveau du contrôle d'admission ;
- une modification de la manière de construire la requête agrégée.

4.1.6.1 Visualisation des politiques

La figure 4.4 propose une comparaison entre les politiques d'anticipation pour une SS_j. Dans cette figure, la politique **sans anticipation** correspond à ce qu'on a présenté dans les paragraphes précédents (la représentation correspond à celle de la figure 4.3). Nous représentons aussi la valeur du champs CB des requêtes agrégées et le service reçu par SS_j avec la politique **anticipation simple**. Nous ne représentons pas la politique **anticipation double** sur le graphe afin d'en faciliter la lecture.
Comme nous le voyons dans la figure 4.4, le champs CB d'une requête agrégée envoyée par SS_j à l'instant t_{rj}^{k-1} peut couvrir :

- le service dû à l'instant de l'envoi de la requête avec la politique **sans anticipation** ;
- le service dû à $t_{rj}^{k-1} + FS$ avec la politique **anticipation simple** ;
- le service dû à $t_{rj}^{k-1} + 2FS$ avec la politique **anticipation double**.

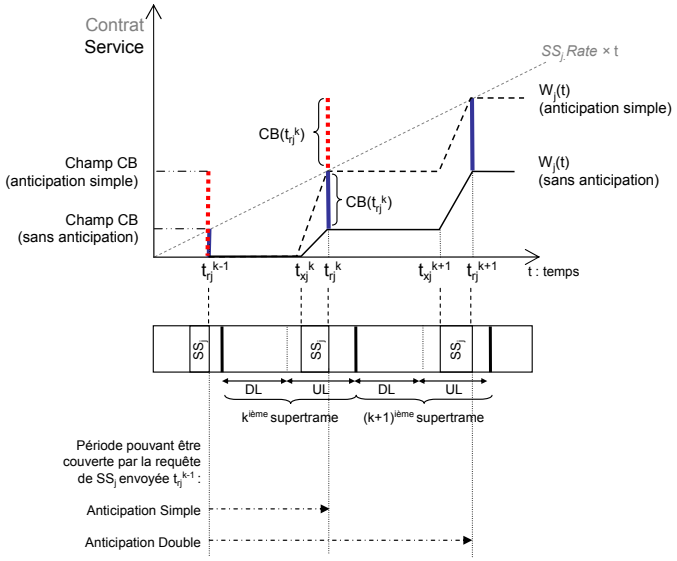

FIG. 4.4 – *Le champ* Contracted Bytes *avec les politiques* ***sans anticipation*** *et* ***anticipation simple****. La politique* ***anticipation simple*** *permet d'approcher l'enveloppe définie par* Rate

Notre représentation considère que les instants d'envoi de la requête agrégée sont périodiques (tous les FS) ce qui n'est évidemment pas le cas tout le temps.
Ainsi le service reçu approche, avec la politique **anticipation simple** l'enveloppe définie par le contrat. Ces politiques permettent de réduire les délais d'accès moyens.

4.1.6.2 Sur la construction de la requête agrégée

La modification de la manière de construire la requête implique directement une modification du service fourni. Donnons une formule générale des champs de la requête aggrégée (construite à l'instant t_{rj}^k)) :

$$CB(t_{rj}^k) = min\{(SS_j.Rate \times t - W(t_{rj}^k)) + \varepsilon_{UGS}, \sum Q_{SS_j}(t_{rj}^k)\}$$
$$\text{avec } \sum Q_{SS_j}(t_{rj}^k) = Q_{UGS}(t_{rj}^k) + Q_{rtPS}(t_{rj}^k) + Q_{nrtPS}(t_{rj}^k) + Q_{BE}(t_{rj}^k)$$
$$\text{et } AB(t_{rj}^k) = \sum Q_{SS_j}(t_{rj}^k) - CB(t_{rj}^k)$$

La valeur de t définit le degré d'anticipation. Nous instancions t à trois valeurs représentants les trois politiques d'approvisionnement que nous préconisons. La figure 4.3 donne une représentation de la valeur de **CB** avec la politique **Sans anticipation** ; la figure 4.4 donne une représentation de la valeur de **CB** avec les politiques **Anticipation simple** et **Anticipation double** :

- **Sans anticipation** la requête prend en compte, pour le calcul du champs CB, le service dû à l'instant de construction de la requête : $t = t_{rj}^k$.
- **Anticipation simple** la requête prend en compte, pour le calcul du champs CB, le service dû à l'instant : $t = t_{rj}^k + FS$.
- **Anticipation double** la requête prend en compte, pour le calcul du champs CB, le service dû à l'instant : $t = t_{rj}^k + 2FS$.

Il est cependant important de noter qu'en cas d'anticipation, la SS est contrainte de ne pas dépasser, sur l'ensemble de sa requête, le nombre d'octets effectivement mis en attente dans

les files de la SS (prenant donc en compte les files nrtPS et BE), nous évitons ainsi tout gaspillage des ressources.

4.2 Sur le comportement de type serveur *Latency Rate*-LR de la proposition

4.2.1 Les serveurs de type LR

Un serveur de type LR est un modèle d'algorithme d'ordonnancement de trafic défini dans (Stiliadis 98) par Stiliadis et Varma. Le comportement d'un algorithme respectant ce modèle dérive de deux paramètres, la latence et le débit alloué. Il permet de donner des garanties sur les délais de bout en bout, les dimensions des files d'attente ainsi que sur les débits moyens. Nous donnons dans la suite la définition du modèle ainsi que les différentes garanties qu'il accorde aux flux.
Soit une session i, nous notons ρ_i le débit alloué à la session i. $A_i(\tau, t)$ spécifie les arrivées de la session i dans l'intervalle $(\tau, t]$ et $W_i(\tau, t)$ le montant du service reçu par la session i dans le même intervalle. Nous définissons $Q_i(t)$ comme le volume de trafic de la session i placé en file d'attente :

$$Q_i(t) = A_i(\tau, t) - W_i(\tau, t)$$

Lorsque $Q_i(t)$ est strictement positif, nous disons que la session a du retard (la session est *backlogged*). Une période de *backlog* (ou une période de retard) de la session i est toute période durant laquelle du trafic pour la session arrive dans la file d'attente. Nous définissons aussi la période de charge de la session i comme étant l'intervalle maximale $(\tau_1, \tau_2]$ pour lequel, pour tout $t \in (\tau_1, \tau_2]$ l'ensemble des arrivées de la session dépassent le service théorique dû au débit alloué :

$$A_i(\tau_1, t) \geq \rho_i(t - \tau_1)$$

Il y a donc une différence notable entre la période de retard de la session (*backlogged period*) et la période de charge de la session (*busy period*). La première dépend de l'état véritable du système avec le service qu'il reçoit alors que la deuxième n'est définie que de façon théorique par rapport à la valeur du débit alloué ρ_i. Soit la $j^{ième}$ période de charge commençant à l'instant τ, $W_{i,j}(\tau, t)$ est le service reçu par la session i servant le trafic arrivé pendant la $j^{ième}$ période de charge.
Un serveur de type *LR* est caractérisé, pour chaque session, par deux paramètres : la latence Θ_i et le débit alloué ρ_i. Un serveur est dit de type *LR* si, pour une $j^{ième}$ période de charge débutant à l'instant τ_1 et étant donné τ_2 l'instant où le dernier bit arrivé pendant la $j^{ième}$ période de charge quitte le serveur, l'inégalité suivante est satisfaite pour tout instant $t \in [\tau_1, \tau_2]$:

$$W_{i,j}(\tau_1, t) \geq max(0, \rho_i(t - \tau_1 - \Theta_i))$$

Un serveur de type *LR* est donc un serveur qui permet de donner une enveloppe par le bas du service offert à une session pendant toute période de charge de cette session. La figure 4.5 présente un exemple du comportement d'un serveur *LR*.

4.2.2 Démonstration

Considérons un flux i et considérons la $m^{ième}$ période de backlog du flux i qui débute à s^m. Durant cette période de backlog, le système d'ordonnancement local de la station SS transmet

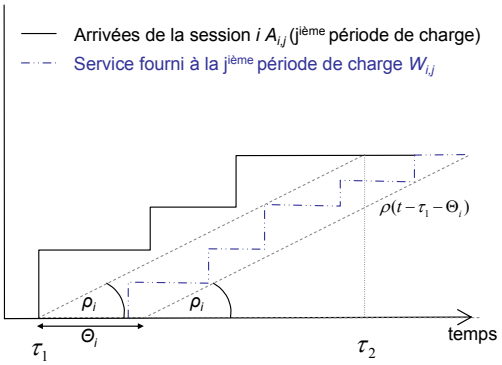

FIG. 4.5 – *Comportement d'un serveur de type* LR

de façon cyclique (à chaque supertrame k) une requête de bande passante $BwReq(t_{ri}^k)$ avec

$$BwReq(t_{ri}^k) = min\{Q_i(t_{ri}^k), max\{0, g_i(t_{ri}^k - s^m) - W_i(s^m, t_{ri}^k)\}\} \tag{4.2}$$

avec g_i un débit moyen que le réseau cherche à garantir au flux i.
Si chacune des requêtes de bande passante exprimée au nom du flux i est acceptée par la BS et allouée à la prochaine supertrame : le comportement du système d'ordonnancement vis à vis du flux i suit le modèle LR avec $\rho_i = g_i$ et $\theta_i = 2FS + ULS$, FS est la durée d'une supertrame et ULS est la durée maximale d'une sous-trame montante.

Démonstration. Pour montrer ce comportement LR du système d'ordonnancement vis à vis du flux i, il faut que : pour toute *"Busy Period"* n du flux i qui débute à l'instant α^n et finit à β^n, la quantité de trafic appartenant au flux arrivé depuis le début de la *"Busy Period"* (i.e. α^n) qui a été servie depuis α^n jusqu'à un instant t de la *"Busy Period"* (i.e. $W_{i,n}(\alpha^n, t)$) respecte l'inégalité suivante :

$$W_{i,n}(\alpha^n, t) \geq max\{0, \rho_i(t - \alpha^n - \theta_i)\} \tag{4.3}$$

avec $\Theta_i = 2FS + ULS$.
Deux premiers cas sont à distinguer :

- **Cas 1 :** si $t \leq \alpha^n + \theta_i$, alors l'inégalité est vérifiée ;
- **Cas 2 :** si $t > \alpha^n + \theta_i$

Pour ce deuxième cas, il existe au moins une période de transmission du flux i depuis l'instant α^n. Notons t_{ei}^k l'instant qui clôt la dernière période de transmission du flux i qui précède t. Notons t_{ri}^k l'instant où la requête de bande passante associée à la $k^{ième}$ période de transmission (qui s'achève à t_{ei}^k) est envoyée.

$$BwReq(t_{ri}^k) = min\{B_i(t_{ri}^k), max\{0, \rho_i(t_{ri}^k - s^m) - W_i(s^m, t_{ri}^k)\}\}$$

Nous pouvons également déduire que :

$$t - t_{ri}^k < \theta_i \tag{4.4}$$

En effet, le pire scénario est lorsque $t = t_{ei}^{k+1}$ (instant infinitésimal précédant la fin de la période de transmission).

Pour $n > 1$, notons par τ l'instant de fin de transmission du dernier paquet de la *"Busy Period"* $(n-1)$. Nous allons démontrer deux relations qui seront nécessaires à la suite de la démonstration :

$$W_i(s^m, \tau) \leq \rho_i(\beta^{n-1} - s^m) \tag{4.5}$$

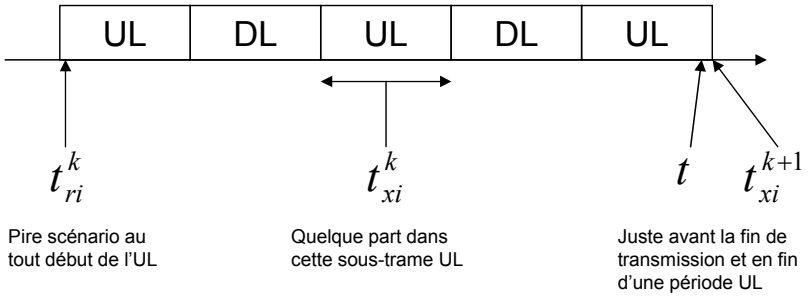

FIG. 4.6 – *Différents instants de transmission et de requêtes*

et que :

$$\tau < t \tag{4.6}$$

Commençons par la preuve de l'inégalité 4.5 :
Puisqu'une arrivée appartient nécessairement à une *"Busy Period"* (arrivée instantanée à un débit infini car livrée par l'application ou par un régulateur de trafic). On a donc :

$$W_i(s^m, \tau) = \sum_{j=n-L}^{n-1} A_i(\alpha^j, \beta^j) - W_{i,n-L}(\alpha^{n-L}, s^m) \tag{4.7}$$

avec $(n-L)\dots(n-1)$ représentant toutes les *"Busy Period"* précédent n qui se sont succédées durant la $m^{ième}$ *Backlog Period* considérée (qui débute à s^m), l'égalité 4.7 représente le cas où une partie de la $(n-L)^{ième}$ *"Busy Period"* a été servie avant la $m^{ième}$ *Backlog Period* (voir figure 4.7). Par définition d'une *Busy Period* j avec les limites α^j, β^j :

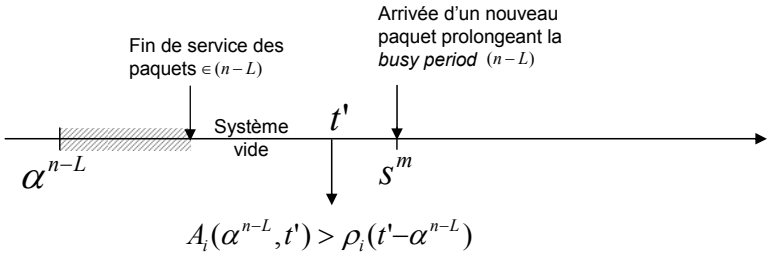

FIG. 4.7 – *Période de* Backlog

$$\sum_{j=n-L}^{n-1} A_i(\alpha^j, \beta^j) \leq \sum_{j=n-L}^{n-1} \rho_i(\beta^j - \alpha^j)$$

De plus, puisque tout le traffic qui est arrivé durant $[\alpha^{n-L}, s^m[$ a été servi (le système est vide à s^m) :

$$W_{i,n-L}(\alpha^{n-L}, s^m) = A_i(\alpha^{n-L}, s^m) \geq \rho_i(s^m - \alpha^{n-L})$$

