UNIVERSITE LIBRE DE BRUXELLES
Faculté des Sciences appliquées

Année académique 2004-2005

Analyse de performance de *handover* vertical entre réseaux UMTS et 802.11

DIRECTEUR DE MEMOIRE : Esteban Zimányi

MEMOIRE DE FIN D'ETUDES
PRESENTE PAR Xavier Hick EN VUE
DE L'OBTENTION DU GRADE
D'INGENIEUR CIVIL Informaticien

Etat de l'art

Plusieurs références traitent d'une partie du handover inter-système.
[b] traite notamment de la discussion d'un algorithme de handover pour une interconnexion entre réseaux hétérogènes. [c] propose un schéma de gestion de mobilité multicouche pour le support de réseaux IP dans un contexte AAA. [d] discute de mécanismes et protocoles de gestion de mobilité basés sur les couches réseau, réseau et liaison et liaison.
Enfin, [e] analyse et discute les performances du handover entre UMTS et WLAN au moyen du protocole de mobilité mSCTP.
Ce projet veut considérer une approche basée sur en comparant les architectures UMTS et 802.11, ainsi que les protocoles de mobilité mentionnés dans [d] tout en tenant compte des particularités de chaque protocole envisagé. Ce projet se démarque de [a][b][c][d] en développant un organe de décision placé dans le terminal mobile (l'entité de handover) ainsi que par le développement d'une architecture d'interconnexion UMTS/802.11 basée sur le protocole Mobile IP dans le simulateur logiciel NS-2.

Résumé

Cette recherche tend à analyser les performances de handover vertical entre les réseaux UMTS et 802.11 (WLAN) en définissant une architecture de développement pour le support de la mobilité entre ces deux réseaux.
Tout d'abord, les architectures et principes de fonctionnement des réseaux UMTS et 802.11 sont détaillés et comparés.
Ensuite, les principaux protocoles de mobilité sont envisagés comme support de handover entre UMTS et 802.11.
Enfin, une architecture de simulation est présentée suivie de l'analyse de performance modélisée par le simulateur NS-2 dans un contexte de communication orientée connexion.

Remerciements

Je tiens à remercier Laurent Schumacher, professeur aux Facultés Universitaires Notre Dame de la Paix à Namur pour ses remarques et conseils concernant l'UMTS ; Olivier De Mey, doctorant aux Facultés Universitaires Notre Dame de la Paix à Namur pour son soutien moral et ses explications sur l'extension EURANE ; Jean-Michel Dricot, doctorant à l'U.L.B. pour ses conseils et explications sur le fonctionnement de NS-2 et en particulier de 802.11 dans NS-2.

Enfin, je remercie ma famille et mes proches pour le soutien qu'ils m'ont apportés.

Table des matières

Liste des figures

UMTS

L'UMTS pour "Universal Mobile Telecommunications System" désigne une norme cellulaire numérique de troisième génération compatible avec la norme actuelle de deuxième génération plus connue sous le nom GSM (Global System for Mobile communications).
Cette norme est en cours de développement par le Partenariat de Projet 3ème Génération (3GPP), un rassemblement de plusieurs organisations développeuses de standards – l'ETSI[1] (Europe), l'ARIB/TTC[2] (Japon), l'ANSI[3] T-1 (USA), la TTA[4] (Corée du Sud) et le CWTS[5] (Chine).
Pour atteindre l'accord de toutes ces organisations, 3GPP introduit l'UMTS en phases et releases annuelles. La première release, introduite en décembre 1999, définit les améliorations et transitions pour les réseaux GSM existants. Les introductions significatives ont été les définitions de l'UTRAN (UMTS Terrestrial Radio Access Network) et de l'UE (User Equipment), l'UE communiquant avec l'UTRAN via la technologie radio WCDMA[6]. Par opposition, la définition du CN (Core Network) est adoptée du GSM, facilitant l'introduction de la technologie UMTS dans l'architecture GSM [1][2][3][4].

Architecture réseau

Le réseau UMTS est constitué d'entités logiques assurant chacune une fonction précise. Ces entités sont groupées en UTRAN chargé des fonctions « radio », UE l'interface utilisateur exploitant le medium radio et le CN responsable de la commutation et du routage d'appels et de données vers des réseaux extérieurs [3][4].

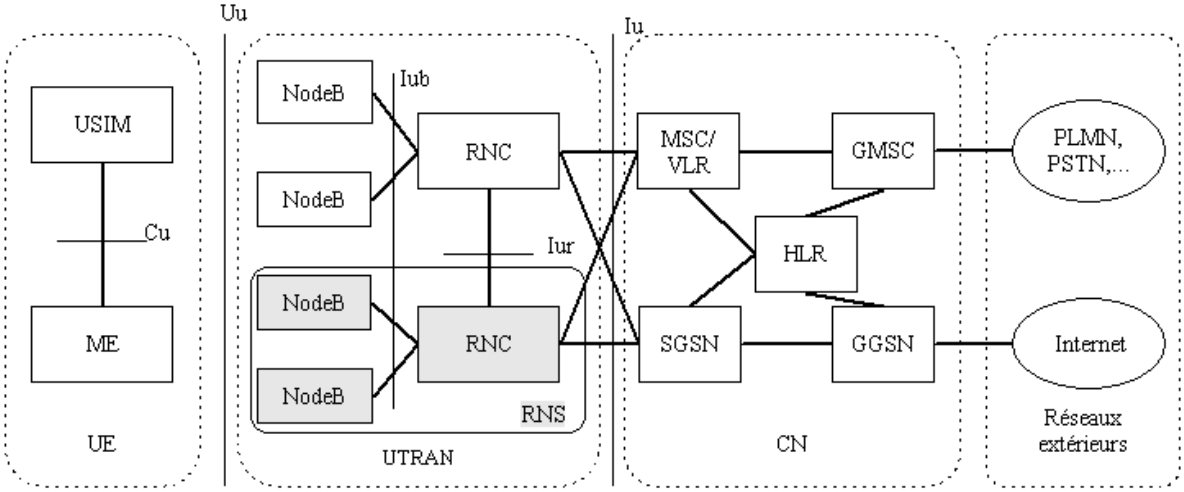

Figure 1 – Architecture d'un réseau UMTS [3]

UE (User Equipment)

L'UE consiste en un ME (Mobile Equipment) et un USIM (UMTS Subscriber Identity Module). Le ME est le terminal radio employé pour la communication radio sur l'interface Uu (l'interface air WCDMA). L'USIM est une carte à puce conservant l'identité de l'abonné, effectuant des algorithmes d'authentification, et stockant d'autres informations requises par le terminal.

[1] European Telecommunications Standards Institute
[2] Association of Radio Industries and Business/Telecommunication Technology Committee
[3] American National Standards Institute Standards Committee
[4] Telecommunications Technology Association
[5] Chinese Wireless Telecommunication Standard
[6] Wideband Code Division Multiple Access

UTRAN (UMTS Terrestrial Radio Access Network)

L'UTRAN introduit le NodeB (aussi appelé Base Station) et le RNC (Radio Network Controller). Il est divisé en systèmes réseau radio individuels (RNS), où chaque RNS est contrôlé par un RNC. Le RNC est connecté à un ensemble de NodeBs, chaque NodeB pouvant servir une ou plusieurs cellules. Une cellule peut être divisée en secteurs, notamment en zone urbaine. La zone de couverture d'un NodeB représente un ou plusieurs secteurs/cellules.

Le NodeB, équivalent du Base Transceiver Station (BTS) de GSM, convertit les signaux de l'interface radio Uu en un flux de données et le transmet au RNC par l'interface Iub. Dans la direction opposée, il prépare les données parvenant au RNC pour le transport sur l'interface radio.

Le RNC est le nœud central dans l'UTRAN et l'équivalent du Base Station Controller (BSC) de GSM. Il contrôle un ou plusieurs NodeBs à travers l'interface Iub et est responsable du contrôle des ressources radio dans son domaine (i.e. l'ensemble des NodeBs qui lui sont connectés).

On distingue 2 rôles logiques pour un RNC : SRNC et DRNC.

Le SRNC (Serving RNC) d'un mobile est le RNC transportant les données utilisateur par l'interface Iu. Le SRNC est l'intermédiaire final pour la signalisation entre l'UE et l'UTRAN. Tout UE est associé à un seul SRNC.

Le DRNC (Drift RNC) est un RNC contrôlant les cellules utilisées par un mobile et qui relaye les données depuis/vers le SRNC.

CN (Core Network)

Le CN, issu de GSM, est divisé en réseau à commutation de paquet (PS), réseau à commutation de circuits (CS) et un Home Location Register (HLR).

Le réseau PS consiste en un Serving GPRS Support Node (SGSN) et une Gateway GPRS Support Node (GGSN). Le SGSN est responsable du routage de paquets dans le réseau PS ainsi que de l'authentification et du cryptage pour les utilisateurs.
Le GGSN est utilisé comme la passerelle pour les réseaux externes à commutation de paquet tels que l'Internet, les LANs, les WANs, les réseaux GPRS, les réseaux ATM etc. et donc termine la fonction de routage du SGSN.