Puisque s^m appartient à la *Busy Period*$(n-L)$. Ainsi, en reprenant l'égalité 4.7,

$$W_i(s^m,\tau) \leq \sum_{j=n-L}^{n-1} \rho_i(\beta^j - \alpha^j) - \rho_i(s^m - \alpha^{n-L})$$
$$\leq \sum_{j=n-L+1}^{n-1} \rho_i(\beta^j - \alpha^j) + \rho_i(\beta^{n-L} - \alpha^{n-L} - s^m + \alpha^{n-L})$$

$$W_i(s^m,\tau) \leq \rho_i(\beta^{n-1} - s^m) \tag{4.8}$$

Montrons maintenant que si $t > \alpha^n + \theta_i$, alors nécessairement $\tau < t$ (avec n>1). Notons k' le numéro de la première requête de bande passante qui suit l'instant β^{n-1}. Nous avons :

$$t_{ri}^{k'} < t_{ei}^{k'} < t \qquad (t > \alpha^n + \theta_i > \beta^{n-1} + \theta_i)$$

avec $\theta_i = 2FS + ULS$. Nous avons aussi :

$$W_i(s^m, t_{ei}^{k'}) = W_i(s^m, t_{ri}^{k'}) + min\{B_i(t_{ri}^{k'}), \rho_i(t_{ri}^{k'} - s^m) - W_i(s^m, t_{ri}^{k'})\}$$

Deux cas à considérer :

- **Si** $B_i(t_{ri}^{k'}) \geq \rho_i(t_{ri}^{k'} - s^m) - W_i(s^m, t_{ri}^{k'})$ alors

$$W_i(s^m, t_{ei}^{k'}) = W_i(s^m, t_{ri}^{k'}) + \rho_i(t_{ri}^{k'} - s^m) - W_i(s^m, t_{ri}^{k'})$$
$$= \rho_i(t_{ri}^{k'} - s^m) \geq \rho_i(\beta^{n-1} - s^m)$$

  En se référant à l'inégalité 4.8, nous avons $\tau = t_{ei}^{k'}$.
- **Si** $B_i(t_{ri}^{k'}) < \rho_i(t_{ri}^{k'} - s^m) - W_i(s^m, t_{ri}^{k'})$, nous avons $BwReq(t_{ri}^{k'}) = B_i(t_{ri}^{k'})$, puisque le système d'ordonnancement de la BS accordera cette bande passante, donc $\tau = t_{ei}^{k'} < t$

Reprenons le fil de la démonstration pour prouver l'inégalité 4.3, si $t > \alpha^n + \theta_i$. Puisque $\tau < t$ (c'est le cas pour $n > 1$ ainsi que pour n=1 auquel cas la *Busy Period* 0 n'existe pas, nous pouvons donc considérer que $\tau = \alpha^1 = s^1$, τ vérifie bien cette inégalité). Nous avons :

$$W_{i,n}(s^m, t) = W_i(s^m, t) - W_i(s^m, \tau) \tag{4.9}$$

Cette égalité est également vraie pour $n = 1$ ($W_i(s^m, t) = W_i(s^1, s^1) = 0$).

Deux cas se présentent :

- **Cas 1 où** $\rho_i(t_{ri}^k - s^m) - W_i(s^m, t_{ri}^k) \leq B_i(t_{ri}^k)$ **:** on a alors :

$$W_i(s^m, t) = W_i(s^m, t_{ri}^k) + \rho_i(t_{ri}^k - s^m) - W_i(s^m, t_{ri}^k)$$
$$W_i(s^m, t) = \rho_i(t_{ri}^k - s^m) \tag{4.10}$$

  En combinant les équations 4.9 et 4.5,4.8 et 4.10 :

$$W_{i,n}(s^m, t) \geq \rho_i(t_{ri}^k - s^m) - \rho_i(\beta^{n-1} - s^m)$$

  puisque $\beta^{n-1} \leq \alpha^n$

$$W_{i,n}(s^m, t) \geq \rho_i(t_{ri}^k - s^m) - \rho_i(\alpha^n - s^m)$$
$$\geq \rho_i(t_{ri}^k - \alpha^n)$$

  en se référant à l'équation 4.4

$$W_{i,n}(s^m, t) \geq \rho_i(t - \alpha^n - \theta_i)$$

– **Cas 2 où** $\rho_i(t^k_{ri} - s^m) - W_i(s^m, t^k_{ri}) < B_i(t^k_{ri})$ **:** cela veut dire que tout ce qui est arrivé entre α^n et t^k_{ri} sera pris en compte dans la requête de bande passante envoyée à t^k_{ri}. Ainsi :

$$W_{i,n}(\alpha^n, t) = W_{i,n}(\alpha^{n-1}, \tau) + A_i(\alpha^n, t^k_{ri})$$

Le premier terme ($W_{i,n}(\alpha^{n-1}, \tau)$) équivaut à tout ce qui appartient à la *Busy Period* $(n-1)$ (ce terme est nul si $n = 1$ ou si $\tau < \alpha^n$). Ainsi

$$W_{i,n}(\alpha^n, t) \geq A_i(\alpha^n, t^k_{ri}) \geq \rho_i(t^k_{ri}, \alpha^n)$$

D'après l'équation 4.4, $t - t^k_{ri} < \theta_i$. Donc :

$$W_{i,n}(\alpha^n, t) \geq \rho_i(t - \alpha^n - \theta_i)$$

□

4.2.3 Propriétés des algorithmes d'ordonnancement de type LR

Les algorithmes d'ordonnancement de type LR présentent des propriétés intéressantes du point de vue de la QoS. Considérons de nouveau le flux i, servi par une stratégie d'ordonnacement qui a un comportement vis-à-vis de ce flux de type $LR(\rho_i, \theta_i)$. Une première propriété intéressante concerne le débit : un débit moyen de ρ_i est garanti au flux i. Considérons maintenant que le trafic du flux i qui arrive à l'ordonnanceur est borné et est conforme au modèle (σ_i, ρ_i)[3] avec σ_i représentant une taille de rafale maximale et ρ_i un débit sur le long terme (un tel modèle peut être obtenu en appliquant au trafic entrant un régulateur de type *token bucket*). Les propriétés suivantes peuvent être garanties au flux i :

– le délai d'attente d'un paquet du flux est borné, avec pour borne $\frac{\sigma_i}{\rho_i} + \theta_i$;
– le volume du trafic du flux i en attente de service a une borne maximale valant $\sigma_i + \rho_i\theta_i$;
– le trafic en sortie est conforme à un modèle $(\sigma_i^{sortie}, \rho_i)$ avec : $\sigma_i^{sortie} = \sigma_i + \theta_i\rho_i$.

Nous pouvons également citer une dernière propriété intéressante dans le cadre d'un réseau hybride, celle de la transitivité : une cascade de serveurs de type LR est équivalente à un serveur de type LR préservant ainsi les propriétés que nous venons d'exposer.

4.2.4 Lien de notre proposition avec les serveurs LR

Le système d'allocation que nous avons proposé offre à tous les flux UGS et rtPS d'une station un débit égal, sur le long terme, à la somme de leurs exigences de débits moyens exprimées.

En effet, par construction, et grâce à l'algorithme de contrôle d'admission que nous avons proposé, la requête de bande passante contenue dans le champs CB de la requête agrégée envoyée par une station sera satisfaite dans la supertrame qui suit l'envoi de la requête. Nous rappelons que la partie CB de la requête agrégée est exprimée dans l'équation 4.1 et est équivalente à l'équation 4.2 avec comme définition du flux : l'agrégat de flux UGS et rtPS d'une station. D'après la démonstration, le comportement de notre système d'allocation vis-à-vis des flux UGS et rtPS d'une station est de type LR avec pour paramètre de débit $SS_j.Rate$ et pour paramètre de délai $2FS + ULS$.
Notre système d'allocation va donc hériter des propriétés que nous avons énoncées au paragraphe 4.2.3 concernant les garanties de débit, de délai d'accès ainsi que le profil de trafic en sortie de SS_j pour l'agrégat des flux UGS et rtPS. Nous avons d'ailleurs observés ces garanties dans les simulations que nous présentons dans le paragraphe suivant.

[3]Il est à noter que des résultats équivalents existent pour des modèles plus élaboré, intégrant un paramètre supplémentaire (en plus de σ_i et de ρ_i) : le débit de crête.

4.3 Analyse

4.3.1 Contexte de l'analyse

Nous présentons dans ce paragraphe l'analyse que nous avons effectuée pour évaluer notre solution. Le but de cette analyse est d'évaluer la performance de notre solution dans l'absolu par rapport aux objectifs que nous annonçons, à savoir : une simplicité et une efficacité dans un cadre où un grand nombre de connexions sont simultanément actives. La simplicité est aisément montrée : le système de requête est plus simple au vu de l'introduction de la requête agrégée en place du système de requêtes de WiMAX bien plus complexe ; les algorithmes d'ordonnancement que nous utilisons sont eux aussi courants et assez simples. L'analyse de l'efficacité sera réalisée par simulation. Nous observerons, dans divers scénarios qui se rapportent au cadre que nous préconisons, la performance du contrôle d'admission, l'utilisation du réseau, le respect des garanties données aux flux contraints et les caractéristiques du service que notre solution offre. Cette efficacité sera aussi étudiée en comparaison à une solution de la littérature (Freitag 07) qui peut être considérée comme représentative des solutions existantes. Nous utilisons, pour effectuer l'analyse de notre proposition, le simulateur ns-2 (ns2 05). Nous utilisons le module de WiMAX qu'a développé une équipe de l'Univesidade Estadual de Campinas (Freitag 08). Nous avons intégré au module notre solution et l'avons testé. Nous présentons en premier la solution (Freitag 07), nous donnerons ensuite les scénarios de simulation, les résultats et l'analyse.
Nous avons donc choisi de comparer notre solution à un fonctionnement standard de WiMAX (en termes de requêtes de bande passante) auquel est couplé un algorithme d'ordonnancement de la voie montante. Freitag et. al. utilisent les mécanismes standard de requête de WiMAX. La solution qu'ils proposent gère la bande passante montante en considérant chaque connexion dans l'ordre de priorité instauré par WiMAX. L'algorithme d'ordonnancement qu'ils proposent est sensé respecter les besoins des connexions en termes de bande passante.

4.3.2 Scénarios de simulation

Le tableau 4.1 présente les choix de paramètres effectués pour ces simulations. Dans un réseau où les stations fonctionnent à un débit physique de 40 Mbps, nous avons 1 BS et 10 SS avec, au sein de chaque SS, 10 flux montants simultanément actifs. Parmi ces flux :

- 3 sont des flux à débit faible et constant (portant par exemple une application de voix), utilisant le service UGS ;
- 3 sont des flux à débit moyen et variant (une application de visioconférence par exemple), utilisant le service rtPS ;
- 3 sont des flux à débit moyen sans contraintes temporelles utilisant le service nrtPS ;
- 1 flux à débit fort et constant (une application de transfert de fichier), utilisant le service BE.

En nous basant sur ce profil applicatif qui correspond au contexte qui nous intéresse, nous faisons varier la taille de la supertrame afin de voir son effet sur les différents aspects de fonctionnement de notre proposition.

Nous présentons dans ce qui suit les résultats des scénarios de simulation présentés. Les résultats sont présentés sous l'une des formes suivantes :

- une fonction de distribution de probabilités cumulative (CDF) des délais d'accès des flux à contraintes temporelles (UGS et rtPS) ;

Paramètre	Valeur
Nombre de BS	1
Nombre de SS	10
mode UL/DL	TDD
ratio UL/DL	1/1
Débit Physique	40 Mbps
Nombre de Flux par SS	10
UGS(Nombre par SS ;Débit ;Mode)	(3 ;24kbps ;CBR)
rtPS(Nombre par SS ;Débit ;Mode)	(3 ;64kbps ;ON-OFF)
nrtPS(Nombre par SS ;Débit ;Mode)	(3 ;64kbps ;CBR)
BE(Nombre par SS ;Débit ;Mode)	(1 ;1Mbps ;CBR)
Taille de la Supertrame (FS - *Frame Size*	Varie de 2 à 20 ms
Politique d'approvisionnement	Sans Anticipation
Durée de la simulation	20 s

TAB. 4.1 – *Spécification de la simulation*

– une fonction de distribution de probabilités cumulative (CDF) des délais d'accès des flux nrtPS et BE ;
– le débit utile des flux UGS, rtPS, nrtPS ou BE.

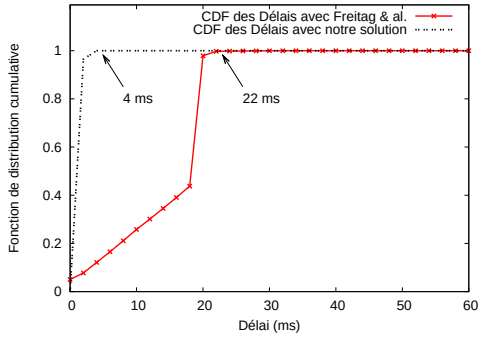

FIG. 4.8 – *Délais d'UGS et rtPS FS=2ms*

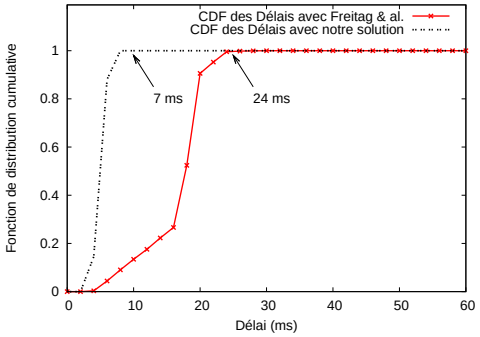

FIG. 4.9 – *Délais d'UGS et rtPS FS=5ms*

4.3.3 Résultats et interprétation

Nous analyserons dans ce paragraphe les différents résultats de cet ensemble de simulations. Nous rappelons qu'entre les différentes simulations, seule la durée de la supertrame WiMAX (FS) a été modifiée. Nous retrouvons l'effet de cette modification sur deux aspects : les délais d'accès des différents flux et l'admissibilité des flux. Nous analyserons ces deux aspects ainsi que la performance de notre solution du point de vue des débits utiles dans ce qui suit.