Le réseau CS consiste en un Mobile Service Switching Centre (MSC)/Visitor Location Register (VLR) et une Gateway MSC (GMSC). Le MSC/VLR sert comme un commutateur et une base de données. Le MSC est responsable de la signalisation requise pour l'établissement, la fermeture et le maintien des connexions. Il est aussi chargé des fonctions mobile radio telles que le reroutage d'appel ainsi que de l'allocation de canaux radio pour un mobile.
Le VLR est contrôlé par le MSC et est employé pour gérer les utilisateurs en « roaming » dans la zone associée aux MSC. Il stocke les informations transmises par le HLR responsable des utilisateurs.

Le GMSC a une fonction similaire au GGSN. Il sert de passerelle pour les réseaux externes à commutation de circuit comme d'autres PLMNs, les réseaux téléphoniques publiques (PSTNs), les réseaux ISDN etc.

Le HLR est une base de données située dans le système domiciliataire de l'utilisateur (i.e. l'opérateur principal de l'utilisateur) qui conserve toutes les informations relatives à l'utilisateur. Ces informations sont rassemblées dans un profil de service et consistent entre autres en un numéro de téléphone, une clé d'authentification, les services autorisés, les zones de roaming interdites etc. Le HLR contient aussi des informations sur la localisation de l'UE.

Interfaces

Le standard UMTS spécifie des interfaces entre les entités logiques du réseau UMTS. Ces interfaces standardisées sont ouvertes. Les fabricants d'entités du réseau UMTS doivent donc inclure le support des protocoles définis pour chaque interface. Ceci assure la compatibilité entre entités UMTS issues de fabricants différents [3].
Ces interfaces sont :
Cu : l'interface électrique entre l'USIM et le ME
Uu : l'interface radio WCDMA à travers laquelle l'UE accède à l'UTRAN.
Iu : cette interface connecte l'UTRAN au CN. On distingue Iu-CS l'interface Iu pour le domaine à commutation de circuits et Iu-PS cette interface pour le domaine à commutation de paquets.
Iur : l'interface d'interconnexion des RNCs notamment pour le support de handover entre RNCs ou encore pour interconnecter un SRNC et un DRNC.
Iub : l'interface connectant un NodeB à un RNC [3]
Les piles de protocoles supportées par ces interfaces sont données dans l'Annexe UMTS.

Scénarios d'utilisation

Pour comprendre le fonctionnement d'UMTS, on considère différents scénarios d'utilisation. Les scénarios d'utilisation décrivent le comportement et les interactions des entités du réseau UMTS avant, pendant et après une communication. Les scénarios d'utilisation sont basés sur les états de service de l'UE.
L'UE peut se trouver dans 3 états de services différents : détaché, connecté ou inactif. L'UE est en état détaché quand il est éteint. Aucune communication entre l'UE et le réseau n'est alors possible. L'UE ne peut rien envoyer ni recevoir. Pour pouvoir communiquer avec le réseau, l'UE doit s'attacher au réseau en l'allumant, en sélectionnant une cellule à laquelle s'attacher, et s'attacher à cette cellule. Quand l'UE est attachée au réseau, il passe en état connecté et peut débuter la communication, ou passer en état inactif s'il ne demande pas de communication [5].
Les états sont différents selon que l'UE soit en mode CS ou PS. On considère 6 scénarios d'utilisation importants : l'attachement au réseau, la connexion CS, l'inactivité CS, la connexion PS, l'inactivité PS et le détachement du réseau. Ces scénarios sont détaillés par la suite.

Attachement au réseau

Le processus d'attachement au réseau débute quand l'utilisateur allume l'UE. Il doit alors entrer un code PIN pour s'authentifier à l'USIM. Si l'authentification USIM est validée, l'UE cherche un NodeB (dans une cellule ou secteur de cellule) auquel s'attacher. La procédure d'attachement est toujours initiée par l'UE. Quand l'UE trouves un NodeB auquel s'attacher, il se synchronise avec lui, et tente ensuite de s'y attacher en envoyant une requête d'attachement au réseau, en particulier au RNC. Le réseau répond en envoyant l'identification USIM de l'UE au HLR pour l'informer de la requête d'attachement de l'UE.
Le HLR et l'USIM partagent une clé secrète de 128bits, que le HLR applique à un nombre aléatoire. Le résultat et le nombre aléatoire sont alors envoyés au réseau. Le réseau envoie le nombre aléatoire à l'UE. De façon semblable, l'USIM applique la clé secrète au nombre

aléatoire, ensuite l'UE renvoie le résultat au réseau. Si le résultat envoyé par l'UE est identique au résultat envoyé par le HLR, le réseau accepte l'UE et l'attache au réseau. Enfin, le réseau récupère le profil de l'utilisateur depuis le HLR et le transmet au VLR associé au NodeB d'attachement.

Connexion CS

Après le processus d'attachement au réseau, l'UE peut effectuer le processus de connexion CS. Le processus de connexion CS comporte aussi bien l'établissement de communication que la réception de communication. Le point commun de l'établissement et de la réception de communication est qu'ils nécessitent tous deux une connexion de signalisation entre l'UE et le CN.

Pour établir une communication – procédure d'appel d'un poste téléphonique d'un réseau externe, une connexion CS doit d'abord être faite. L'UE signale, pour ce faire, au MSC qu'elle nécessite une connexion CS à un numéro particulier. Le MSC consulte le profil de l'utilisateur dans le VLR pour déterminer si l'utilisateur a la permission d'appeler le numéro. Si l'appel est permis, le MSC vérifie s'il dispose de circuits disponibles et si l'UTRAN dispose de ressources pour supporter la communication. Si c'est le cas, il établit la connexion CS de l'UE, par l'interface air, passant par l'UTRAN jusqu'au MSC du CN. Le MSC commute alors l'appel au GMSC, qui se charge de la commutation dans le réseau externe CS. Le réseau externe CS effectue alors les fonctions de commutations nécessaires pour diriger l'appel à la destination.
Quand la communication est terminée, le MSC et le GMSC produisent un Call Detail Record (CDR). Le CDR contient les informations concernant l'identité des parties appelante et appelée, les ressources utilisées, etc. et est transmis au serveur de facturation.

Pour la réception de communication – procédure d'appel de l'UE depuis un réseau externe, la procédure est différente. D'abord, l'appel est routé à travers le réseau externe vers le GMSC. Le GSMC détermine alors le HLR contenant le profil de l'utilisateur appelé sur base de son numéro de téléphone. Le HLR connaît la « location area » de l'UE (un ensemble de cellules dans lesquelles l'UE est susceptible de se trouver et d'être appelée), et est de ce fait capable d'envoyer une requête de numéro indiquant le MSC de destination au VLR responsable de cette location area. Le VLR renvoie le numéro du MSC, et le HLR transmet alors le numéro au GMSC. Le GMSC est maintenant capable de router l'appel jusqu'au MSC. A parti du VLR, le MSC connaît le RNC responsable de la location area de l'UE appelée et peut dès lors demander à ce RNC l'établissement d'un canal vers l'UE. Le RNC signale alors l'appel à l'UE dans la dernière location area connue et établit une connexion à l'UE à travers le NodeB quand l'UE répond à l'appel. Quand la liaison de transmission est établie, l'UE se met à sonner. Quand l'utilisateur décroche, la communication est commutée par la liaison établie.

Quand la connexion de signalisation pour les services CS est fermée, lors de la clôture de communication ou lors d'un défaut de liaison radio, l'UE peut être ordonnée par le réseau de passer en mode inactif CS. Alternativement, l'UE peut passer en état détaché soit sur ordre du réseau soit par l'utilisateur. [4][5][6]

Inactivité CS

Si la connexion de signalisation des services CS est fermée, l'UE passe de l'état connecté CS à l'état inactif CS. Le réseau arrête le traçage de localisation de l'UE et l'UE écoute le canal de diffusion des cellules. Tant que l'UE reste dans la même location area, la situation demeure inchangée. Si l'UE se déplace dans une nouvelle location area, il informe le MSC de son

changement de localisation. La mise à jour de localisation est stockée dans le HLR et copiée dans le VLR attachée au MSC.
Si l'utilisateur désire effectuer un appel, l'UE passe en état de connexion CS et effectue la procédure d'établissement de communication. S'il y'a une communication entrante pour l'UE, le RNC le signale à l'UE. Quand l'UE répond, le RNC établit la connexion et l'UE se met à sonner. Quand l'utilisateur décroche, la communication est commutée. [4][5][6]

Connexion PS

Une alternative à la connexion CS est la connexion PS. Le processus de connexion PS comporte également l'établissement de communication et la réception de communication. Une connexion de signalisation entre l'UE et le CN est requise, comme pour les connexions CS.

Pour établir une communication, une connexion PS doit être établie. L'UE active d'abord le contexte PDP (Packet Data Protocol) dans le GGSN. Un contexte PDP est un ensemble de paramètres définissant les réseaux de paquets qu'un utilisateur peut employer pour transmettre des données. La liste des contextes PDP permis pour l'utilisateur est stockée dans le HLR. Pour activer le contexte PDP, l'UE établit une connexion par le RNC jusqu'au SGSN et envoie un message de requête d'établissement de connexion à un réseau PS externe. Le SGSN transmet la requête au GGSN, qui interroge alors le HLR pour vérifier si l'utilisateur est autorisé à accéder aux réseaux PS externes. Si l'utilisateur est autorisé, le GGSN active le contexte et informe l'UE en incluant une adresse IP. L'activation du contexte crée un tunnel IP fixe vers lequel les paquets de données sortants sont envoyés au RNC et transmis ensuite au GGSN. Le GGSN commute alors l'appel dans le réseau PS externe, qui effectue les fonctions de commutation nécessaires pour diriger l'appel à destination. Le tunnel est actif jusqu'à ce que l'UE désactive le contexte soit en fermant l'application soit en se déconnectant du SGSN.
Le SGSN est informé continuellement de la routing area actuelle de l'UE (la routing area est l'équivalent PS de la location area). Si l'UE change de routing area pour une area avec un nouveau SGSN responsable de cette area, la route dans le GGSN est adaptée à cette area.
Grâce à la requête au HLR, le SGSN et le GGSN sont conscients de la qualité de service (QoS) demandée pour le transfert de paquet et sont capables d'établir un chemin de transfert de paquet conformément à la QoS désirée.
Les catégories de QoS pour les connexions PS sont conversationnel (voix), streaming (streaming vidéo), interactif (navigation web) et background (transfert de fichier, e-mails). Quand la communication est terminée, le SGSN génère un enregistrement de facturation sur base du contexte PDP et l'envoie au serveur de facturation.

Pour recevoir un appel PS, un autre processus est requis. D'abord, l'appel entrant est routé à travers le réseau PS externe jusqu'au GGSN. Le GGSN détermine alors le HLR dans lequel le profil de l'utilisateur appelé est stocké sur base de son numéro de téléphone. Le GGSN interroge ensuite le HLR et détermine si l'UE est attachée au réseau et a activé un contexte PDP. Si l'UE n'est pas attaché au réseau, l'appel est rejeté. Si l'UE est attachée au réseau mais ne dispose pas d'un contexte PDP, l'UE doit être localisée et recevoir un signal d'activation de contexte PDP. Le HLR connaît la routing area de l'UE. Il connaît également le SGSN responsable de l'UE. Le GGSN obtient cette information et vérifie le profil de l'utilisateur dans le HLR concernant l'attachement au réseau de l'UE et le statut du contexte PDP. Le GGSN est à présent capable de router l'appel vers le SGSN. Le SGSN connaît le RNC responsable de la routing area et demande au RNC d'établir un canal vers l'UE. Le RNC appelle l'UE dans la dernière routing area connue et établit une connexion à l'UE à travers le

NodeB utilisé par l'UE lors de sa réponse à l'appel. Une fois la liaison établie, l'UE reçoit l'appel PS et la communication débute.

Quand la connexion de signalisation pour les services PS est fermée, lors de la clôture du service PS ou lors d'activité très faible ou de défaut de liaison radio, l'UE peut être ordonnée à passer en état inactif PS. Alternativement, l'UE peut passer en état de détachement réseau. [4][5][6]

Inactivité PS

Si l'UE est restée inactive durant un certain temps, elle passe de l'état connecté PS à l'état inactif PS. Le réseau arrête le traçage de localisation de l'UE, et l'UE écoute simplement le canal de diffusion des cellules. Si l'UE se déplace dans une nouvelle routing area, il informe le SGSN de son changement de localisation. La mise à jour de localisation est stockée dans le HLR et copiée dans le VLR.
La connexion logique entre le GGSN et l'UE est maintenant aussi en état inactif PS si le contexte PDP n'a pas été désactivé. Si l'utilisateur désire faire un appel et que le contexte PDP est toujours actif, l'UE passe en état connecté PS et débute l'appel. Si le contexte PDP est inactif, l'UE doit d'abord passer en état connecté PS et activé le contexte PDP avant d'effectuer l'appel. S'il arrive une communication entrante pour l'UE et que le contexte PDP est toujours activé, l'UE passe automatiquement en état connecté PS à la réception de l'appel. Si le contexte PDP est inactif pour un appel entrant, le RNC signale à l'UE d'activer le contexte PDP dans la dernière routing area connue. Quand l'UE répond au signal du RNC, le RNC établit une connexion à l'UE et l'appel entrant est dirigé vers l'UE. Alternativement, l'UE peut passer en état détaché.

Détachement du réseau

Quand l'UE ne nécessite plus de services du réseau UMTS, il peut passer en état détaché. Pour ce faire, il envoie une requête de détachement au réseau. Le détachement du réseau peut aussi être initié par le réseau soit explicitement en demandant le détachement à l'UE soit implicitement en détachant l'UE sans l'avertir. Le détachement implicite s'effectue lorsque le réseau ne parvient plus à atteindre l'UE durant un certain temps ou après la déconnexion du lien logique. Quand le détachement du réseau est invoqué, les données transitant dans le réseau et destinées à l'UE sont supprimées. [5]

Gestion de la mobilité

La mobilité dans les réseaux cellulaires tels que UMTS impliquent 2 mécanismes : la gestion de la localisation de l'UE et la gestion de handover. La gestion de localisation est le mécanisme permettant de conserver la localisation de l'UE indépendamment de ses connexions actives et la gestion de handover est le mécanisme de transfert de connexion active d'un UE d'un canal radio vers un autre.

Gestion de localisation

Afin de transférer une connexion entrante à un UE inactif, le réseau doit continuellement mettre à jour ses informations de localisation de l'UE. Le procédé de mise à jour des informations de localisation est défini pour les services CS et PS.
Pour les services CS, le réseau est divisé en Location Areas (LA). Une LA consiste en un nombre de cellules dans lesquelles l'UE peut se déplacer sans mettre à jour sa localisation. Tous les NodeBs d'une même LA diffusent un identifiant spécifique, le LA Index, qui est reçu par l'UE. L'UE devient consciente de son changement de LA quand cet identifiant

change. Il exécute consécutivement une mise à jour Location Area Update avec le MSC qui transmet ensuite l'information au HLR.
Pour les services PS, le réseau est divisé en zones plus petites que les cellules, appelées Routing Areas (RA). Une RA est englobée dans une LA. Quand l'UE devient consciente du changement de RA, une mise à jour Routing Area Update est exécutée avec le SGSN qui transmet l'information au HLR [2].
L'UE devient consciente (y compris en mode inactif) du changement de RA ou LA au moyen du protocole RRC (Protocole RRC (Radio Resource Control)explicité dans l'Annexe UMTS). Le réseau connaît la localisation de l'UE en mode inactif via la LA ou RA enregistrée dans le HLR. Lors de l'établissement d'une connexion avec l'UE, le réseau peut obtenir sa localisation à la cellule près en forçant l'UE à passer en état Cell_FACH ou Cell_DCH [3].

Gestion de handover

Pour transférer une connexion active d'un canal radio vers un autre, ce dans la même cellule ou d'une cellule à une autre, le réseau effectue un handover de connexion.
De façon similaire à la gestion de localisation, le traitement de handover est défini pour les services CS et PS.
Pour les services CS, les handovers peuvent être implémentés comme soft handover, softer handover et hard handover.
Le soft handover est effectué lors du mouvement de l'UE d'une cellule à une autre. L'UE communique initialement avec le NodeB(1) lié à un RNC dans sa cellule actuelle. Ensuite, l'UE se déplace vers un NodeB(2) lié au même RNC dans une autre cellule et commence à recevoir les mêmes données des 2 NodeBs. En se déplaçant, l'UE surveille périodiquement la qualité du signal provenant d'autres cellules. Si la différence de puissance du signal entre le NodeB(2) et le NodeB(1) atteint un seuil appelé marge de handover durant un certain laps de temps, le RNC responsable des 2 NodeBs établit une connexion avec le NodeB(2). Quand la puissance du signal du NodeB(1) devient inférieure à celle du NodeB(2), la connexion vers le NodeB(1) est supprimée. Il s'est produit un soft handover.

Un softer handover consiste en un soft handover excepté que le handover s'effectue lors du déplacement de l'UE d'un secteur à un autre secteur d'un même NodeB.

Un hard handover a lieu lorsque la connexion à la cellule courante est rompue avant que l'établissement d'une connexion à une autre cellule ne soit fait. Il existe plusieurs types de hard handover : inter-fréquentiel, intra-fréquentiel et inter-système.
Un hard handover inter-fréquentiel est un handover entre 2 fréquences différentes dans la même cellule ou entre cellules adjacentes.
Un hard handover intra-fréquentiel est effectué quand l'interface Iur entre 2 RNCs n'est pas disponible pour un soft handover. Un handover est alors effectué d'une cellule appartenant à un RNC vers une autre cellule appartenant à un autre RNC en conservant la même fréquence.
Un hard handover inter-système est effectué quand il est nécessaire de changer de technologie d'accès radio, par exemple de UMTS vers GSM.

Pour les services PS, le seul type de handover défini est la re-sélection de cellule. Elle se produit dans la situation suivante. L'UE surveille périodiquement la qualité du signal des cellules à sa portée lorsqu'il se déplace. L'UE doit envoyer périodiquement (toutes les 0.67ms) un rapport de mesure au SRNC. Le SRNC, après réception du rapport, initie un handover si la qualité de réception du signal d'une autre cellule dépasse un seuil et que la qualité de réception de la cellule courante est insatisfaisante. Le SRNC ordonne ensuite au DRNC contrôlant l'autre cellule de réserver des ressources radio. Le DRNC renvoie un

message de commande de handover incluant les détails des ressources allouées à l'UE. Quand l'UE réceptionne la commande de handover, il établit une connexion avec la nouvelle cellule conformément aux paramètres inclus dans le message de commande de handover. L'UE confirme le bon déroulement du handover en envoyant un message handover terminé au DRNC, après quoi le DRNC initie la libération de l'ancienne connexion. Enfin, quand la re-sélection de cellule est terminée, l'UE initie la procédure de mise à jour de RA. Vu que UMTS utilise la transmission et réception continues en état connecté PS, l'UE ne peut mesurer la qualité de réception d'autres cellules simultanément à l'envoi et la réception de données si l'UE dispose d'un seul récepteur radio. Pour surmonter cet obstacle, il existe le mode compressé.