4.3.3.1 Les délais d'accès

Nous comparons la performance de notre solution, au niveau des délais d'accès, par rapport à la performance de l'algorithme d'ordonnancement proposé par Freitag et. al. Il est clair

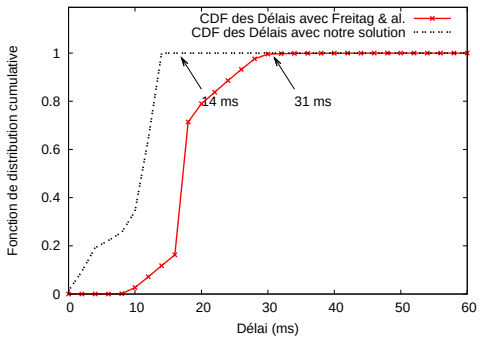

FIG. 4.10 – *Délais d'UGS et rtPS FS=10ms*

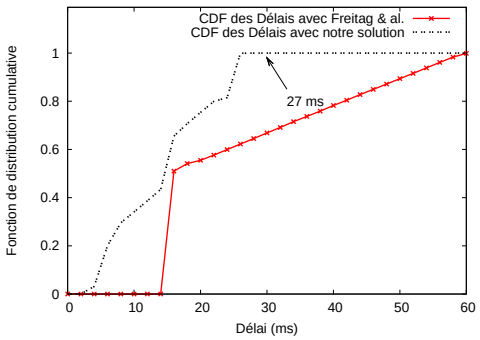

FIG. 4.11 – *Délais d'UGS et rtPS FS=20ms*

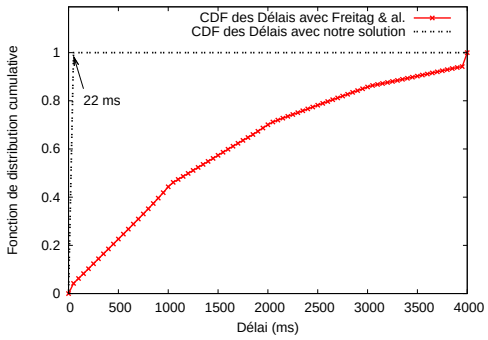

FIG. 4.12 – *Délais de nrtPS FS=2ms*

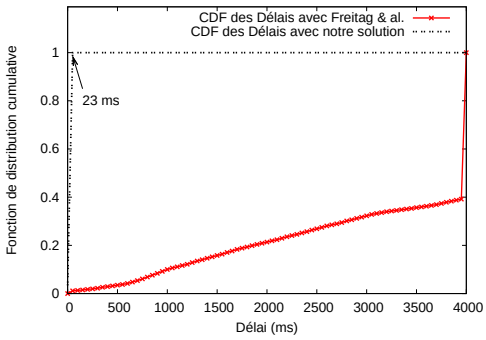

FIG. 4.13 – *Délais de BE FS=2ms*

en observant les différents graphes représentant les délais (les figures 4.8, 4.9, 4.10 et 4.11 pour les flux à contraintes temporelles ; les figures 4.12, 4.14, 4.16 et 4.18 pour les flux nrtPS et les figures 4.13, 4.15, 4.17 et 4.19 pour les flux BE) que notre solution permet d'obtenir de meilleurs délais d'accès pour tous les types de flux, et ce quelque soit la durée de la supertrame. Pour les flux à contraintes temporelles, la garantie sur le délai d'accès est assurée : 100% des paquets ont un délai d'accès inférieur au double de la durée de la supertrame (quelle que soit cette durée). Cette propriété provient de la caractéristique de notre solution obligeant une demande provenant d'un flux à contraintes d'être provisionnée de façon immédiate. Pour les autres types de flux nous atteignons des délais d'accès inférieurs à ceux atteints par la solution de Freitag et. al. La différence entre les bornes maximales des délais des flux non-contraints est conséquente. Cette différence est due à la modification que nous apportons au système de requête et de provisions. En effet, dans notre solution, les flux non-contraints ont l'occasion de s'exprimer sur leurs besoins de façon permanente. A chaque fois que la SS envoie une requête agrégée, celle-ci contient tous les besoins de la station ; ainsi, si la BS peut fournir à une SS ses besoins supplémentaires de façon ponctuelle elle pourra le faire. Dans WiMAX, les flux non-contraints s'expriment avec une fréquence moindre, impliquant des délais d'accès plus grands.

4.3.3.2 Le débit utile

Notre solution permet d'avoir, en règle générale une meilleure utilisation du réseau lorsqu'elle est comparée à l'utilisation du réseau réalisée dans les mêmes conditions par l'algorithme d'ordonnancement de Freitag et. al. Les graphes présentant les résultats avec des durées de

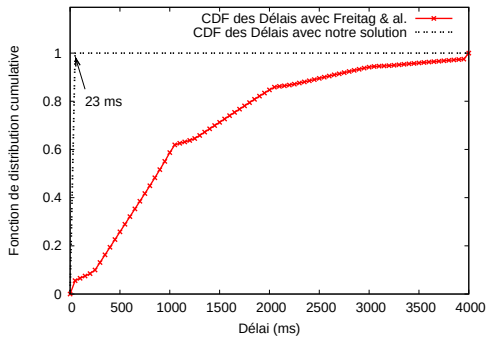

FIG. 4.14 – *Délais de nrtPS FS=5ms*

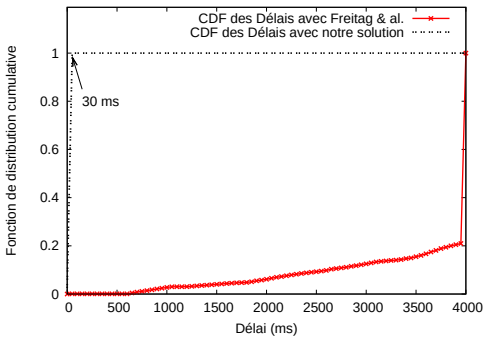

FIG. 4.15 – *Délais de BE FS=5ms*

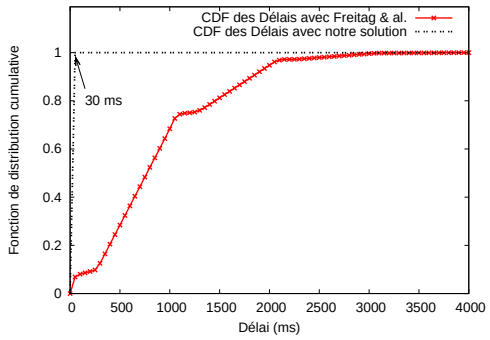

FIG. 4.16 – *Délais de nrtPS FS=10ms*

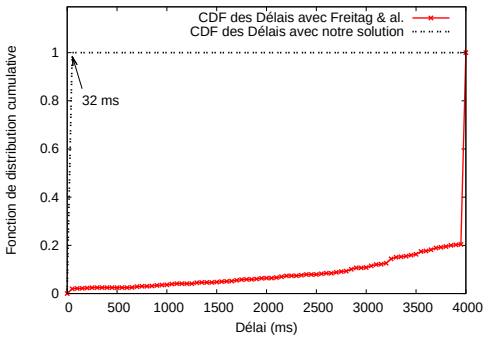

FIG. 4.17 – *Délais de BE FS=10ms*

supertrame allant de 5 ms à 20 ms (figures 4.24 à 4.35) le montrent clairement. Pour les flux à contraintes temporelles, les résultats sont équivalents. Le fonctionnement de l'algorithme de Freitag et. al. provoque un décalage en début de simulation faisant que les flux à contraintes temporelles ne sont pas tous immédiatement servis mais le seront au bout d'un certain temps. Un aspect intéressant des résultats de notre solution est la stabilité relative du service offert aux flux non contraints. Si nous comparons le service offert par notre solution à celui offert par l'algorithme de Freitag et. al. dans les figures 4.22,4.23,4.26,4.27, 4.30,4.31,4.34 et 4.35 ; nous observons un service plus stable offert par notre solution (le service est non seulement plus stable mais aussi plus important). Ceci est encore une fois dû au fait que nous donnons aux flux non-contraints la possibilité de s'exprimer sur leurs besoins de façon permanente, la BS est constamment au courant de l'état des files de toutes les SS lui permettant de les servir sans gaspiller les ressources. La signalisation proposée par WiMAX fait que la BS manque à des moments où elle peut se permettre de mieux servir certains flux non contraints des informations sur l'état des files de ces flux.

4.3.3.3 L'admissibilité des flux

Freitag et al. ne proposant pas de contrôle d'admission, nous analysons l'admissibilité de notre solution de façon générale. Notre solution se positionne dans un scénario où un nombre important de connexions sont simultanément actifs. Ceci implique qu'il ne faut pas que l'admissibilité des flux pose problème (surtout l'admissibilité des flux à contraintes temporelles). L'admissibilité dans notre solution est fortement reliée à la durée de la supertrame WiMAX. En effet, nous imposons une contrainte sur le délai de réponse aux requêtes des flux à

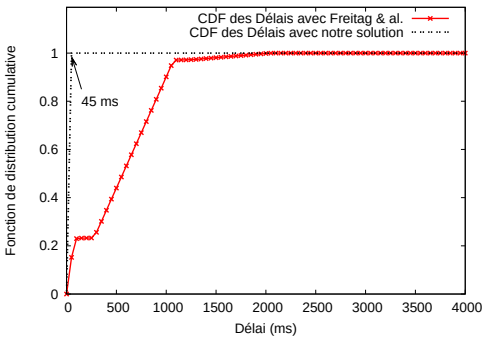

FIG. 4.18 – *Délais de nrtPS FS=20ms*

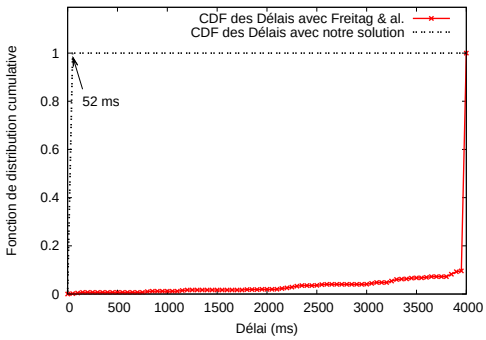

FIG. 4.19 – *Délais de BE FS=20ms*

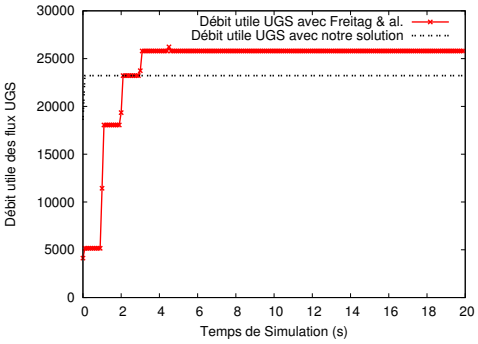

FIG. 4.20 – *Débit utile d'UGS FS=2ms*

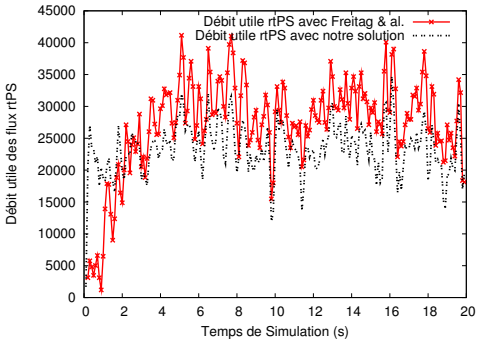

FIG. 4.21 – *Débit utile de rtPS FS=2ms*

contraintes temporelles. Cette contrainte permet d'avoir une garantie sur les délais d'accès au prix d'une admissibilité moindre de ce type de flux. Nous parvenons quand même à un taux d'admissibilité élevé. Les simulations que nous présentons plus haut voient un nombre de connexions de l'ordre de la centaine simultanément actifs au sein du réseau. Ceci n'est pas le cas pour les simulations où la durée de la trame est de 2 ms. En effet, dans ce scénario, notre algorithme refuse une connexion UGS ainsi que deux connexions rtPS afin de pouvoir assurer les garanties. Aucune connexion n'essuie de refus dans les scénarios avec une durée de supertrame plus grande. Notre algorithme de contrôle d'admission est moins contraint au niveau de la réservation de ressources lorsque la durée de la supertrame est plus grande. Il n'en reste pas moins vrai que, si l'on jette un regard critique sur les résultats, le contrôle d'admission est le point à améliorer de notre solution. Plusieurs pistes sont à explorer que nous exposerons dans la section suivante.

CONCLUSIONS ET PERSPECTIVES

Nous avons proposé, dans ce chapitre une modalité de gestion de la bande passante montante pour IEEE 802.16 WiMAX. Cette modalité a pour but de simplifier les procédures d'ordonnancement. La clef du fonctionnement de notre proposition réside en deux points :

- pour les flux à contraintes temporelles, une obligation d'allocation de la bande passante requise dans un délai faible ;
- pour l'ensemble des flux, une mise à jour périodique et fréquente de l'état des files des

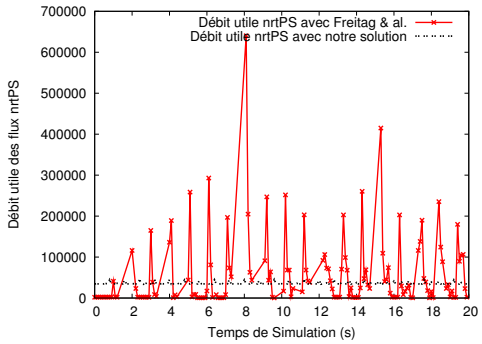

FIG. 4.22 – *Débit utile de nrtPS FS=2ms*

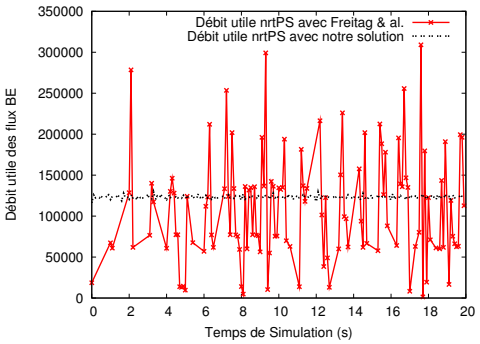

FIG. 4.23 – *Débit utile de BE FS=2ms*

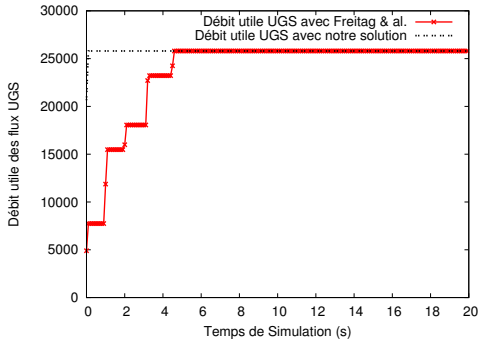

FIG. 4.24 – *Débit utile d'UGS FS=5ms*

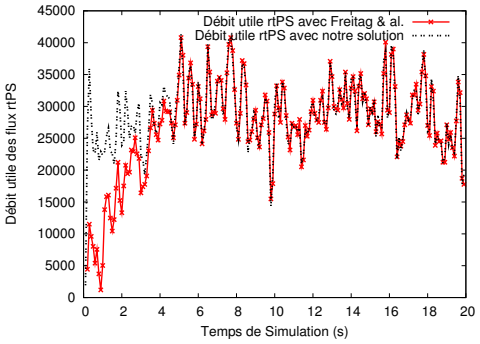

FIG. 4.25 – *Débit utile de rtPS FS=5ms*

SS transmise à la BS (grâce à la requête agrégée envoyée périodiquement) permettant à cette dernière de gérer au mieux les allocations.

Nous avons analysé notre proposition par simulation et l'avons comparé au fonctionnement standard de WiMAX couplé à un ordonnancement de la voie montante représentatif des ordonnancements par flux. Nous avons montré un certain nombre des propriétés de notre proposition : malgré sa simplicité, notre proposition donne aux flux à contraintes temporelles des garanties de délais faibles. Ces dernières dépendent directement de la durée de la supertrame WiMAX. Elle parvient aussi à servir correctement l'ensemble des autres flux admis. Les simulations ont montré que notre proposition est meilleure que la proposition de la littérature du point de vue de l'utilisation du réseau et des délais d'accès des flux à contraintes temporelles. Il y a cependant un point sur lequel notre proposition peut être critiquée : l'admissibilité des flux lorsque la taille de la supertrame est de 2 ms. En effet, avec cette durée de la supertrame, le fonctionnement de notre modalité va offrir des délais de transfert très faibles (comparés aux besoins) au prix d'une admissibilité moindre. Si ces faibles délais ne sont pas exigés (ce qui est généralement le cas), mais qu'il y a une contrainte sur la durée de la supertrame (à 2 ms), nous avons exploré la possibilité de relâcher la contrainte sur les délais d'allocations qui, au lieu d'être toutes les supertrames, peut couvrir plusieurs supertrames. Au lieu d'avoir l'allocation des octets respectant le contrat dans la supertrame suivant la requête, cette allocation peut s'étendre sur plusieurs supertrames. Il est important cependant de noter que, dans ce cas, les délais moyens seront plus importants (mais resteront dans l'ordre que nous préconisons) et la complexité des algorithmes d'ordonnancement sera elle aussi plus importante. Pour des durées de supertrame élevées (20 ms), nous avons observé

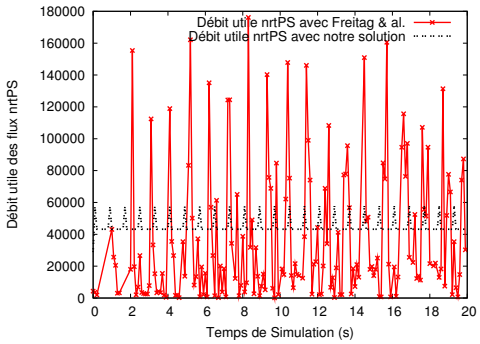

FIG. 4.26 – *Débit utile de nrtPS FS=5ms*

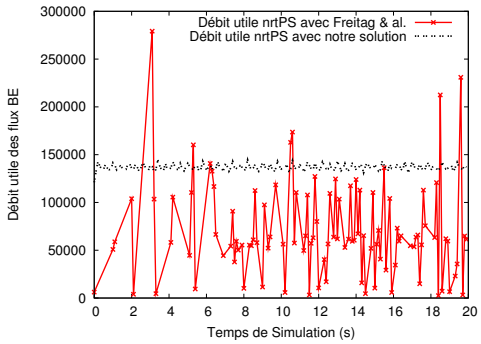

FIG. 4.27 – *Débit utile de BE FS=5ms*

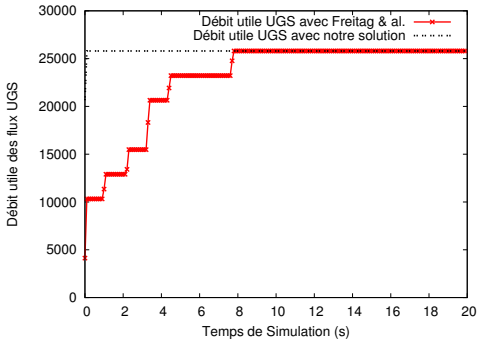

FIG. 4.28 – *Débit utile d'UGS FS=10ms*

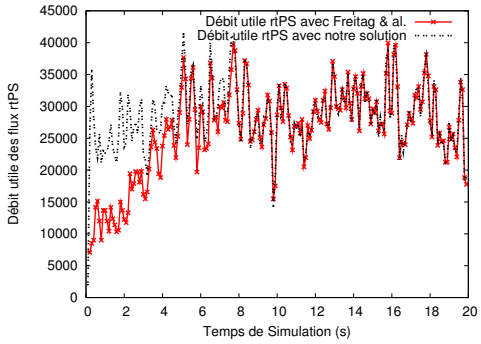

FIG. 4.29 – *Débit utile de rtPS FS=10ms*

des délais qui peuvent atteindre la trentaine de ms. Ces délais dépassent l'objectif souhaité. Nous préconisons dans ce cas d'employer une des politiques d'anticipation que nous avons présentées afin de réduire les délais moyens d'accès.

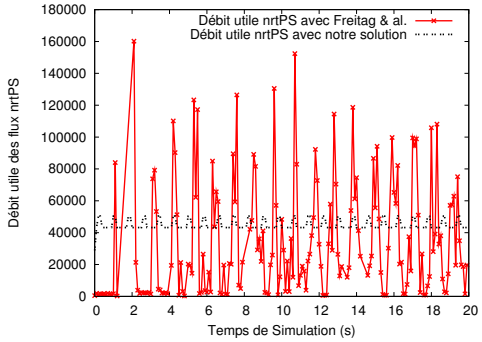

FIG. 4.30 – *Débit utile de nrtPS FS=10ms*

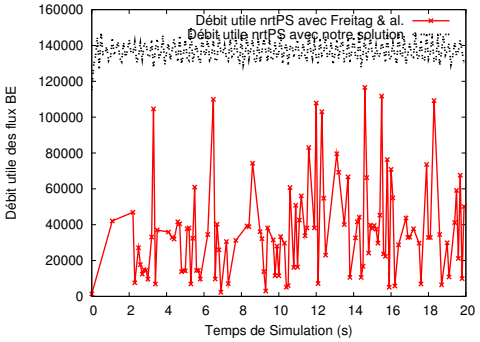

FIG. 4.31 – *Débit utile de BE FS=10ms*

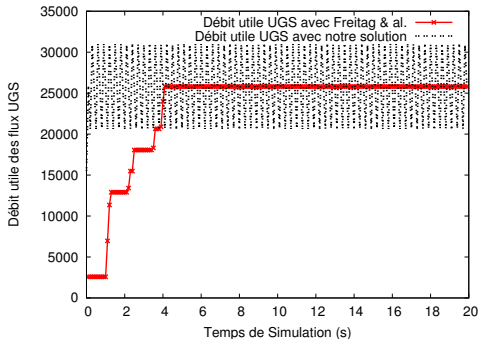

FIG. 4.32 – *Débit utile d'UGS FS=20ms*

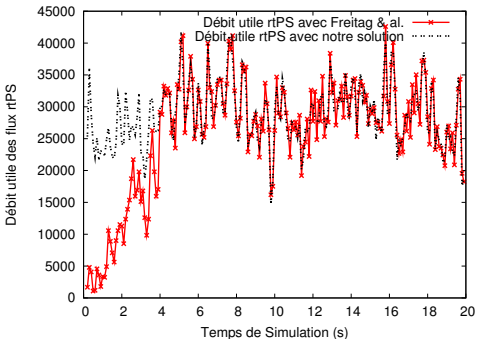

FIG. 4.33 – *Débit utile de rtPS FS=20ms*

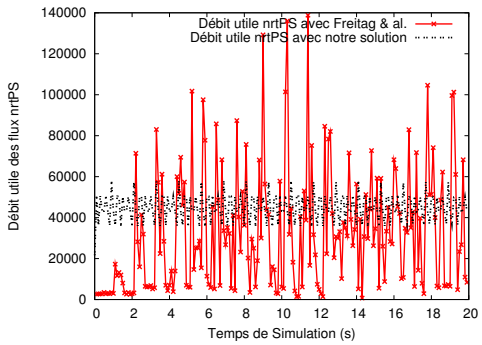

FIG. 4.34 – *Débit utile de nrtPS FS=20ms*

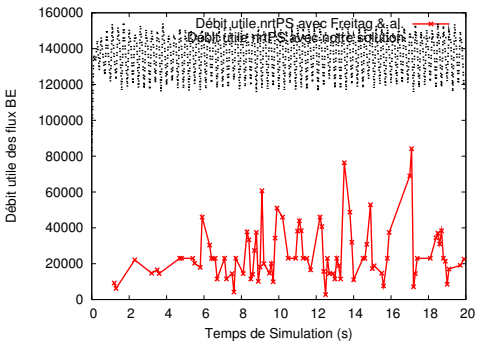

FIG. 4.35 – *Débit utile de BE FS=20ms*

Conclusion générale

Jacob : It only ends once.
Anything that happens before
that is just progress.
"Lost" saison 5, épisode 16

Les technologies sans fil sont devenues, ces dernières années, un phénomène de société. Le sentiment de liberté qui accompagne l'utilisation de telles technologies les ont rendues attrayantes aux yeux du grand public. Ainsi, les technologies sans-fil ont pris une place importante dans le quotidien des sociétés développées ou en voie de développement. D'un autre côté, les technologies sans-fil, de par la facilité de leur mise en place et la flexibilité qu'elles offrent, ont trouvé aussi leur chemin vers les milieux industriels et sont de plus en plus utilisées dans certains systèmes de commande pour les réseaux industriels. Cependant, ces technologies ne répondent pas tout à fait aux exigences en termes de Qualité de Service des applications et des utilisateurs qu'elles servent ; elles représentaient un des maillons faibles, en terme de Qualité de Service, dans une chaîne de bout en bout. Notre travail s'est donc inscrit dans cette problématique : contribuer aux mécanismes de Qualité de Service dans les réseaux d'accès sans fil.

Bilan

Les contributions de ce travail se sont articulées autour des deux technologies sans fil que nous avons étudiées : WiFi et WiMAX. Le but étant de proposer, plus particulièrement pour ces deux technologies, des mécanismes de gestion de la bande passante permettant d'améliorer la qualité de service.

Pour WiFi

Nous avons proposé un algorithme de contrôle d'admission pour IEEE 802.11 EDCA. L'algorithme de contrôle d'admission que nous proposons est de type hybride : c'est donc un algorithme qui base la décision sur des mesures faites de l'état du réseau ainsi que sur un modèle analytique représentant le schéma d'accès. Notre algorithme utilise, dans son fonctionnement, un modèle du schéma d'accès EDCA que nous avons développé. Notre modèle se distingue des modèles existants par la prise en compte explicite du mécanisme de gestion des collisions virtuelles d'EDCA. Ceci est fondamental dans le cas de figure où nous nous plaçons : un réseau en mode infrastructure où le point d'accès sert de passerelle entre les stations du réseau sans fil et le réseau Internet. En effet, un trafic conséquent va dans ce cas traverser le point d'accès et des collisions virtuelles sont susceptibles de se produire. Le modèle développé se présente sous deux formes : une générale prenant en compte l'ensemble des caractéristiques du schéma d'accès ; l'autre réduite, équivalente à la forme générale en terme de représentation du schéma d'accès, et en même temps plus simple et plus facile à prendre en main. L'algorithme de contrôle d'admission a été évalué par simulation sous ns-2, mis en œuvre puis intégré dans une solution de QoS de bout en bout (EuQoS). Il doit maintenant être évalué dans le cadre de cette solution.

Pour WiMAX

Nous nous sommes intéressés, pour WiMAX, au système d'ordonnancement de la voie montante dans le but de proposer une implémentation des classes de services WiMAX (UGS, rtPS, nrtPS et BE) pour un réseau WiMAX jouant le rôle de lien *backhaul*. Dans ce cas de figure, le problème de mise à l'échelle est à prendre en compte dans la conception du système d'ordonnancement. Nous avons proposé dans ce cadre une modalité simple de gestion de la bande passante montante pour WiMAX. Cette modalité couvre les principaux mécanismes nécessaires à la fourniture d'une Qualité de Service pour WiMAX : nous avons défini un algorithme de contrôle d'admission, un nouveau système de requête de bande passante et les algorithmes d'ordonnancement conséquents au niveau de la BS et de la SS. Notre proposition se distingue des solutions proposées dans la littérature par :

- Des requêtes de bande passante agrégée à l'échelle de la station (par opposition aux requêtes individuelles des flux considérées par les solutions existantes). Ces requêtes sont communiquées à la BS fréquemment.
- Une gestion des allocations de la bande passante au niveau de la BS par station (par opposition à une gestion par flux des solutions existantes).

Par l'introduction de l'agrégation, notre solution se distingue par sa simplicité et une meilleure résistance à l'échelle. En plus, notre solution offre la possibilité de garantir aux flux à contraintes temporelles des propriétés de QoS contraignantes du point de vue des délais et de la gigue tout en garantissant une meilleure utilisation du réseau. Cette meilleure utilisation est obtenue grâce à une meilleure expression des besoins en bande passante et une meilleure gestion des flux nrtPS et BE. Ces propriétés intéressantes ont pu être observées suite à une analyse comparative par simulation de notre proposition à des solutions conventionnelles.

Perspectives

Plusieurs pistes méritent d'être explorées, suite au travail que nous venons de présenter.

En plus des perspectives sur le court terme que nous avons présentées dans les conclusions des chapitres, d'autres points doivent être étudiés parmi lesquelles :

- Nos travaux ont supposé une certaine stabilité de l'environnement radio dans lesquels les stations opèrent et n'ont pas explicitement considéré de changement de mode physique occasionnant un changement du débit physique. Il est à noter que ces changements peuvent être pris en compte facilement par nos solutions en jouant sur des paramètres des modèles par exemple. Il est néanmoins nécessaire d'approfondir la réflexion à ce niveau.
- Tel qu'énoncé dans l'introduction, un cadre du travail effectué est les réseaux hétérogènes sans fil WiMAX-WiFi. Il faudrait étudier la combinaison du contrôle d'admission sur WiFi et de la modalité de gestion de bande passante sur WiMAX dans un réseau hétérogène sans fil. La réflexion devra porter sur les correspondances (*mapping*) à faire pour passer d'un réseau à l'autre. Il faut aussi s'intéresser aux conséquences qu'impliquent l'intégration de nos solutions dans une architecture de QoS de bout en bout.

Sur un plus long terme, plusieurs directions intéressantes peuvent être explorées :

- Le *cross-layering* : afin d'offrir la QoS, nous avons la conviction qu'il est fondamental de coordonner les actions de la couche MAC et PHY voire des couches supérieures. Il serait donc intéressant d'explorer, dans quelles mesures nos propositions peuvent exploiter

les informations en provenance de la couche physique et dans quelles mesures elles pourraient exploiter d'éventuelles configurations de la couche physique.

– Les réseaux sans fil ont une autre caractéristique que nous avons occultée pendant ces travaux : la mobilité. En effet, dans des applications où les utilisateurs sont mobiles (même à grande vitesse pour les réseaux véhiculaires), il faut réfléchir aux moyens à employer pour qu'un utilisateur passant d'une cellule à une autre puisse continuer à bénéficier d'une qualité de service satisfaisante.