Le mode compressé est une méthode qui crée des courts instants d'inactivité dans la transmission et la réception. Pour maintenir un bit rate perçu par l'utilisateur constant, le bit rate de transmission est augmenté ou compressé avant et après l'instant d'inactivité. Un bit rate constant est nécessaire pour des services tels que la voix, mais pour des services de données, un bit rate constant n'est pas requis. La transmission peut dès lors être simplement reportée pour créer l'instant d'inactivité.

L'UE emploie le mode compressé si il ne dispose que d'un récepteur radio. Si l'UE contient davantage de récepteurs radio, il peut utiliser chaque récepteur en parallèle, en effectuant des mesures sur d'autres cellules/réseaux d'accès tout en communiquant avec l'UTRAN sans employer le mode compressé.

L'inconvénient du hard handover inter-système et de la re-sélection de cellule est que ces 2 procédés ne fonctionnent que pour des handovers entre UMTS et GSM. Pour le handover entre UMTS et WLAN, le support d'autres protocoles est requis [3][4].

802.11

802.11 fait partie de la famille IEEE 802, un ensemble de spécifications pour technologies LAN. Les spécifications IEEE 802 se basent sur les 2 couches basses du modèle OSI. La spécification 802.11 inclut la couche 802.11 MAC et 3 couches physiques à radio-fréquence à spectre étalé (spread-spectrum) et une couche physique à diffusion infrarouge. 802.11a décrit une couche physique basée sur le multiplexage de fréquences orthogonales (OFDM). Voir annexe 802.11 pour les détails de couches physiques.

Architecture réseau

L'architecture réseau 802.11 consiste basiquement en un ou plusieurs « Basic Service Sets » (BSSs) liés par un Système de Distribution (DS).

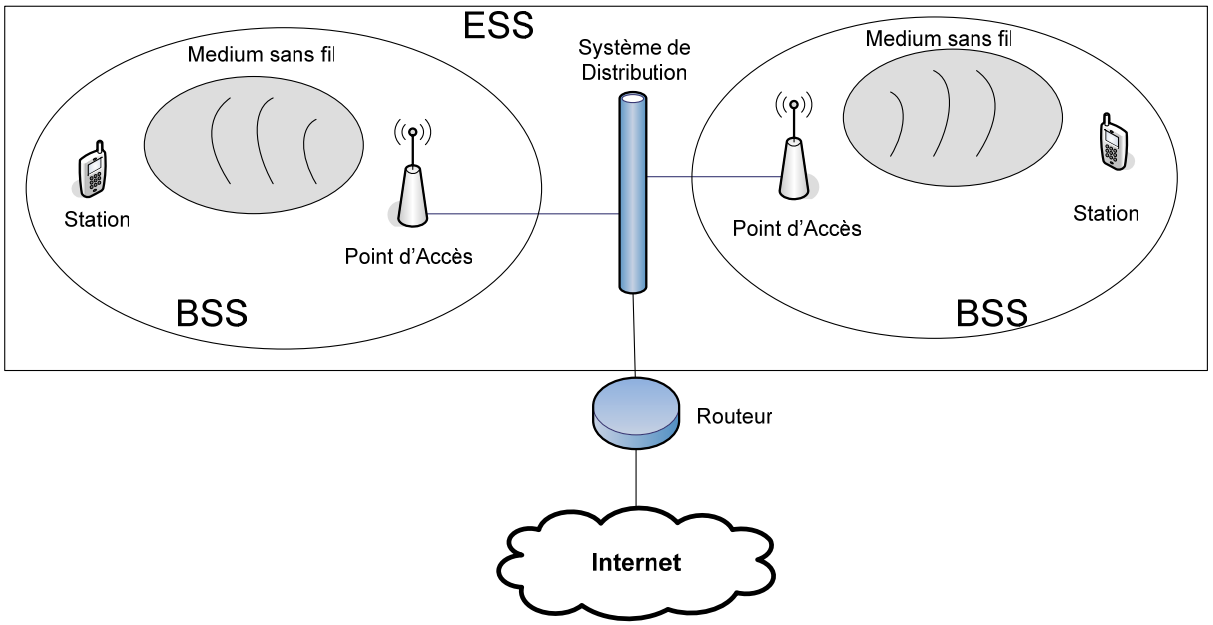

Figure 2 – Architecture 802.11 [8]

Basic Service Set

Le BSS consiste en un groupe de stations (STAs) qui sont sous le contrôle d'une même fonction de coordination. Les STAs sont des appareils informatiques (ordinateurs portables, pocket PCs, etc.) disposant d'une interface réseau sans fil. Les STAs communiquent à travers le medium sans fil (air). La zone géographique couverte par l'AP est appelée la Basic Service Area (BSA). La BSA est analogue à une cellule dans le réseau UMTS.
Dans un réseau infrastructure, toutes les STAs communiquent en envoyant tout le trafic vers un point centralisateur appelé le Point d'Accès (AP). L'AP contrôle la communication dans le BSS et fournit aussi la connectivité réseau avec d'autres BSSs. L'AP est analogue au NodeB du réseau UMTS.
A l'opposé du BSS infrastructure, dans un BSS indépendant (IBSS), connu aussi sous le nom de réseau ad-hoc, les STAs communiquent entre elles sans AP comme intermédiaire.
Pour interconnecter un réseau 802.11 à d'autres réseaux, un AP est requis. Dès lors, seul le réseau infrastructure est considéré par la suite.

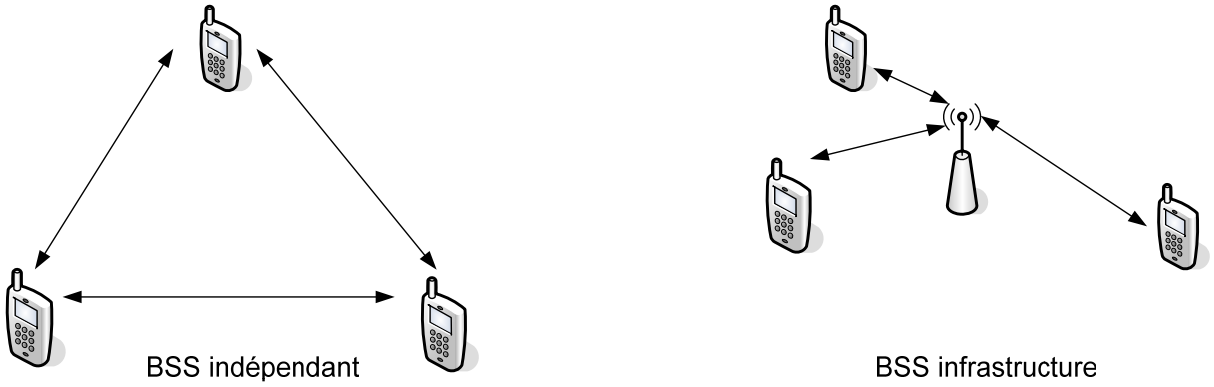

Figure 3 – Configuration BSS [8]

Dans un Infrastructure BSS, les STAs sont tenues de rester à portée de l'AP pour pouvoir communiquer entre elles.
Dans un Indepedant BSS, les STAs communiquent directement entre elles et doivent dès lors être à portée les unes des autres.

Système de distribution

Le DS est un composant logique permettant la communication entre BSSs interconnectés.
L'interconnexion de plusieurs BSSs utilisant un DS est appelée un Extended Service Set (ESS). Le DS est responsable du traçage de la localisation des STAs dans l'ESS et de la transmission de données, destinée à une STA, à l'AP associé à cette STA, l'AP se chargeant de la transmission à la STA. L'ESS permet une plus grande couverture que le BSS. La zone de couverture d'un ESS, appelée « extended service area », correspond à l'ensemble des BSAs des BSS constituant l'ESS. L'ESS ne fournit pas seulement l'accès sans fil pour une multitude d'utilisateurs mais aussi une passerelle d'accès pour les utilisateurs sans fil à un réseau filaire tel que l'Internet. Ceci est fait par l'intermédiaire d'un routeur. [7][8][9]

Scénarios d'utilisation

Pour comprendre le fonctionnement de 802.11, il est utile, tout comme pour UMTS, de considérer les différents scénarios d'utilisation. Les scénarios d'utilisation pour 802.11 sont basés sur les états de services dans lesquels la STA peut être.
La STA peut être dans les 3 états : détaché, connecté et inactif. L'état détaché est lorsque la STA est éteinte et qu'aucune communication n'est possible entre la STA et le réseau. Pour avoir accès au réseau, la STA doit d'abord être allumée, scanner les réseaux disponibles, décider quel réseau joindre parmi les réseaux disponibles puis s'authentifier et s'associer avec le réseau choisi. Quand la STA est connectée à un réseau, elle peut envoyer et recevoir des trames de données commutées paquet, ou passer en mode d'économie d'énergie c-a-d l'état inactif. Enfin, la STA peut choisir de ne plus utiliser les ressources du réseau et de s'en détacher ou encore le réseau peut décider de détacher la STA pour diverses raisons. [9]
On distingue 4 scénarios d'utilisation : l'attachement au réseau, les connexions PS, l'inactivité PS et le détachement du réseau. Ces scénarios sont détaillés par la suite.