Sur un point moins technique mais tout aussi important et au risque de paraître crédule et un peu cliché, je ne peux m'empêcher de citer Rabelais : "Science sans conscience n'est que ruine de l'âme". Lorsque en introduction, je parlais de l'effet des technologies sur lesquelles nous travaillons sur la société, ou sur la physiologie de l'humain, ce n'était pas banal. La technologie continuera à évoluer, mais il faut garder un esprit critique face à ces avancées technologiques qui certes sont d'un grand service pour l'être humain, mais qui peuvent s'avérer dangereuses pour l'Homme et pour son environnement.

Calcul de la matrice des probabilités en régime stationnaire

A

Sommaire

Cette annexe donne les différentes étapes nécessaires au calcul de la matrice des probabilités en régime stationnaire de la chaîne de Markov du comportement d'une catégorie d'accès présentée au chapitre 3.

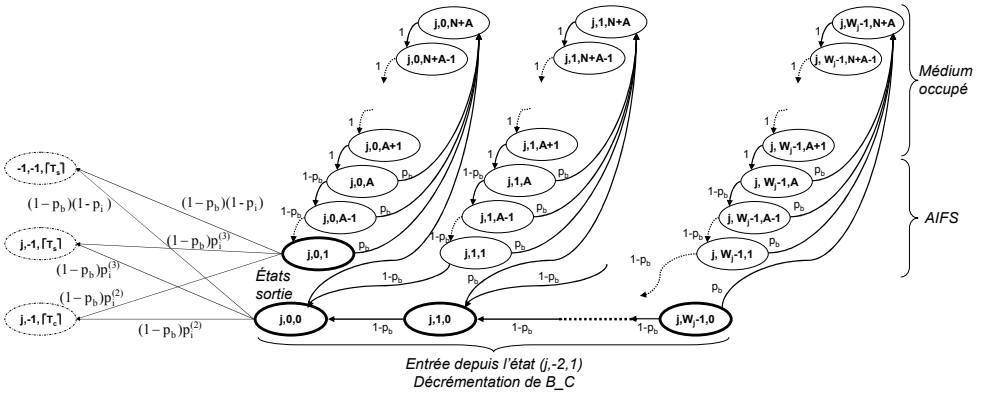

FIG. A.1 – *Bloc Backoff :* $0 \leq j \leq m + h$

La méthode que nous employons consiste à exprimer les probabilités des états en fonction de l'un d'entre eux (en exploitant les probabilités de transitions entre états). Cette expression est possible puisque le modèle général est reinitialisable.
Nous notons $b_{j,k,d}$ la probabilité de l'état (j, k, d). La probabilité de l'état en fonction de laquelle nous écrivons les probabilités des autres états est la probabilité $b_{0,-2,1}$. Ce choix est arbitraire, il correspond cependant au premier état à partir duquel une tentative d'accès est possible, le choix nous paraît donc pertinent.
Les probabilités des états en régime stationnaire (qui sont celles que nous recherchons) sont ensuite obtenues en écrivant l'équation d'harmonie (la somme des probabilités de tous les états vaut 1).
Cette méthode est réalisée en exploitant les blocs de base du modèle (AIFS, Backoff, résultats de la tentative d'accès). Nous écrivons en premier lieu les probabilités par bloc et par étage de backoff en fonction de $b_{j,-2,1}$ avant d'écrire cette probabilité en fonction de $b_{0,-2,1}$.

A.1 Le bloc backoff

Nous rappelons dans la figure A.1 le modèle du bloc backoff. Nous pouvons écrire : $\forall j, 0 \leq j \leq m + h \quad \forall k, 0 \leq k \leq W_j - 1$

Pour les états où le médium est occupé par une autre AC pendant le Backoff, les probabilités de transitions valent 1 :

$$b_{j,k,d} = b_{j,k,N+A} \quad A \leq d \leq N + A$$

Pour la période AIFS suite à l'occupation ci-dessus, les probabilités des transitions entre ces états et l'état (j, k, A) sont $(1 - p_b)$, nous avons donc :

$$b_{j,k,d} = (1 - p_b)^{A-d} b_{j,k,A} \quad 1 \leq d \leq A \tag{A.1}$$

Nous écrivons ensuite la probabilité de l'état $(j,k,N+A)$, en fonction des probabilités de ses états prédécesseurs (avec une probabilité d'occupation p_b pour les transitions) :

$$\begin{aligned} b_{j,k,N+A} &= p_b \sum_{d=0}^{A} b_{j,k,d} \\ &\quad \text{en séparant et intégrant l'équation A.1} \\ &= p_b \sum_{d=1}^{A} \left((1-p_b)^{A-d} b_{j,k,N+A} \right) + p_b b_{j,k,0} \\ &= p_b \frac{1-(1-p_b)^A}{p_b} b_{j,k,N+A} + p_b b_{j,k,0} \\ &= \left(1-(1-p_b)^A \right) b_{j,k,N+A} + p_b b_{j,k,0} \end{aligned}$$

donc :

$$b_{j,k,N+A} = \frac{p_b}{(1-p_b)^A} b_{j,k,0} \quad \forall j, 0 \leq j \leq m+h \tag{A.2}$$

(A.3)

Nous avons aussi, pour l'état de décrémentation de backoff $(j, W_j - 1, 0)$ (ayant pour seule prédécesseur $(j,-2,1)$ avec une probabilité $\frac{1-p_b}{W_j+1}$) :

$$b_{j,W_j-1,0} = \frac{1-p_b}{W_j+1} b_{j,-2,1}$$

Les autres états de décrémentation de backoff ont, en plus de l'état $(j,-2,1)$, deux prédécesseurs provenant du décompte de backoff (avec une probabilité $(1-p_b)$:

$$b_{j,k,0} = (1-p_b)(b_{j,k+1,1} + b_{j,k+1,0}) + \frac{1-p_b}{W_j+1} b_{j,-2,1} \forall k \neq W_j - 1$$

La somme $b_{j,k,0} + b_{j,k,1}$ peut être simplifiée en utilisant les équations A.1 et A.2 :

$$\begin{aligned} b_{j,k,0} + b_{j,k,1} &= (1-p_b)^{A-1} b_{j,k,N+A} + b_{j,k,0} \\ &= (1-p_b)^{A-1} \frac{p_b}{(1-p_b)^A} b_{j,k,0} + b_{j,k,0} \\ &= \frac{1}{1-p_b} b_{j,k,0} \end{aligned} \tag{A.4}$$

Revenons à $b_{j,k,0} \forall k \neq W_j - 1$ en intégrant l'équation A.4 :

$$b_{j,k,0} = b_{j,k+1,0} + \frac{1-p_b}{W_j+1} b_{j,-2,1}$$

Nous pouvons donc généraliser $b_{j,k,0} \forall k, 0 \leq k \leq W_j - 1$

$$\begin{aligned} b_{j,k,0} &= b_{j,W_j-1,0} + (W_j - 1 - k) \frac{1-p_b}{W_j+1} b_{j,-2,1} \\ &= \frac{1-p_b}{W_j+1} b_{j,-2,1} + (W_j - 1 - k) \frac{1-p_b}{W_j+1} b_{j,-2,1} \\ &= \frac{W_j - k}{W_j+1} (1-p_b) b_{j,-2,1} \end{aligned}$$

Récapitulons donc :

$$b_{j,k,0} = \frac{W_j - k}{W_j + 1}(1 - p_b)b_{j,-2,1}, \quad \forall j, \forall k \quad \textit{Les états de décompte} \tag{A.5}$$

$$\begin{aligned} b_{j,k,d} &= \frac{p_b}{(1-p_b)^A} b_{j,k,0}, \forall j, \quad \forall k, A \leq d \leq N + A \\ &= \frac{p_b}{(1-p_b)^{A-1}} \frac{W_j - k}{W_j + 1} b_{j,-2,1} \quad \textit{Les états "médium occupé"} \end{aligned} \tag{A.6}$$

$$\begin{aligned} b_{j,k,d} &= (1-p_b)^{A-d} b_{j,k,A}, \forall j, \quad \forall k, 1 \leq d \leq A \\ &= \frac{p_b}{(1-p_b)^{d-1}} \frac{W_j - k}{W_j + 1} b_{j,-2,1} \quad \text{Les états de la décompte d'AIFS} \end{aligned} \tag{A.7}$$

$$\tag{A.8}$$

Nous laissons ce bloc à ce stade, nous y reviendrons ultérieurement.

A.2 Les blocs "Résultats d'une tentative de transmission"

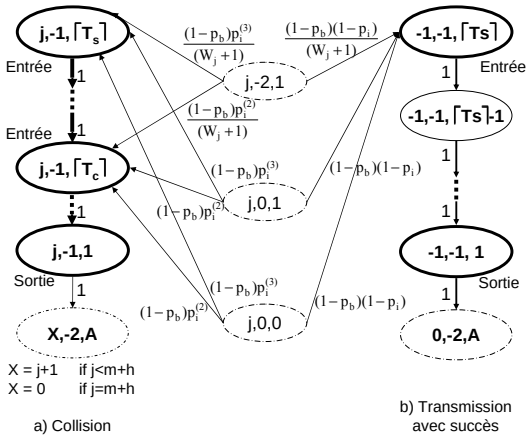

FIG. A.2 – *Résultats d'une tentative effective de transmission :* $0 \leq j \leq m + h$

Nous rappelons le modèles des blocs "résultats d'une tentative de transmission".

A.2.1 Le bloc transmission avec succès

Nous pouvons simplement écrire (grâce aux transitions avec une probabilité de 1) :

$$b_{-1,-1,d} = b_{-1,-1,\lceil T_s \rceil}, \quad \forall d, 1 \leq d \leq \lceil T_s \rceil$$

L'état $(-1, -1, \lceil T_s \rceil)$ est précédé par tous les états où une tentative d'accès est possible, avec la probabilité de transition correspondant à une transmission avec succès. Nous faisons la

somme sur tous les étages de Backoff :

$$b_{-1,-1,\lceil T_s \rceil} = \sum_{j=0}^{m+h} \left[\frac{(1-p_b)(1-p_i)}{W_j+1} b_{j,-2,1} + (1-p_b)(1-p_i)(b_{j,0,1}+b_{j,0,0}) \right]$$

puis en intégrant le résultat A.5

$$= \sum_{j=0}^{m+h} \left[\frac{(1-p_b)(1-p_i)b_{j,-2,1} + (1-p_b)(1-p_i)W_j b_{j,-2,1}}{W_j+1} \right]$$

$$= (1-p_b)(1-p_i) \sum_{j=0}^{m+h} \frac{1+W_j}{W_j+1} b_{j,-2,1}$$

$$= (1-p_b)(1-p_i) \sum_{j=0}^{m+h} b_{j,-2,1}$$

Récapitulons donc :

$$b_{-1,-1,d} = (1-p_b)(1-p_i) \sum_{j=0}^{m+h} b_{j,-2,1}, \quad \forall d, 1 \leq d \leq \lceil T_s \rceil$$

Nous laissons ce bloc à ce stade, nous y reviendrons ultérieurement.

A.3 Le bloc collision

Là encore, nous pouvons simplement écrire (grâce aux probabilités à 1 des transitions) :

$$b_{j,-1,d} = b_{j,-1,\lceil T_s \rceil} \quad \forall j, 0 \leq j \leq m+h, \lceil T_c \rceil < d \leq \lceil T_s \rceil$$

et :

$$b_{j,-1,d} = b_{j,-1,\lceil T_c \rceil} \quad \forall j, 0 \leq j \leq m+h, 1 \leq d \leq \lceil T_c \rceil$$

Nous intégrons les probabilités des différents types de collisions :

$$b_{j,-1,\lceil T_s \rceil} = \frac{(1-p_b)p_i^{(3)}}{W_j+1} b_{j,-2,1} + (1-p_b)p_i^{(3)}(b_{j,0,1}+b_{j,0,0})$$

et :

$$b_{j,-1,\lceil T_c \rceil} = \frac{(1-p_b)p_i^{(2)}}{W_j+1} b_{j,-2,1} + (1-p_b)p_i^{(2)}(b_{j,0,1}+b_{j,0,0}) + b_{j,-1,\lceil T_s \rceil}$$

$$= \frac{(1-p_b)p_i}{W_j+1} b_{j,-2,1} + (1-p_b)p_i(b_{j,0,1}+b_{j,0,0})$$

De la même façon que pour le bloc Transmission avec succès, nous avons :

$$b_{j,-1,\lceil T_s \rceil} = (1-p_b)p_i^{(3)} \left[\frac{1}{W_j+1} b_{j,-2,1} + \frac{1}{1-p_b}(1-p_b)\frac{W_j}{W_j+1} b_{j,-2,1} \right]$$

$$= (1-p_b)p_i^{(3)} b_{j,-2,1}$$

et :

$$b_{j,-1,\lceil T_c \rceil} = (1-p_b)p_i b_{j,-2,1}$$

Récapitulons :

$$\forall j, 0 \leq j \leq m+h$$
$$b_{j,-1,d} = (1-p_b)p_i^{(3)} b_{j,-2,1} \quad \text{pour} \quad \lceil T_c \rceil + 1 \leq d \leq \lceil T_s \rceil$$
$$b_{j,-1,d} = (1-p_b)p_i b_{j,-2,1} \quad \text{pour} \quad 1 \leq d \leq \lceil T_c \rceil$$

Nous laissons ce bloc à ce stade, nous y reviendrons ultérieurement.

A.4 Le bloc AIFS

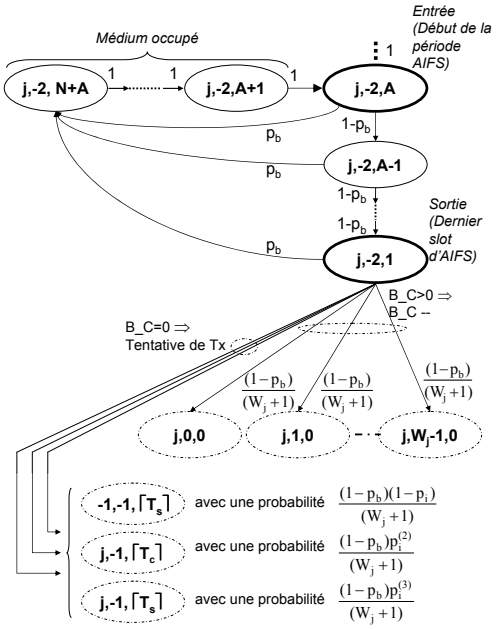

FIG. A.3 – *Bloc AIFS :* $0 \leq j \leq m+h$

Nous rappelons dans la figure A.3 le bloc AIFS du modèle. Nous pouvons écrire, comme pour la partie AIFS du bloc Backoff $\forall j, 0 \leq j \leq m+h$:

$$b_{j,-2,d} = b_{j,-2,N+A} \quad \text{pour} \quad A+1 \leq d \leq N+A$$

Pour les états de décompte d'AIFS (avec une probabilité $(1-p_b)$) nous avons :

$$b_{j,-2,d} = (1-p_b)^{A-d} b_{j,-2,A} \quad \text{pour} \quad 1 \leq d \leq A$$

Nous avons aussi :

$$b_{j,-2,N+A} = \sum_{d=1}^{A} p_b b_{j,-2,d}$$

est la somme des passages à un état occupé du médium

$$= p_b \sum_{d=1}^{A} (1-p_b)A - d b_{j,-2,A}$$
$$= p_b b_{j,-2,A} \frac{1-(1-p_b)^A}{p_b}$$
$$= (1-(1-p_b)^A) b_{j,-2,A}$$

et en faisant le lien avec l'étage de backoff précédent, nous avons, pour l'état $(j, -2, A)$:

$$\begin{aligned} b_{j,-2,A} &= b_{j,-2,N+A} + b_{j-1,-1,1}, \forall j \neq 0 \\ &= (1-(1-p_b)^A) b_{j,-2,A} + b_{j-1,-1,1} \\ &= \frac{1}{(1-p_b)^A} b_{j-1,-1,1} \end{aligned}$$

Sachant que :

$$b_{j,-2,1} = (1-p_b)^{A-1} b_{j,-2,A}$$

nous avons :

$$b_{j,-2,A} = \frac{1}{(1-p_b)^{A-1}} b_{j,-2,1}$$

et :

$$\begin{aligned} b_{j,-2,1} &= (1-p_b)^{A-1} \frac{1}{(1-p_b)^A} b_{j-1,-1,1} \\ &= \frac{1}{1-p_b} b_{j-1,-1,1} \\ &= p_i b_{j-1,-2,1} \end{aligned}$$

donc :

$$b_{j,-2,1} = p_i^j b_{0,-2,1}$$

Ce dernier résultat nous permettra de connecter chaque étage de backoff au premier ($j = 0$). Nous avons aussi pour $j = 0$:

$$\begin{aligned} b_{0,-2,1} &= (1-p_b)^{A-1} b_{0,-2,A} \\ b_{0,-2,A} &= \frac{1}{(1-p_b)^{A-1}} b_{0,-2,A} \end{aligned}$$

Récapitulons donc :

$$\forall j, 0 \leq j \leq m+h$$

$$\begin{aligned} b_{j,-2,d} &= b_{j,-2,N+A} \quad \text{pour} \quad A+1 \leq d \leq N+A \\ &= (1-(1-p_b)^A) b_{j,-2,A} \\ &= \frac{(1-(1-p_b)^A)}{(1-p_b)^{A-1}} b_{j,-2,1} \\ b_{j,-2,d} &= \frac{(1-(1-p_b)^A)}{(1-p_b)^{A-1}} p_i^j b_{0,-2,1} \quad \text{Pour les états d'occupation du médium} \end{aligned}$$

$$\begin{aligned} b_{j,-2,A} &= \frac{1}{(1-p_b)^{A-1}} b_{j,-2,1} \\ b_{j,-2,A} &= \frac{1}{(1-p_b)^{A-1}} p_i^j b_{0,-2,1} \quad \text{Pour le premier état AIFS} \end{aligned}$$

$$\begin{aligned} b_{j,-2,d} &= (1-p_b)^{A-d} b_{j,-2,A} \quad \text{pour} \quad 1 \leq d \leq A \\ &= \frac{(1-p_b)^{A-d}}{(1-p_b)^{A-1}} b_{j,-2,1} \\ b_{j,-2,d} &= \frac{1}{(1-p_b)^{d-1}} p_i^j b_{0,-2,1} \quad \text{Pour les états de décompte d'AIFS} \end{aligned}$$

$$b_{j,-2,1} = p_i^j b_{0,-2,1}$$

A.5 Récapitulatif total

En utilisant la dernière formule obtenue pour le bloc AIFS ($b_{j,-2,1} = p_i^j b_{0,-2,1}$), nous pouvons généraliser les formules des probabilités de tous les états en fonction de $b_{0,-2,1}$. Nous obtenons donc :

A.5.1 Le bloc backoff

$$\begin{aligned} b_{j,k,0} &= \frac{W_j - k}{W_j + 1}(1-p_b) p_i^j b_{0,-2,1}, \quad \forall j, \forall k \\ b_{j,k,d} &= \frac{W_j - k}{W_j + 1} \frac{p_b}{(1-p_b)^{d-1}} p_i^j b_{0,-2,1}, \quad \forall j, \quad \forall k, \quad 1 \leq d \leq A \\ b_{j,k,d} &= \frac{W_j - k}{W_j + 1} \frac{p_b}{(1-p_b)^{A-1}} p_i^j b_{0,-2,1}, \quad \forall j, \quad \forall k, \quad A \leq d \leq N+A \end{aligned}$$

A.5.2 Le bloc transmission avec succès

$$\begin{aligned} b_{-1,-1,d} &= (1-p_b)(1-p_i)\sum_{j=0}^{m+h} p_i^j b_{0,-2,1}, \quad \forall d, 1 \leq d \leq \lceil T_s \rceil \\ &= (1-p_b)(1-p_i) b_{0,-2,1} \sum_{j=0}^{m+h} p_i^j \\ &= (1-p_b)(1-p_i) b_{0,-2,1} \frac{1-p_i^{m+h+1}}{1-p_i} \\ b_{-1,-1,d} &= (1-p_b)(1-p_i^{m+h+1}) b_{0,-2,1} \end{aligned}$$

A.5.3 Le bloc collision

$$\begin{aligned} &\forall j, 0 \leq j \leq m+h \\ &\qquad b_{j,-1,d} = (1-p_b) p_i^{(3)} p_i^j b_{0,-2,1} \quad \text{pour} \quad \lceil T_c \rceil + 1 \leq d \leq \lceil T_s \rceil \\ &\qquad b_{j,-1,d} = (1-p_b) p_i p_i^j b_{0,-2,1} \quad \text{pour} \quad 1 \leq d \leq \lceil T_c \rceil \end{aligned}$$

A.5.4 Le bloc AIFS

$$\begin{aligned} &\forall j, 0 \leq j \leq m+h \\ &\qquad b_{j,-2,d} = \frac{(1-(1-p_b)^A)}{(1-p_b)^{A-1}} p_i^j b_{0,-2,1} \quad \text{pour} \quad A+1 \leq d \leq N+A \\ &\qquad b_{j,-2,d} = \frac{1}{(1-p_b)^{d-1}} p_i^j b_{0,-2,1} \quad \text{pour} \quad 1 \leq d \leq A \end{aligned}$$

A.6 Équation d'harmonie

Le système ainsi obtenu n'est pas solvable. Nous utilisons pour le résoudre l'équation d'harmonie de la chaîne de Markov (la somme des probabilités en régime stationnaire de tous les états vaut 1).

Nous avons donc :

$$1 = \sum_{j=0}^{m+h} \left[\sum_{k=0}^{W_j - 1} \left(\sum_{d=0}^{N+A} b_{j,k,d} \right) + \sum_{d=1}^{N+A} (b_{j,-2,d}) + \sum_{d=1}^{\lceil T_s \rceil} (b_{j,-1,d}) \right] + \sum_{d=1}^{\lceil T_s \rceil} (b_{-1,-1,d})$$

Afin de faciliter la lecture et la compréhension du calcul, nous procéderons par bloc, c'est à dire que pour chaque bloc nous effectuerons la somme de tous les états possibles avant de donner la somme totale. Dans ces sommes, nous remplaçons la probabilité de chacun des états par sa formule obtenue en fonction de $b_{0,-2,1}$.

A.6.1 Le bloc backoff

La somme de tous les états de backoff donne le développement suivant :

$$\begin{aligned}
\sum_{j=0}^{m+h}\left[\sum_{k=0}^{W_j-1}\left(\sum_{d=0}^{N+A} b_{j,k,d}\right)\right] &= \sum_{j=0}^{m+h}\left[\sum_{k=0}^{W_j-1}\left(\frac{W_j-k}{W_j+1}(1-p_b)p_i^j b_{0,-2,1}\right.\right.\\
&\quad + \sum_{d=1}^{A}\frac{W_j-k}{W_j+1}p_b p_i^j b_{0,-2,1}\frac{1}{(1-p_b)^{d-1}}\\
&\quad \left.\left. + \sum_{d=A+1}^{N+A}\frac{W_j-k}{W_j+1}p_b p_i^j b_{0,-2,1}\frac{1}{(1-p_b)^{A-1}}\right)\right]\\
&= \sum_{j=0}^{m+h}\left[\sum_{k=0}^{W_j-1}\left(\left(\frac{W_j-k}{W_j+1}p_i^j b_{0,-2,1}\right)\right.\right.\\
&\quad \left.\left.\left((1-p_b)+\frac{1-(1-p_b)^A}{(1-p_b)^{A-1}}+\frac{Np_b}{(1-p_b)^{A-1}}\right)\right)\right]\\
&= \frac{1+Np_b}{(1-p_b)^{A-1}}b_{0,-2,1}\sum_{j=0}^{m+h}\left[\frac{p_i^j}{W_j+1}\sum_{k=0}^{W_j-1}(W_j-k)\right]\\
&= \frac{1+Np_b}{(1-p_b)^{A-1}}b_{0,-2,1}\sum_{j=0}^{m+h}\left[\frac{p_i^j}{W_j+1}\frac{W_j(W_j+1)}{2}\right]\\
&= \frac{1+Np_b}{2(1-p_b)^{A-1}}b_{0,-2,1}\sum_{j=0}^{m+h}p_i^j W_j
\end{aligned}$$

A.6.2 Le bloc AIFS

La somme de tous les états AIFS donne le développement suivant :

$$\begin{aligned}
\sum_{j=0}^{m+h}\left[\sum_{d=1}^{N+A} b_{j,-2,d}\right] &= \sum_{j=0}^{m+h}\left[\sum_{d=1}^{A}\left(\frac{1}{(1-p_b)^{d-1}}p_i^j b_{0,-2,1}\right)\right.\\
&\quad \left. + \sum_{d=A+1}^{N+A}\left(\frac{1}{(1-p_b)^{A-1}}p_i^j b_{0,-2,1}\right)\right]\\
&= \sum_{j=0}^{m+h}\left[p_i^j b_{0,-2,1}\left(\frac{1-(1-p_b)^A}{p_b(1-p_b)^{A-1}}+\frac{N(1-(1-p_b)^A)}{(1-p_b)^{A-1}}\right)\right]\\
&= \frac{(1-(1-p_b)^A)(1+Np_b)}{p_b(1-p_b)^{A-1}}\frac{1-p_i^{m+h+1}}{1-p_i}b_{0,-2,1}
\end{aligned}$$

A.6.3 Le bloc collision

La somme de tous les états de collision nous donne :

$$\sum_{j=0}^{m+h}\left[\sum_{d=1}^{\lceil T_s\rceil} b_{j,-1,d}\right] = \sum_{j=0}^{m+h}\left[\sum_{d=1}^{\lceil T_c\rceil}\left((1-p_b)p_i^{j+1}b_{0,-2,1}\right)\right.$$
$$\left.+\sum_{d=\lceil T_c\rceil+1}^{\lceil T_s\rceil}\left((1-p_b)p_i^{j}p_i^{(3)}b_{0,-2,1}\right)\right]$$

$$= (1-p_b)\left(\lceil T_c\rceil p_i + (\lceil T_s\rceil - \lceil T_c\rceil)p_i^{(3)}\right)\frac{1-p_i^{m+h+1}}{1-p_i}b_{0,-2,1}$$

A.6.4 Le bloc Transmission avec succès

La somme de tous les états de transmission avec succès donne :

$$\sum_{d=1}^{\lceil T_s\rceil} b_{-1,-1,d} = \lceil T_s\rceil(1-p_b)(1-p_i^{m+h+1})b_{0,-2,1}$$

A.6.5 La somme générale

Nous obtenons donc, en sommant toutes les sommes précédentes :

$$\begin{aligned}1 = {} & \frac{1+Np_b}{2(1-p_b)^{A-1}}b_{0,-2,1}\sum_{j=0}^{m+h}p_i^jW_j \\ & + \frac{(1-(1-p_b)^A)(1+Np_b)}{p_b(1-p_b)^{A-1}}\frac{1-p_i^{m+h+1}}{1-p_i}b_{0,-2,1} \\ & + (1-p_b)\left(\lceil T_c\rceil p_i + (\lceil T_s\rceil - \lceil T_c\rceil)p_i^{(3)}\right)\frac{1-p_i^{m+h+1}}{1-p_i}b_{0,-2,1} \\ & + \lceil T_s\rceil(1-p_b)(1-p_i^{m+h+1})b_{0,-2,1}\end{aligned}$$

$$\begin{aligned}b_{0,-2,1} = {} & \left[\frac{1+Np_b}{2(1-p_b)^{A-1}}\sum_{j=0}^{m+h}p_i^jW_j\right. \\ & + \frac{(1-(1-p_b)^A)(1+Np_b)}{p_b(1-p_b)^{A-1}}\frac{1-p_i^{m+h+1}}{1-p_i} \\ & + (1-p_b)\left(\lceil T_c\rceil p_i + (\lceil T_s\rceil - \lceil T_c\rceil)p_i^{(3)}\right)\frac{1-p_i^{m+h+1}}{1-p_i} \\ & \left. + \lceil T_s\rceil(1-p_b)(1-p_i^{m+h+1})\right]^{-1}\end{aligned}$$

L'ensemble de ces résultats seront utiles pour l'utilisation analytique du modèle général de EDCA.

Bibliographie

[80211 97] 80211. *IEEE Std 802.11-1997 Information Technology- telecommunications And Information exchange Between Systems-Local And Metropolitan Area Networks-specific Requirements-part 11 : Wireless Lan Medium Access Control (MAC) And Physical Layer (PHY) Specifications*. IEEE Std 802.11-1997, pages i–445, Nov 1997. (Cité page 7.)

[80211 07] 80211. *IEEE Standard for Information technology-Telecommunications and information exchange between systems-Local and metropolitan area networks-Specific requirements - Part 11 : Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*. IEEE Std 802.11-2007 (Revision of IEEE Std 802.11-1999), pages C1–1184, 12 2007. (Cité pages 7, 8, 9 et 15.)

[802.11e 05] 802.11e. *IEEE Standard for Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Specific requirements Part 11 : Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment 8 : Medium Access Control (MAC) Quality of Service Enhancements*. IEEE Std 802.11e-2005 (Amendment to IEEE Std 802.11, 1999 Edition (Reaff 2003)), pages 0–189, 2005. (Cité pages 7, 9, 12, 15, 17, 38 et 57.)