Attachement au réseau

Pour s'attacher à un réseau, la STA doit d'abord être allumée par l'utilisateur. Ensuite, la STA doit identifier un réseau compatible. Ce processus d'identification des réseaux existant dans une zone est appelé le scanning.

Scanning

La procédure de scanning est basée sur plusieurs paramètres qui peuvent être soit des valeurs par défaut soit spécifiées par l'utilisateur. Les paramètres incluent : BSSType, BSSID, SSID, ScanType, ChannelList, ProbeDelay, et MinChannelTime, MaxChannelTime.
BSSType spécifie si le scanning porte sur des réseaux ad-hoc, infrastructure ou les 2
BSSID cherche un réseau spécifique à joindre (spécifié par son identifiant) ou par défaut, n'importe quel réseau autorisant la STA à le joindre.
SSID assigne un identifiant à un ESS, ce qui permet à la STA de chercher un ESS spécifique.
ScanType autorise le scanning passif ou actif, détaillé par la suite.
ChannelList spécifie une liste de canaux que la STA peut scanner.
ProbeDelay est un intervalle de temps spécifié avant le début de la procédure de scanning actif.
MinChannelTime et MaxChannelTime sont des intervalles minimum et maximum de temps de scanning sur un canal.
Le scanning peut être soit actif soit passif. En scanning passif, la STA économise la puissance de sa batterie parce qu'elle ne transmet pas. Elle passe simplement d'un canal à un autre sur la ChannelList et attend la trame balise des APs à proximité. La trame balise contient les informations nécessaires à la STA pour s'accorder à un BSS et entamer la communication.
En scanning actif, la STA joue un rôle actif et tente de trouver le réseau au lieu d'attendre l'annonce du réseau. Des trames sondes de requête sont utilisées pour solliciter la réponse de l'AP d'un réseau qui en retour renvoie une trame sonde de réponse.
La trame sonde de requête est destinée à tous les APs appartenant à l'ESS de la STA mais peut aussi être destiné à tous les réseaux d'une zone en utilisant le BSSID.

Quand la procédure de scanning est terminée, un rapport de scanning est généré. Il contient toutes les BSSs que le scanning a découverts et leurs paramètres. Les paramètres incluent le BSSID, SSID, BSSType, l'intervalle de balise, la période DTIM, les paramètres de timing, les paramètres PHY et CF et le BSSBasicRateSet.
L'intervalle de balise spécifie l'intervalle de chaque BSS durant lequel l'AP du BSS peut transmettre des trames balise.
Les trames DTIM sont utilisées comme mécanisme d'économie d'énergie.
Les paramètres de timing contiennent des informations de timing utilisées pour synchroniser le timer de la STA au time du BSS.
Les paramètre PHY et CF contiennent des informations de canal et d'opération sans contention.
Enfin, le BSSBasicRateSet contient une liste de débits que la STA doit supporter pour joindre le réseau.

Quand le rapport de scanning a été généré, la STA peut choisir de joindre un des BSSs. Le choix de joindre un BSS n'autorise cependant pas l'accès au réseau. La STA doit aussi passer une procédure d'authentification et d'association.

Service d'Authentification

Le processus de sélection d'un BSS est une décision basée sur l'implémentation qui peut être déclenchée par des valeurs spécifiques ou par intervention de l'utilisateur. Les critères usuels de décision sont le niveau de puissance et la force du signal.
Quand la STA a décidé de joindre un BSS, l'étape suivante est l'authentification. Il existe 2 approches d'authentification : l'authentification ouverte et l'authentification à clé partagée.
L'authentification ouverte est la seule méthode requise dans les réseaux sans fil. Elle implique que l'AP accepte la STA sans vérifier l'identité de la STA. La STA envoie une trame de

requête d'authentification à l'AP avec son adresse MAC comme unique identifiant/adresse source. L'AP traite alors la requête d'authentification et renvoie une réponse d'authentification à la STA en utilisant l'adresse source.

L'authentification à clé partagée est une méthode d'authentification optionnelle. Si cette méthode d'authentification est utilisée, les entités impliquées doivent implémenter le protocole WEP (Wired Equivalent Privacy) ou WPA (Wifi Protected Access), 2 protocoles de sécurité 802.11 qui crypte les données. L'authentification à clé partagée implique qu'une clé partagée soit distribuée à la STA avant de tenter de s'authentifier. Dans le cas de WEP, la STA envoie d'abord une trame de requête d'authentification à l'AP. l'AP renvoie alors soit une trame de refus d'authentification soit une trame de challenge d'authentification. La trame de challenge contient un texte de 128-bytes. La STA répond à la trame de challenge en cryptant le corps de la trame avec sa clé partagée et en la renvoyant à l'AP. l'AP décrypte alors la trame de challenge réceptionnée avec sa clé partagée et vérifie l'intégrité de la trame. Un message d'authentification positif est renvoyé à la STA si l'intégrité de la trame est intacte.

Une STA doit s'authentifier avec un AP avant de s'y associer. L'authentification n'est cependant pas tenue d'avoir lieu immédiatement avant l'association. Une STA peut s'authentifier avec plusieurs APs durant le processus de scanning. Ainsi la STA est déjà authentifiée quand l'association est requise. Ceci s'appelle la pré-authentification. Elle permet un gain de temps et un roaming entre APs plus aisé comparativement à l'authentification.

Service d'Association / Réassociation

Quand une STA s'est authentifiée ou pré-authentifiée à un AP, elle peut s'associer ou se réassocier à l'AP pour avoir accès au réseau.
L'association est une procédure d'enregistrement de la localisation de la STA, qui est utilisée par le DS pour transmettre les trames destinées à la STA au bon AP. une STA ne peut s'associer qu'avec un AP à la fois. La procédure d'association est initiée par la STA, qui envoie une trame de requête d'association à l'AP. L'AP traite alors la requête en fonction de paramètres dépendant de l'implémentation. Il n'existe aucune spécification sur la façon de déterminer si une association doit être permise. Généralement, la place requise pour la mise en tampon des trames dans l'AP est considérée. Si l'AP autorise la requête d'association, il répond avec un message d'association positive contenant un AssociationID (AID). L'AID est employé pour identifier logiquement la STA quand des trames mises en tampon doivent être remises à la STA. Enfin, l'association est terminée et l'AP peut traiter des trames pour la STA.

La réassociation est le processus de déplacer une association d'un ancien AP vers un nouvel AP. il diffère de l'association dans le sens où les APs interagissent. La réassociation est initiée par la STA. La STA surveille la qualité du signal de son AP actuel de même que la qualité du signal des autres APs dans le même ESS. Si la STA décide qu'un autre AP est un meilleur choix, la STA entame la réassociation. La décision d'effectuer une réassociation est basée sur des facteurs dépendant des fabricants. La STA envoie une requête de réassociation au nouvel AP contenant l'adresse de l'ancien AP. le nouvel AP communique ensuite avec l'ancien AP pour déterminer si une association précédente existait. Si l'ancien AP ne vérifie pas qu'il a authentifié la STA, le nouvel AP renvoie une trame de désauthentification à la STA et clôt la procédure. Sinon, le nouvel AP répond avec un message de réassociation positive contenant un AID. Le nouvel AP contacte ensuite l'ancien AP et termine la procédure de réassociation. L'ancien AP envoie les trames mises en tampon pour la STA vers le nouvel AP et supprime

son association avec la STA. Le nouvel AP commence à traiter des trames pour la STA et la réassociation est terminée.
Il faut remarquer que durant le processus de réassociation, la STA n'est associée qu'avec un AP à la fois durant toute la procédure.

Connexion PS

Une fois que la STA est associée/réassociée avec un AP, elle peut commencer à envoyer et recevoir des trames de données aussi connues comme Mac Protocol Data Units (MPDUs).

Service MSDU Delivery

Pour envoyer des trames de données en configuration infrastructure, toutes les trames doivent passer par l'AP, y compris les trames à destinations d'autres STAs dans la même service area. Les STAs utilisent le service MSDU (MAC Service Data Unit) Delivery pour envoyer les trames de données. Ce service définit 2 fonctions de coordination : la distributed coordination function (DCF) et la point coordination function (PCF). La DCF est une fonction de distribution obligatoire pour les réseaux infrastructure et ad-hoc alors que la PCF est une fonction optionnelle pour les réseaux infrastructure seulement.

Méthode d'accès au medium

La DCF est une méthode d'accès fondamentale que les STAs doivent supporter. Elle supporte le transfert asynchrone de données basé sur le best effort. La DCF est basée sur le principe de contention, ce qui signifie que toutes les STAs accèdent au même medium de transmission. La DCF est aussi basée sur le principe CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance). La détection de porteuse (Carrier Sense) implique que les STAs surveillent le medium pour déterminer s'il est libre ou occupé.
Deux types de détection de porteuse sont spécifiées : la détection physique de porteuse et la détection virtuelle de porteuse.
La détection physique de porteuse est fournie par la couche physique, qui détecte la présence d'autre utilisateurs sur le medium en analysant les paquets détectés et en mesurant la puissance du signal des autres sources.
La détection virtuelle de porteuse est fournie par le Network Allocation Vector (NAV). Le NAV est un timer fixé par les STAs spécifiant la durée durant laquelle le medium sera réservé. Quand le NAV égale 0, la détection virtuelle de porteuse indique que le medium est libre. Le medium est marqué occupé si l'un ou l'autre mécanisme de détection de porteuse indique que le medium est occupé.