[802.11k 08] 802.11k. *IEEE Standard for information technology- Telecommunications and information exchange between systems- Local and metropolitan area networks-Specific requirements Part 11 : Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 1 : Radio Resource Measurement of Wireless LANs*. IEEE Std 802.11k-2008 (Amendment to IEEE Std 802.11-2007), pages c1–222, 12 2008. (Cité page 7.)

[802.11r 08] 802.11r. *IEEE Standard for Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Specific requirements part 11 : Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment 2 : fast Basic Service Set (BSS)*. IEEE Std 802.11r-2008 (Amendment to IEEE Std 802.11-2007 as amended by IEEE Std 802.11k-2008), pages c1–108, 15 2008. (Cité page 8.)

[802.16 04] 802.16. *IEEE Standard for Local and Metropolitan area Networks Part16 : Air Interface for fixed broadband Wireless Access Systems*, 2004. (Cité pages 19, 25 et 87.)

[802.16e 05] 802.16e. *IEEE Standard for Local and Metropolitan area Networks Part16 : Air Interface for fixed broadband Wireless Access Systems - Amendment 2 : Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands*, 2005. (Cité pages 24 et 26.)

[ABIResearch 09] ABIResearch. *Allied Business Intelligence, Inc.*, 2009. URL `http://www.abiresearch.com/home.jsp`. (Cité page 1.)

[Alavi 05] H.S. Alavi, M. Mojdeh & N. Yazdani. *A Quality of Service Architecture for IEEE 802.16 Standards*. 2005 Asia-Pacific Conference on Communications, pages 249–253, Oct. 2005. (Cité page 26.)

[Alexander 06] Michael Alexander & Peter Suppan. *An architecture for SDP-based bandwidth resource allocation with QoS for SIP in IEEE 802.16 networks*. In Q2SWinet '06 : Proceedings of the 2nd ACM international workshop on Quality of service & security for wireless and mobile networks, pages 75–82, New York, NY, USA, 2006. ACM. (Cité page 27.)

[Ali-Yahiya 08] T. Ali-Yahiya, T. Bullot, A.-L. Beylot & G. Pujolle. *A Cross-Layer Based Autonomic Architecture for Mobility and QoS Supports in 4G Networks*. 5th IEEE Consumer Communications and Networking Conference, 2008. CCNC 2008, pages 79–83, Jan. 2008. (Cité page 27.)

[Ansel 04] Pierre Ansel, Qiang Ni & Thierry Turletti. *An efficient scheduling scheme for IEEE 802.11e*. In Procceedings of WiOpt (Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks. IEEE, 2004. (Cité page 18.)

[ARCEP 09] ARCEP. *Autorité de Régulation des Communications électroniques et des Postes*, 2009. URL `http://www.arcep.fr/`. (Cité page 1.)

[Assi 08] C.M. Assi, A. Agarwal & Yi Liu. *Enhanced Per-Flow Admission Control and QoS Provisioning in IEEE 802.11e Wireless LANs*. IEEE Transactions on Vehicular Technology, vol. 57, no. 2, pages 1077–1088, March 2008. (Cité page 18.)

[Bai 06] Aleksander Bai, Bjørn Selvig, Tor Skeie & Paal Einar Engelstad. *A Class Based Dynamic Admitted Time Limit Admission Control Algorithm for 802.11e EDCA*. In Stefan Arbanowski, editeur, Proceedings of the 6th International Workshop on Applications and Services in Wireless Networks (ASWN), volume 1, pages 243–249. Fraunhofer FOKUS, 2006. (Cité page 18.)

[Bai 07] A. Bai, T. Skeie & P. E. Engelstad. *A Model-Based Admission Control for 802.11e EDCA using Delay Predictions*. 21st IEEE International Performance, Computing, and Communications Conference, 2007, vol. 1, pages 226–235, 2007. (Cité page 18.)

[Beizer 71] B. Beizer. The architecture and enginneering of digital computer complexes. Plenum Press, New York, 1971. (Cité page 44.)

[Bianchi 00] G. Bianchi. *Performance Analysis of the IEEE 802.11 Distributed Coordination Function*. IEEE Journal on selected areas in communications, vol. 18, March 2000. (Cité pages 17, 31 et 66.)

[Casetti 05] C. Casetti, C.-F. Chiasserini, M. Fiore & M. Garetto. *Notes on the inefficiency of 802.11e HCCA*. IEEE 62nd Vehicular Technology Conference, 2005. VTC-2005-Fall, vol. 4, pages 2513–2517, Sept. 2005. (Cité page 17.)

[Chen 04] Jie Chen, Tiejun Lv & Haitao Zheng. *Cross-layer design for QoS wireless communications*. Proceedings of the 2004 International Symposium on Circuits and Systems, 2004. ISCAS '04, vol. 2, pages II–217–20 Vol.2, May 2004. (Cité page 17.)

[Chen 05] Jianfeng Chen, Wenhua Jiao & Hongxi Wang. *A service flow management strategy for IEEE 802.16 broadband wireless access systems in TDD mode*. In IEEE International Conference on Communications, 2005. ICC 2005, volume 5, pages 3422–3426 Vol. 5, 2005. (Cité page 27.)

[Chou 05] Chun-Ting Chou, S.N. Shankar & K.G. Shin. *Achieving per-stream QoS with distributed airtime allocation and admission control in IEEE 802.11e wireless LANs*. INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies, vol. 3, pages 1584–1595 vol. 3, March 2005. (Cité page 18.)

[Cicconetti 06] C. Cicconetti, L. Lenzini, E. Mingozzi & C. Eklund. *Quality of service support in IEEE 802.16 networks*. IEEE Network, vol. 20, no. 2, pages 50–55, March-April 2006. (Cité page 26.)

[Cisco 09a] Cisco. *Cisco Aironet 1130 AG Series - Products and services*, 2009. URL `http://www.cisco.com/en/US/products/ps6087/index.html`. (Cité page 81.)

[Cisco 09b] Cisco. *Cisco Aironet 802.11a/b/g Wireless CardBus Adapter*, 2009. URL `http://www.cisco.com/en/US/prod/collateral/wireless/ps6442/ps4555/ps5818/product_data_sheet09186a00801ebc29.html`. (Cité page 81.)

[Clifford 06] P. Clifford, K. Duffy, J. Foy, D.J. Leith & D. Malone. *Modeling 802.11e for data traffic parameter design*. 4th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks, 2006, pages 1–10, April 2006. (Cité page 17.)

[Dangerfield 06] I. Dangerfield, D. Malone & D.J. Leith. *Experimental Evaluation of 802.11e EDCA for Enhanced Voice over WLAN Performance*. 2006 4th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks, pages 1–7, April 2006. (Cité page 17.)

[Delicado 06] Jesus Delicado, Luis Orozco-Barbosa, Francisco Delicado & Pedro Cuenca. *A QoS-aware protocol architecture for WiMAX*. Canadian Conference on Electrical and Computer Engineering, 2006. CCECE '06, pages 1779–1782, May 2006. (Cité page 26.)

[Delicado 08] Jesus Delicado, Francisco Delicado & Luis Orozco-Barbosa. *Study of the IEEE 802.16 contention-based request mechanism*. Journal on Telecommunication Systems, 2008, vol. 38, no. 1, pages 19–27, June 2008. (Cité page 27.)

[Engelstad 05] P.E. Engelstad & O.N. Osterbo. *Delay and Throughput Analysis of IEEE 802.11e EDCA with Starvation Prediction*. The IEEE Conference on Local Computer Networks, 2005. 30th Anniversary, pages 647–655, Nov. 2005. (Cité page 17.)

[EuQoS 06] EuQoS. *EuQoS - End-to-end Quality of Service support over heterogeneous networks*, 2006. URL `http://www.euqos.eu`. (Cité page 79.)

[Fallah 04] Yaser Pourmohammadi Fallah, Anwar Elfeitori & Hussein Alnuweiri. *A Unified Scheduling Approach for Guaranteed Services over IEEE 802.11e Wireless LANs*. International Conference on Broadband Networks, vol. 0, pages 375–384, 2004. (Cité page 18.)

[Fallah 05] Yaser Pourmohammadi Fallah & Hussein Alnuweiri. *A controlled-access scheduling mechanism for QoS provisioning in IEEE 802.11e wireless LANs*. In Q2SWinet '05 : Proceedings of the 1st ACM international workshop on Quality of service & security in wireless and mobile networks, pages 122–129, New York, NY, USA, 2005. ACM. (Cité page 18.)

[Forum 04] WiMAX Forum. *Business Case Models for Fixed Broadband Wireless Access based on WiMAX Technology and the 802.16 Standard*, 2004. (Cité page 84.)

[Freitag 07] J. Freitag & N.L.S. da Fonseca. *Uplink Scheduling with Quality of Service in IEEE 802.16 Networks*. IEEE Global Telecommunications Conference, 2007. GLOBECOM '07, pages 2503–2508, Nov. 2007. (Cité page 99.)

[Freitag 08] J. Freitag & N.L.S. da Fonseca. *Simulator for WiMAX networks*. Simulation Modelling Practice and Theory, vol. 16, no. 7, pages 817–833, Aug. 2008. (Cité page 99.)

[Gao 05] Deyun Gao, Jianfei Cai & King Ngi Ngan. *Admission control in IEEE 802.11e wireless LANs*. IEEE Network, vol. 19, pages 6–13, July-August 2005. (Cité pages 18 et 65.)

[Garg 03] P. Garg, R. Doshi, R. Greene, M. Baker, M. Malek & Xiaoyan Cheng. *Using IEEE 802.11e MAC for QoS over wireless*. Proceedings of the 2003 IEEE International Performance, Computing, and Communications Conference, pages 537–542, April 2003. (Cité page 17.)

[Ghazal 08] S. Ghazal, L. Mokdad & J. Ben-Othman. *A Real Time Adaptive Scheduling Scheme for Multi-Service Flows in WiMAX Networks*. IEEE Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008, pages 1–5, 30 2008-Dec. 4 2008. (Cité page 26.)

[Kobliakov 06] V.A. Kobliakov, A.M. Turlikov & A.V. Vinel. *Distributed Queue Random Multiple Access Algorithm for Centralized Data Networks*. IEEE Tenth International Symposium on Consumer Electronics, 2006. ISCE '06, pages 1–6, 2006. (Cité page 26.)

[Kong 04] Zhenning Kong, Danny H. K. Tsang, Brahim Bensaou & Deyun Gao. *Performance Analysis of IEEE 802.11e Contention-Based Channel Access*. IEEE Journal on selected areas in communications, vol. 22, no. 10, pages 2095–2106, December 2004. (Cité pages 17, 31, 35 et 40.)

[Kuo 08] Wen-Kuang Kuo. *Supporting multimedia streaming and best-effort data transmission over IEEE 802.11e EDCA*. International Journal on Network Management, vol. 18, no. 3, pages 209–228, 2008. (Cité page 18.)

[Lee 05] Jeng Farn Lee, Wanjiun Liao & Meng Chang Chen. *A per-class QoS service model in IEEE 802.11e WLANs*. Second International Conference on Quality of Service in Heterogeneous Wired/Wireless Networks, 2005, pages 8 pp.–26, Aug. 2005. (Cité page 18.)

[Lee 07] S. Lee, T. Kim, S. Yoon, J. Lee, M.L. Huang, K.E. Pyun & S.-C. Park. *Adaptive Physical Rate based Admission Control at EDCA in IEEE 802.11e WLANs*. Electronics Letters, vol. 43, no. 22, pages 1208–1209, 2007. (Cité page 18.)

[Li 04] Ming Li, Hua Zhu, S. Sathyamurthy, I. Chlamtac & B. Prabhakaran. *End-to-end framework for QoS guarantee in heterogeneous wired-cum-wireless networks*. First International Conference on Quality of Service in Heterogeneous Wired/Wireless Networks, 2004. QSHINE 2004, pages 140–147, Oct. 2004. (Cité page 19.)

[Lim 04] L.W. Lim, R. Malik, P.Y. Tan, C. Apichaichalermwongse, K. Ando & Y. Harada. *A QoS scheduler for IEEE 802.11e WLANs*. First IEEE Consumer Communications and Networking Conference, 2004. CCNC 2004, pages 199–204, Jan. 2004. (Cité page 18.)

[Liu 08] Cheng-Yueh Liu & Yaw-Chung Chen. *An Adaptive Bandwidth Request Scheme for QoS Support in WiMAX Polling Services*. In ICDCSW '08 : Proceedings of the 2008 The 28th International Conference on Distributed Computing Systems Workshops, pages 60–65, Washington, DC, USA, 2008. IEEE Computer Society. (Cité page 26.)

[Loyola 08] L. Loyola, H.S. Lichte, I. Aad, J. Widmer & S. Valentin. *Increasing the Capacity of IEEE 802.11 Wireless LAN through Cooperative Coded Retransmissions*. IEEE Vehicular Technology Conference, 2008. VTC Spring 2008, pages 1746–1750, May 2008. (Cité page 18.)

[Malli 04] M. Malli, Qiang Ni, T. Turletti & C. Barakat. *Adaptive fair channel allocation for QoS enhancement in IEEE 802.11 wireless LANs*. 2004 IEEE International Conference on Communications, vol. 6, pages 3470–3475 Vol.6, June 2004. (Cité page 18.)

[Mangold 02] S. Mangold, Sunghyun Choi, P. May & G. Hiertz. *IEEE 802.11e - fair resource sharing between overlapping basic service sets*. The 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 2002, vol. 1, pages 166–171 vol.1, Sept. 2002. (Cité page 17.)

[Mangold 03] S. Mangold, Sunghyun Choi, G.R. Hiertz, O. Klein & B. Walke. *Analysis of IEEE 802.11e for QoS support in wireless LANs*. IEEE Wireless Communications, vol. 10, no. 6, pages 40–50, Dec. 2003. (Cité page 17.)

[Marchese 07a] M. Marchese & M. Mongelli. *Neural bandwidth allocation function (NBAF) control scheme at WiMAX MAC layer interface*. International Journal of Communications Systems, 2007, vol. 20, no. 9, pages 1059–1079, 2007. (Cité page 26.)

[Marchese 07b] M. Marchese & M. Mongelli. *Optimal Bandwidth Provision at WiMAX MAC Service Access Point on Uplink Direction*. IEEE International Conference on Communications, 2007. ICC '07, pages 80–85, June 2007. (Cité page 26.)

[Masri 06] Mohamad El Masri. *IEEE 802.11e : the problem of the virtual collision management within EDCA*. IEEE Infocom Student Workshop '06, April 2006. (Cité page 3.)

[Masri 07a] Mohamad El Masri, Guy Juanole & Slim Abdellatif. *Revisiting the Markov chain model of IEEE 802.11e EDCA and introducing the virtual collision phenomenon*. International Conference on Wireless Information Networks and Systems, ICETE WINSYS '07, Jul. 2007. (Cité pages 3 et 63.)

[Masri 07b] Mohamad El Masri, Guy Juanole & Slim Abdellatif. *A synthetic model of IEEE 802.11e EDCA*. International Conference on Latest Advances in Networks, ICLAN '07, Dec. 2007. (Cité page 3.)

[Masri 08a] Mohamad El Masri, Slim Abdellatif & Guy Juanole. *Proposal of a Novel Bandwidth Management Framework for IEEE 802.16 Based on Aggregation*. 2nd IFIP International Conference on New Technologies, Mobility and Security, 2008. NTMS '08, Nov. 2008. (Cité page 3.)

[Masri 08b] Mohamad El Masri, Guy Juanole & Slim Abdellatif. *Hybrid admission control algorithm for IEEE 802.11e EDCA : analysis*. International Conference on Networks, ICN '08, April 2008. (Cité page 3.)

[Masri 08c] Mohamad El Masri, Guy Juanole & Slim Abdellatif. *On enhancing a hybrid admission control algorithm for IEEE 802.11e EDCA*. IFIP International Conference on Personal Wireless Communications, IFIP PWC '08, Oct. 2008. (Cité page 3.)

[Masri 09a] Mohamad El Masri & Slim Abdellatif. *Managing the virtual collision in IEEE 802.11e EDCA*. 8th IFAC International Conference on Fieldbuses anf Networks in Industrial and Embedded Systems, FET 2009, May 2009. (Cité page 3.)

[Masri 09b] Mohamad El Masri, Slim Abdellatif & Guy Juanole. *An uplink bandwidth management framework for IEEE 802.16 with QoS guarantees*. 8th International IFIP-TC6 Networking Conference, NETWORKING 2009, May 2009. (Cité page 3.)

[Mukul 06] R. Mukul, P. Singh, D. Jayaram, D. Das, N. Sreenivasulu, K. Vinay & A. Ramamoorthy. *An adaptive bandwidth request mechanism for QoS enhancement in WiMax real time communication*. 2006 IFIP International Conference on Wireless and Optical Communications Networks, pages 5 pp.–5, 2006. (Cité page 27.)

[Ng 05] Anthony C. H. Ng, David Malone & Douglas Leith. *Experimental Evaluation of TCP Performance and Fairness in an 802.11e Test-bed*. In Proceedings of the ACM SIGCOMM 2005 Workshop on experimental approaches to wireless network design and analysis (E-WIND-05), Philadelphia, PA, August 2005. (Cité page 17.)

[Ni 07] Qiang Ni, A. Vinel, Yang Xiao, A. Turlikov & Tao Jiang. *Wireless Broadband Access : WiMAX and Beyond - Investigation of Bandwidth Request Mechanisms under Point-to-Multipoint Mode of WiMAX Networks*. IEEE Communications Magazine, vol. 45, no. 5, pages 132–138, May 2007. (Cité page 26.)

[Nie 05] J. Nie, X. He, Z. Zhou & C. Zhao. *Communication with bandwidth optimization in IEEE 802.16 and IEEE 802.11 hybrid networks.* IEEE International Symposium on Communications and Information Technology, 2005. ISCIT 2005, vol. 1, pages 27–30, October 2005. (Cité page 27.)

[ns2 05] ns2. *The network Simulator - ns-2*, 2005. URL `http://www.isi.edu/nsnam/ns/`. (Cité pages 70, 84 et 99.)

[Owezarski 08] Philippe Owezarski, Pascal Berthou, Yann Labit & David Gauchard. *LaasNetExp : a generic polymorphic platform for network emulation and experiments*. In Proceedings of the 4th International Conference on Testbeds and research infrastructures for the development of networks & communities, TridentCom '08, pages 1–9, ICST, Brussels, Belgium, Belgium, 2008. ICST. (Cité page 79.)

[Park 03] Seyong Park, Kyungtae Kim, D.C. Kim, Sunghyun Choi & Sangjin Hong. *Collaborative QoS architecture between DiffServ and 802.11e wireless LAN*. The 57th IEEE Semiannual Vehicular Technology Conference, 2003. VTC 2003-Spring, vol. 2, pages 945–949 vol.2, April 2003. (Cité page 19.)

[Pong 03] Dennis Pong & Tim Moors. *Call admission control for IEEE 802.11 contention access mechanism*. IEEE Global Telecommunications Conference, GLOBECOM '03, December 2003. (Cité page 66.)

[Racaru 08] Stelian-Florin Racaru. *Conception et validation d'une architecture de signalisation pour la garantie de qualité de service dans l'Internet multi-domaine, multi-technologie et multi-service*. PhD thesis, INSA de Toulouse, Oct. 2008. 08565. (Cité page 79.)

[Ramos 05] N. Ramos, D. Panigrahi & S. Dey. *Quality of service provisioning in 802.11e networks : challenges, approaches, and future directions*. IEEE Network, vol. 19, no. 4, pages 14–20, July-Aug. 2005. (Cité page 18.)

[Ramos 07] Naomi Ramos, Debashis Panigrahi & Sujit Dey. *Dynamic adaptation policies to improve quality of service of real-time multimedia applications in IEEE 802.11e WLAN networks*. Wireless Networks, vol. 13, no. 4, pages 511–535, 2007. (Cité page 18.)

[Romdhani 03] L. Romdhani, Qiang Ni & T. Turletti. *Adaptive EDCF : enhanced service differentiation for IEEE 802.11 wireless ad-hoc networks*. 2003 IEEE Wireless Communications and Networking, 2003. WCNC 2003, vol. 2, pages 1373–1378 vol.2, March 2003. (Cité page 18.)

[Sharma 06] G. Sharma, A. Ganesh & P. Key. *Performance Analysis of Contention Based Medium Access Control Protocols*. 25th IEEE International Conference on Computer Communications, INFOCOM 2006, pages 1–12, April 2006. (Cité page 17.)

[Soundararajan 07] Srivathsan Soundararajan & Prathima Agrawal. *A scheduling algorithm for IEEE 802.16 and IEEE 802.11 hybrid networks*. Fourth International Conference on Broadband Communications, Networks and Systems, 2007. BROADNETS 2007, pages 320–322, Sept. 2007. (Cité page 27.)

[Stiliadis 98] Dimitrios Stiliadis & Anujan Varma. *Latency-rate servers : a general model for analysis of traffic scheduling algorithms*. IEEE/ACM Transactions on Networks, vol. 6, no. 5, pages 611–624, 1998. (Cité page 94.)

[Takizawa 07] Y. Takizawa, N. Taniguchi, S. Yamanaka, A. Yamaguchi & S. Obana. *Traffic Control for Cognitive Wireless Networks Composed of IEEE802.11 and IEEE802.16*. IEEE Wireless Communications and Networking Conference, 2007.WCNC 2007, pages 3831–3837, March 2007. (Cité page 27.)

[TKN-TUB 07] TKN-TUB. *An IEEE 802.11e EDCA and CFB Simulation Model for ns-2*, 2007. URL `http://www.tkn.tu-berlin.de/research/802.11e_ns2/`. (Cité page 70.)

[Velayutham 06] A. Velayutham & J.M. Chang. *An enhanced Alternative to the IEEE 802.11e MAC Scheme*. IOWA State University Report, 2006. (Cité page 18.)

[Vinel 06] A. Vinel, Ying Zhang, Qiang Ni & A. Lyakhov. *WLC22-4 : Efficient Request Mechanism Usage in IEEE 802.16*. IEEE Global Telecommunications Conference, 2006. GLOBECOM '06, pages 1–5, 27 2006-Dec. 1 2006. (Cité page 26.)

[Wang 05a] Gang Wang, Shunliang Mei, R. Yu & Zhengyong Zhang. *An improved analytical model of 802.11e EDCA with variable packet length*. IEEE 62nd Vehicular Technology Conference, 2005. VTC-2005-Fall, vol. 2, pages 1122–1126, Sept., 2005. (Cité page 17.)

[Wang 05b] H. Wang, W. Li & D.P. Agrawal. *Dynamic admission control and QoS for 802.16 wireless MAN*. 2005 Wireless Telecommunications Symposium, pages 60–66, 2005. (Cité page 26.)

[Wu 06] Haitao Wu, Xin Wang, Qian Zhang & Xuemin Shen. *IEEE 802.11e Enhanced Distributed Channel Access (EDCA) Throughput Analysis*. IEEE International Conference on Communications, 2006. ICC '06, vol. 1, pages 223–228, June 2006. (Cité page 17.)

[Xergias 05] S.A. Xergias, N. Passas & L. Merakos. *Flexible resource allocation in IEEE 802.16 wireless metropolitan area networks*. The 14th IEEE Workshop on Local and Metropolitan Area Networks, 2005. LANMAN 2005, pages 6 pp.–6, Sept. 2005. (Cité page 26.)

[Xiao 04] Yang Xiao, Haizhon Li & Sunghyun Choi. *Protection and guarantee for voice and video traffic in IEEE 802.11e wireless LANs*. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM 2004, vol. 3, pages 2152–2162 vol.3, March 2004. (Cité page 18.)

[Yahiya 08] T.A. Yahiya, A.-L. Beylot & G. Pujolle. *Cross-Layer Multiservice Scheduling for Mobile WiMAx Systems*. IEEE Wireless Communications and Networking Conference, 2008. WCNC 2008, pages 1531–1535, 31 2008-April 3 2008. (Cité page 27.)

[Yang 06] Yaling Yang & Robin Kravets. *Achieving Delay Guarantees in Ad-hoc Networks by Adapting IEEE 802.11 Contention Windows*. In Proceedings of IEEE INFOCOM 2006, 2006. (Cité page 18.)

[Yin 06] Hujun Yin & Siavash Alamouti. *OFDMA : A Broadband Wireless Access Technology*. IEEE Sarnoff Symposium, 2006, pages 1–4, March 2006. (Cité page 20.)

[Zhai 05] Hongqiang Zhai, Xiang Chen & Yuguang Fang. *How well can the IEEE 802.11 wireless LAN support quality of service ?* IEEE Transactions on Wireless Communications, vol. 4, no. 6, pages 3084–3094, Nov. 2005. (Cité page 17.)

[Zhu 05] Hua Zhu & Imrich Chlamtac. *Performance Analysis for IEEE 802.11e EDCF Service Differentiation*. IEEE Transactions on wireless Communications, vol. 4, no. 4, pages 1779–1788, July 2005. (Cité page 31.)

AUTHOR : Mohamad El Masri

TITLE : Contributing to the Quality of Service mechanisms in wireless access Networks

SUPERVISORS : Slim Abdellatif, Guy Juanole

Abstract : This thesis is a contribution to specifying, modelling and evaluating Quality of Service mechanisms destined to Wireless Local and Metropolitan networks. The first part of this work is the modelling, using Markov chains, of the IEEE 802.11e EDCA access scheme. The model we introduce adds, with regards to state of the art models, mechanisms present in the standard which were not accounted for (explicit consideration of the virtual collision phenomena) and corrects misconceptions with regards to the standard (considering the AIFS periods within the Backoff procedure). The model was rendered synthetic in order to make easier it to use (for this purpose we used the Beizer rules of reduction). The model was then used within a hybrid admission control algorithm, using in its decision process both the model and network state measures. The algorithm was validated and compared to other admission control algorithms using ns-2 simulations. In parallel, we proposed a modification of EDCA's behaviour towards virtual collisions making it fairer. This modification was evaluated using the Markov chain model. A second part of this work was the development, for WiMAX, of bandwidth management framework offering QoS guarantees. This framework is made out of three interacting parts : 1- a bandwidth management for WiMAX that can be modelled as a Latency-Rate server ; 2- an aggregated bandwidth request-response mechanism making the management simpler and more flexible ; 3- an admission control algorithm associated to the architecture guaranteeing its correct functioning. The framework was prospectively developed for heterogeneous wireless networks (WiMAX WMAN interconnecting WiFi WLANs). Our work's perspectives are on this track : defining a complete QoS solution for heterogeneous wireless networks.

Keywords : Wireless Networks, WiFi, WiMAX, Quality of Service, Modelling and Simulation

---

AUTEUR : Mohamad El Masri

TITRE : Contribution à la qualité de service dans les réseaux d'accès sans fil

DIRECTEURS DE THESE : Slim Abdellatif, Guy Juanole

LIEU ET DATE DE SOUTENANCE : à Toulouse le 9 juillet 2009

Résumé : La thèse développe une contribution à la spécification, la modélisation et l'évaluation de mécanismes destinés à la fourniture de qualité de service dans les réseaux sans fil locaux et métropolitains. La première partie du travail concerne une modélisation en chaîne de Markov du protocole d'accès EDCA de IEEE 802.11e qui, par rapport aux modèles présents dans la littérature, ajoute des mécanismes du standard qui n'avaient pas été introduits (prise en compte explicite de la collision virtuelle) et corrige des erreurs de conception (prise en compte des périodes AIFS de la procédure de Backoff). Ce modèle a été rendu synthétique pour en faciliter l'usage (réductions réalisées à l'aide des règles de Beizer). Ce modèle a ensuite été utilisé pour définir un algorithme de contrôle d'admission hybride, intégrant dans son processus de décision un modèle analytique et des mesures de l'état du réseau. L'algorithme de contrôle d'admission ainsi développé a été en premier lieu validé puis comparé à d'autres algorithmes de contrôle d'admission par simulation sous ns-2. Notons aussi que nous avons proposé, une modification du comportement de EDCA face à une collision virtuelle assurant une meilleure équité aux catégories d'accès. Cette modification a été évaluée à l'aide du modèle. Une deuxième partie du travail consiste en la proposition pour WiMAX d'une architecture de gestion de bande passante pouvant fournir des garanties de qualité de service. Cette architecture se compose de trois parties interagissantes : 1- une gestion de la bande passante sous WiMAX assimilée à une classe de serveurs dite classe des serveurs latence-débit, 2- un mécanisme de requête-réponse de bande passante agrégée simplifiant la gestion de bande passante et la rendant plus flexible, 3- un protocole de contrôle d'admission associé à l'architecture et qui en garantit le bon fonctionnement. L'architecture ainsi conçue s'inscrit dans une prospective de réseaux hétérogènes sans fil (réseaux métropolitain WiMAX connectant entre eux et à Internet des réseaux locaux WiFi). C'est dans cette optique que s'inscrira la suite de notre travail. Elle consistera en la combinaison des mécanismes que nous proposons afin de fournir une solution complète de qualité de service pour des réseaux hétérogènes sans fil.

Mots-clés : Réseaux sans fil, WiFi, WiMAX, Qualité de Service, Modélisation et Simulation

---

Discipline Administrative : Systèmes Informatiques

Intitulé et adresse du Laboratoire : Laboratoire d'Analyse et d'Architectures des Systèmes - LAAS - CNRS- 7 avenue du colonel roche - 31077 Toulouse