Intervalles de temps inter-trames

Si le medium est libre, la STA initie les préparatifs de transmission. L'accès au medium est contrôlé par l'usage d'intervalles de temps d'espacement inter-trame (IFS). Les intervalles IFS sont spécifiés dans 3 classes de priorité différentes : short IFS (SIFS), PCF-IFS (PIFS) et DCF-IFS (DIFS). SIFS est l'intervalle de temps le plus court et permet donc l'accès de la plus haute priorité au medium suivi par PIFS et DIFS.
Pour l'accès basique (DCF), la STA attend un intervalle de temps DIFS et ensuite teste à nouveau la disponibilité du medium. Si le medium est toujours libre, la STA fixe un champ de durée dans la trame de données et la transmet ensuite. Le champ de durée est utilisé pour faire connaître à toutes les STAs du BSS pendant combien de temps le medium sera occupé et pour que les STAs ajustent leur NAV. Le champ de durée inclut le temps prévu pour la transmission de la trame, l'intervalle de temps IFS pour la trame d'acquittement.

Acquittement positif

Quand la STA réceptrice reçoit la trame de données, elle calcule la somme de contrôle pour vérifier si la trame a été correctement reçue. La STA réceptrice attend un intervalle de temps SIFS et transmet ensuite l'acquittement (ACK) à la STA source. Etant donné que l'intervalle de temps SIFS de la plus haute priorité est inférieur à l'intervalle DIFS, la trame ACK n'entrera pas en collision avec des trames de données. Si la STA source ne reçoit pas la trame ACK, la transmission est considérée perdue et la STA doit retransmettre la trame de données en tentant à nouveau d'accéder au medium.

Si une STA source désire transmettre une trame de données mais que le medium est occupé, la STA attend jusqu'à ce que le medium devienne libre pendant un intervalle de temps DIFS et génère un timer de retenue aléatoire pour programmer une nouvelle tentative de transmission. La STA décrémente alors le timer de retenue jusqu'à ce que le medium devienne occupé à nouveau ou que le timer atteigne zéro.
Si le medium devient occupé et que le timer n'a pas atteint zéro, la STA gèle son timer. Le timer est réactivé lorsque le medium devient à nouveau libre pendant un intervalle DIFS.
Quand le timer atteint zéro, la STA transmet sa trame de données.

Si plusieurs STAs désirent transmettre des données quand le medium est occupé, elles attendent toutes un intervalle de temps DIFS et génère un timer de retenue aléatoire. La STA avec le timer de retenue le plus petit transmettra la première et les autres STAs gèleront leur timer.
Dans le cas où les timers de deux STAs atteignent zéro simultanément, une collision se produira quand elles transmettront leurs trames respectives. Les deux STAs devront alors générer un nouveau timer de retenue.

Fragmentation et réassemblage

Les réseaux sans fil ne peuvent traiter de longues transmissions de données vu le taux d'erreur relativement grand inhérent à l'utilisation du medium sans fil. Les trames de données dépassant un certain seuil de fragmentation sont, dès lors, découpés en plusieurs fragments afin d'améliorer la fiabilité de transmission. Une STA transmettant des trames de données fragmentées doit attendre un intervalle de temps SIFS entre la transmission de deux trames successives. Cette STA a donc une priorité supérieure d'accès au medium par rapport aux autres STAs qui doivent attendre un intervalle DIFS. L'emploi de l'intervalle SIFS pour des trames de données fragmentées assure donc une transmission ininterrompue des fragments de données.

Mécanisme RTS/CTS

La DCF peut être améliorée en implémentant les trames de contrôle Request to Send (RTS)/Clear to Send (CTS). Une STA ne peut percevoir si une collision a eu lieu en transmettant. De ce fait, elle continue de transmettre la trame de données en cours de transmission. Spécifiquement, dans le cas de trames de grandes quantités de données, une part de la bande passante peut être gaspillée.
Si, avant d'envoyer une trame de données, la STA source envoie une trame de contrôle RTS après l'intervalle de temps DIFS à la STA destinataire, alors les autres STAs peuvent ajuster leur NAV, réduisant ainsi les collisions potentielles. La STA destinataire répond avec une trame de contrôle CTS après un intervalle de temps SIFS. La STA source attend ensuite un autre intervalle de temps SIFS puis débute la transmission. Comme pour toute transmission de trame de données, la STA destinataire répond avec une trame ACK après un intervalle SIFS.

L'avantage de l'implémentation RTS/CTS est que seules des collisions de trames RTS ou CTS peuvent se produire. Le gaspillage de bande passante est nettement moindre que lors de collisions de grandes trames de données.

PCF

La PCF est une méthode d'accès optionnelle. Vu que le point coordinator (PC) est implémenté dans l'AP, la PCF est limitée aux réseaux infrastructure. Par opposition à la DCF, la PCF supporte le transfert de données sensible au délai (voix et vidéo par paquets, diffusion audio/vidéo, etc.). La PCF permet la transmission sur le medium sans contention. Le PC contrôle tout le trafic en interrogeant les STAs à tour de rôle. Après avoir été interrogée seulement, une STA est autorisée à transmettre une trame.

Quand la PCF et la DCF coexistent, la capacité de transmission est partagée entre le trafic sans contention durant une période libre de contention (CFP) et le trafic avec contention durant une période de contention (CP).
L'alternance des périodes de service libre de contention et des périodes de service avec contention se répète à intervalle régulier appelé l'intervalle de répétition.

La CFP débute quand le medium est resté libre pendant un intervalle de temps PIFS. Le PC (implémenté dans l'AP) transmet une trame balise incluant la durée maximale de la CFP pour aider les STAs à se synchroniser à l'AP et mettre à jouer leur NAV à la durée maximale de la CFP. Le PC commence ensuite à interroger les STAs de sa liste de participation sur leur désir de transmettre (polling) en envoyant une trame de participation. Les STAs ne sont autorisées à transmettre que si elles ont reçu une trame de participation et consécutivement à un intervalle de temps SIFS. Les STAs peuvent « piggybacker » un ACK de la trame de participation réceptionnée avec une trame de données.
La CFP s'achève quand toutes les STAs ont été interrogées. Le PC annonce cela en envoyant une trame CF End. Les STAs remettent à zéro leur NAV et la CP débute conformément à la DCF.
La CFP peut être raccourcie si peu de STAs y participent et fournir la bande passante restant à la CP.

Quand l'AP reçoit les trames de données, il transfert les trames à la STA destinataire si elle est située dans le même BSS. Si les trames de données sont destinées à une STA d'un autre BSS mais dans le même ESS, le DS transfert les trames à l'AP approprié, qui à son tour transmet les trames à la STA destinataire. Enfin, si les trames de données sont destinées à une STA dans un réseau hors de l'ESS, les trames sont transférées à l'AP. Le DS les achemine jusqu'au routeur qui lui les transmet vers l'Internet, où les paquets sont routés vers la STA destinataire en utilisant des mécanismes de routage standards.

Quand le DS reçoit une trame destinée à une STA, soit depuis le routeur soit depuis un AP de l'ESS, il transfert cette trame au bon AP. l'AP relaye ensuite la trame à la STA destinataire. Si la STA destinataire est en mode d'économie d'énergie (expliqué après), l'AP met en tampon les trames. [7][8][9]

Inactivité PS

L'avantage principal des réseaux sans fil est la mobilité. La mobilité implique cependant que les STAs soient alimentées par des batteries. Pour contrer le gaspillage de cette ressource, les réseaux sans fil permettent de désactiver le l'émetteur/récepteur (transceiver). Quand le transceiver est éteint, la STA est en mode veille, d'économie d'énergie ou inactif.

Quand le transceiver est allumé, la STA est active. L'économie d'énergie optimale dans les réseaux sans fil est obtenue en laissant au maximum le transceiver éteint et un minimum de temps allumé, cela sans sacrifier la connectivité.
L'AP joue un rôle essentiel dans la gestion d'énergie des STAs. Tout d'abord, il est supposé être alimenté électriquement constamment vu qu'il doit rester accessible en permanence.
Ensuite, l'AP est par définition conscient de la localisation des STAs (via le DS) et a accès à l'état de gestion d'énergie des STAs. Ce rôle essentiel mène à deux taches de gestion d'énergie pour l'AP. étant donné que l'AP est conscient de l'état de gestion d'énergie des STAs qui lui sont associées, il peut aussi déterminer si une trame doit être transmise à la STA si elle est active ou mise en tampon si la STA est en mode d'économie d'énergie.
L'AP annonce périodiquement les STAs ayant des trames mises en tampon et en attente de transfert vers elles. Cette annonce du statut du tampon aide à contribuer à l'économie d'énergie. Elle nécessite moins d'énergie pour allumer le récepteur de la STA et écouter le statut du tampon que de transmettre périodiquement des trames d'interrogation à l'AP.
Durant le processus d'association, la STA et l'AP se mettent d'accord sur un intervalle d'écoute. L'intervalle d'écoute est le nombre de périodes balises durant lesquelles la STA peut passer en mode d'économie d'énergie.

Quand l'intervalle d'écoute s'est écoulé, la STA doit s'allumer pour passer en mode actif et écouter le statut du tampon. Si la STA ne repasse pas en mode actif, l'AP peut supprimer les trames destinées à la STA en attente dans son tampon sans avertir la STA.

Quand une STA s'allume, écoute le statut du tampon et découvre que l'AP a des trames en attente pour elle, la STA emploie la trame de contrôle PS-poll pour récupérer les trames du tampon. Une trame PS-poll permet la récupération d'une trame du tampon. La trame récupérée doit être ensuite acquittée par la STA avant d'être supprimée du tampon de l'AP et également avant que la STA ne puisse récupérer la trame en attente suivante. La STA doit rester active jusqu'à la récupération de toutes les trames en tampon dans l'AP qui lui sont destinées.

Détachement du réseau

Quand une STA ne nécessite plus les services de l'AP, elle peut mettre fin à l'association existante en utilisant le service de désassociation. La désassociation est une méthode polie pour supprimer l'association, cependant, le réseau est capable de supporter les STAs quittant le réseau sans désassociation préalable. Quand la STA invoque le service de désassociation, toutes les trames en tampon dans l'AP sont supprimées.
Le service de désassociation peut aussi être employé par l'AP pour informer la STA que l'AP ne lui fournit plus l'accès au réseau, lors de ressources réseau limitées par exemple ou lorsque l'AP s'éteint ou est déconnecté du réseau.
Le processus de désassociation consiste en l'envoi d'une trame d'annonce et peut être invoqué par les deux parties associées, aussi bien par la STA que l'AP. Aucune des deux parties ne peut refuser la fin de l'association.
Si une STA désire quitter un BSS, elle peut envoyer une trame de désauthentification à l'AP pour annoncer à l'AP qu'elle quitte le réseau. Une fois la STA désauthentifiée, elle n'a plus accès au réseau étant donné que le processus de désauthentification met fin à toute association existante. Pour avoir à nouveau accès aux ressources du réseau, la STA doit de ce fait effectuer une fonction d'authentification.
Le service de désauthentification peut aussi être utilisé par l'AP pour empêcher un utilisateur précédemment autorisé à utiliser le réseau.

Le processus de désauthentification consiste en l'envoi d'une trame qui peut être invoquée par les deux parties associées. Aucune des deux parties ne peut refuser la désauthentification. [8][9]

Gestion de Mobilité

La spécification 802.11 traite la mobilité de façon simple par rapport à UMTS. Il n'y a pas de distinction entre gestion de localisation et gestion de handover. Pour illustrer le fonctionnement de la gestion de mobilité, cette partie explicite les différences par rapport à la gestion de la mobilité dans UMTS.

Gestion de localisation

La gestion de localisation dans 802.11 diffère fortement de la gestion de localisation UMTS. Quand une station est associée à un AP de BSS ou ESS, le DS connaît la position de la station dans la basic service area ou l'extended service area. Pour peu que la station demeure dans la basic service area (cas BSS)/extended service area (cas ESS), la station est capable de transmettre et recevoir des trames à un AP.

Gestion de handover

802.11 gère le handover en terme de transitions. Il existe 3 types de transitions différentes : aucune transition, transition BSS et transition ESS.

Aucune transition : tant que la station reste dans l'AP service area, elle n'effectue aucune transition.

Transition BSS : si une station, initialement située dans la basic service area d'un AP1 et associée à cet AP1, sort de cette basic service area pour entrer dans la basic service area de l'AP2 de l'ESS. La station utilise alors le service Reassociation pour s'associer avec l'AP2 qui commence alors à envoyer des trames vers la station. La transition BSS demande une communication entre APs via le protocole IAPP. En effet, lors de la Reassociation, AP2 doit signaler à AP1 que la station lui est à présent associée.

Dans ce type de transition, les basic service area des APs doivent se superposer en partie pour assurer la mobilité des stations.

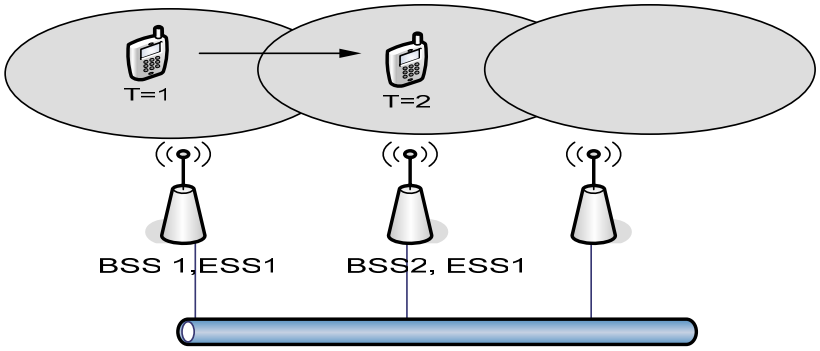

Figure 4 – Transition BSS [8]

Transition ESS : une ESS transition correspond au mouvement d'une station d'un ESS1 vers un ESS2 distinct. 802.11 supporte ce type de transition dans le sens où la station peut s'associer à un AP de l'ESS2 en quittant l'extended service area de ESS1 mais aucune garantie n'est faite quant au maintien de la connexion. Dans la pratique, la connexion est supposée se couper. Ceci signifie que les connexions de couche réseau et supérieures sont rompues. Afin de conserver les connexions de couche réseau, le recours à des protocoles de mobilité est requis.

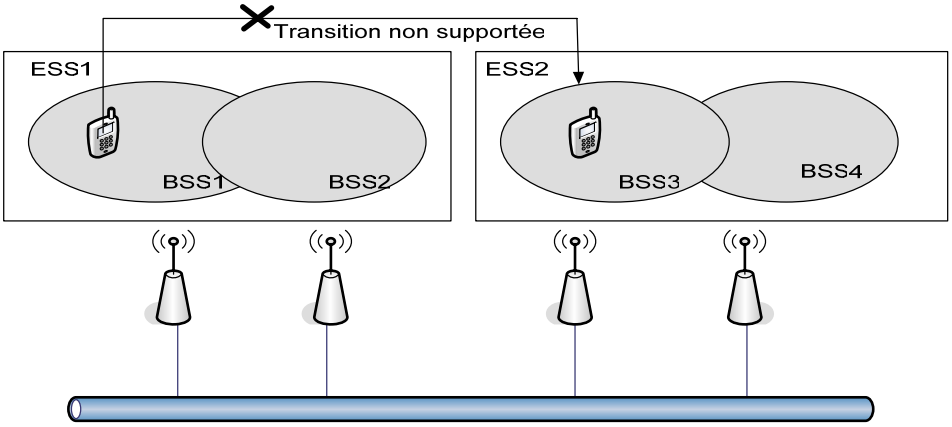

Figure 5 – Transition ESS [8]

Comparaison UMTS/802.11

Les chapitres précédents ont décrit les caractéristiques essentielles des réseaux UMTS et 802.11 en terme d'architecture réseau, de scénarios d'utilisation et de gestion de mobilité.
Ce chapitre détaille les points communs et les différences entre les deux technologies réseau.

Services supportés

La première différence entre les deux technologies réseau, spécifiquement du point de vue utilisateur, est la variété des services supportés. Le standard UMTS supporte les services à commutation de paquet et de circuit tels que la voix, la vidéoconférence, les jeux vidéos, le streaming audio et vidéo, SMS, MMS, email, fax, telnet, jeux interactifs, la navigation web, ftp, etc. alors que 802.11 ne supporte que les services correspondants dans un contexte à commutation de paquet. [3][8]

Débits

La seconde différence concerne les débits de données entre les deux technologies. UMTS supporte des débits allant de 144kbps à 2Mbps voire 10Mbps en transmission HSDPA en fonction des caractéristiques de l'environnement et de la vitesse de déplacement du mobile. Les utilisateurs à haute mobilité, classés comme utilisateurs se déplaçant à plus de 120km/h avec un maximum de 500km/h dans les zones rurales et satellitaires peuvent espérer des débits de 144kbps. Les utilisateurs à mobilité totale, utilisateurs se déplaçant à moins de 120km/h et en environnement urbain outdoor peuvent espérer un débit de 384kbps. Enfin, les utilisateurs peu mobiles, utilisateurs indoor ou se déplaçant à moins de 10km/h, peuvent espérer des débits jusqu'à 2Mbps et 10Mbps avec HSDPA. [3][10][41]
Par comparaison, 802.11 supporte des débits allant de 1Mbps à 54Mbps également en fonction des caractéristiques de l'environnement. Pour une transmission à 1Mbps, la portée maximale indoor est évaluée à 100mètres alors que la portée maximale outdoor en ligne directe est évaluée à 450mètres. Pour un débit de 54Mbps, les portées estimées maximales sont de 30mètres en indoor et 100mètres en outdoor. [8][11][41]

Couverture réseau

La troisième différence fondamentale entre les deux technologies est la couverture globale. UMTS repose sur la technologie cellulaire. Les zones couvertes par les NodeBs sont dépendantes de l'environnement. On peut citer les NodeBs parapluie couvrant de grandes cellules de plusieurs Kms de diamètre, et à l'opposé les NodeBs situés indoor couvrant typiquement de petites cellules de quelques mètres de diamètre. Le territoire est partitionné en cellules plus ou moins grandes. De plus, par l'usage de fréquences de canaux différentes entre cellules, les cellules peuvent se recouvrir les unes les autres.
La possibilité de dimensionner les cellules associées à la réutilisation de fréquences permet au réseau UMTS de supporter un grand nombre d'utilisateurs.
Par l'usage de mécanismes de handover et de gestion de roaming propres au monde cellulaire, le réseau peut fournir la couverture d'un territoire national ou même international.
Les avantages d'un réseau cellulaire sont donc une capacité d'utilisateurs importante, une couverture étendue et une distribution des cellules à la fois en fonction de l'environnement et de la charge estimée.
Par comparaison, 802.11 n'est pas un concept cellulaire. Les réseaux 802.11 fournissent des couvertures locales, telles que des entreprises, hôtels, campus, centres-villes, aéroports par des hots spots.

Contrôle de puissance

Le contrôle de puissance est fortement lié à la couverture réseau. Le réseau UMTS fournit un contrôle de puissance avancé par l'emploi de procédures d'adaptation de la puissance du NodeB et de la puissance du mobile en fonction de la distance qui les sépare et de l'environnement.
A contrario, 802.11 limite la puissance d'émission et de réception à 100mW.
De ce fait, une cellule 802.11 typique est de 50 mètres de diamètre alors qu'une cellule UMTS peut atteindre un diamètre de 35 kilomètres.

Mobilité

La mobilité est aussi un point de comparaison. UMTS traite la mobilité en effectuant des handovers ou des resélections de cellule. Ce mécanisme fonctionne tant que la technologie réseau reste cellulaire. Le hard handover inter-système permet en effet le transfert de connexion entre GSM et UMTS. UMTS fournit donc une mobilité importante et globale entre réseaux UMTS et entre réseaux cellulaires. [3][4]
Par comparaison, 802.11 traite la mobilité différemment. 802.11 supporte la mobilité dans un même BSS et aussi entre BSSs d'un même ESS. Cependant, la mobilité entre plusieurs ESSs n'est pas possible. De ce fait, 802.11 fournit une mobilité faible et locale. [7]
Ni UMTS ni 802.11 ne définissent de mécanisme pour supporter le handover avec 802.11 et UMTS, respectivement. [3][4]

Coût de déploiement

Les coûts de déploiement des deux technologies diffèrent fortement.
Le déploiement UMTS nécessite toute l'infrastructure UTRAN et un NodeB onéreux pour chaque cellule. La licence à payer (plusieurs dizaines de millions €) ainsi que la réservation de fréquences pour l'opérateur sont extrêmement coûteuses. De plus, la licence UMTS est sujette à des règlementations.
A l'opposé, le déploiement 802.11 requiert des AP peu coûteux et n'est sujet à aucune licence ni de réservation de fréquences ni de règlementations. [12]

A ce propos, les organismes responsables du développement des deux technologies réseau sont différents. UMTS est défini par un organisme de standardisation fermé décrit dans l'introduction. Les régulateurs nationaux sont eux chargés de l'administrations des fréquences radio et de vérifier le respect de la licence.
802.11 est défini par un organisme de standardisation ouvert (IEEE) et le régulateur national n'est chargé que de vérifier le bon usage des fréquences. [12]

Origine technologique

Une différence fondamentale entre UMTS et 802.11 est l'origine technologique. UMTS provient de la longue tradition du secteur des télécommunications où les implémentations ont tendance à suivre une approche de développement globale (ex : GSM) – englobant les spécifications des nœuds, des protocoles, des interfaces…
Par comparaison, 802.11 provient du secteur plus jeune des communications paquet où les modèles d'innovation et de diffusion sont qualifiables de moins coordonnés (ex : Internet).

Différences physiques

Couche physique UMTS

UMTS emploie l'interface air standardisée WCDMA, basée sur le Code Division Multiple Access (CDMA).
UMTS implémente une technique CDMA appelée Direct Sequence CDMA qui étend le flux de bits de données utilisateur sur une large bande passante en multipliant chaque bit de données par une séquence de 8 bits appelés chips. Ces séquences de chips sont dérivées de codes d'étalement CDMA.

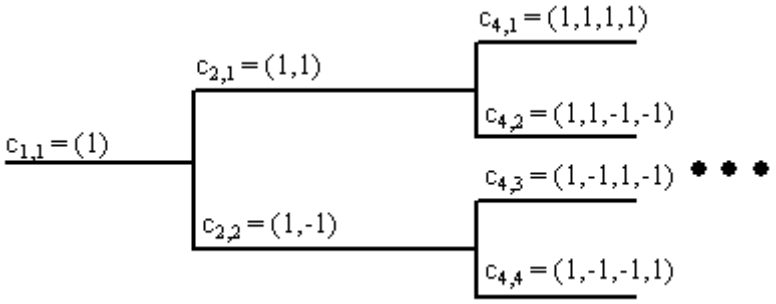

Figure 6 – Arbre de codes d'étalement [47]

Un chip est un digit binaire utilisé par le processus d'étalement. Pratiquement, il n'y a aucune différence entre chip et bit. Les bits sont des données de plus haut niveau alors que les chips sont des nombres binaires utilisés dans le processus de codage.
Le résultat de la multiplication de chaque bit du flux de bits de données utilisateur avec une séquence de chips est un flux de chips avec une amplitude étalée sur un canal occupant une large bande de fréquences. Alors que d'autres systèmes emploient une bande passante de canal de 1MHz, le système UMTS emploie une bande passante de canal de 5MHz, d'où l'appellation Wideband CDMA.
Le flux de chips est transmis simultanément avec d'autres flux de chips dans la même bande de fréquences en trames de longueur 10ms et à un débit de transmission de 3.84Mcps.
Ce débit de chips se traduit en débit de données allant de 144kbps à 2Mbps.
Au récepteur, les bits de données utilisateur sont récupérés du flux de chips à l'aide d'un corrélateur, qui inverse le processus d'étalement.

UMTS opère dans un spectre de fréquence régulé autour de 2GHz. Vu que UMTS supporte 2 modes duplex de fonctionnement : Frequency Division Duplex (FDD) et Time Division Duplex (TDD), l'allocation de spectre en Europe consiste pour le WCDMA FDD aux bandes 1920MHz-1980MHz (sens montant) et 2110-2170MHz (sens descendant) et pour le WCDMA TDD aux bandes 1900-1920MHz et 2010-2025MHz.
Une nouvelle bande de fréquences 2500-2690MHz a été également réservée pour l'usage WCDMA mais n'est pas encore d'application.
Pour l'usage de canaux montant et descendant groupés, un canal de 5MHz montant (dans la bande 1920-1980MHz) et un canal de 5MHz descendant (dans la bande 2110-2170MHz) sont associés. Un total de 12 paires de canaux peut ainsi être défini.
Pour l'usage de canaux non groupés dans les bandes de fréquence 1900-1920MHz et 2010-2025MHz, 7 canaux sont disponibles. [3][4]
Voir Annexe UMTS pour des précisions sur WCDMA.

Couche physique 802.11b

La spécification 802.11b emploie le principe du High Rate Direct Sequence (HR-DS). HR-DS est développé à partir de la méthode de codage 802.11 DS. La méthode DS consiste à étaler le flux de bits de données utilisateur par l'application d'un mot 11-bit Barker, une séquence de bits définie. Chaque bit est codé en additionnant modulo-2 cette séquence Barker. Un bit 1 sera donc représenté par la séquence 01001000111 et 0 par 10110111000. La transmission de flux de chips s'effectue à un débit de 11Mcps. La bande de fréquences 2.412-2.472GHz est divisée en 13 canaux de 5 MHz.

De façon similaire à UMTS, l'étalement d'amplitude entraîne des canaux occupant une large bande de fréquences. L'étalement est engendré sur un canal de 22 MHz. On observe donc un recouvrement partiel des canaux adjacents et seuls 3 canaux sur les 13 sont entièrement isolés les uns des autres. A la réception, le signal est recorrélé c-a-d que le récepteur inverse le processus d'étalement du canal de 22MHz vers un canal de 5MHz.

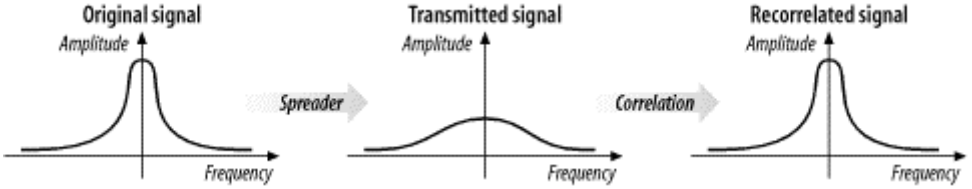

Figure 7 – Etalement DS de 5MHz vers 22MHz et corrélation [8]

La méthode HR-DS emploie les mêmes canaux que DSSS avec une technique de modulation différente. La transmission peut s'effectuer à un débit de 11Mbps.

Pour ce faire, les flux de chips sont dérivés partiellement des données à transmettre, la séquence Barker n'étant plus utilisée.

Voir Annexe 802.11 pour de plus amples informations sur les différentes couches physiques 802.11.

Protocoles de mobilité

Comme mentionné dans les chapitres précédents, UMTS et 802.11 n'ont aucune fonctions propres pour effectuer un handover inter-système entre les technologies UMTS et 802.11.
De ce fait, un protocole de mobilité est nécessaire pour supporter ce type de handover.
Les protocoles de mobilité se présentent à différentes couches du modèle OSI. Chaque couche du modèle a ses fonctions et responsabilités distinctes.

Protocole IP

Le protocole Internet version 4 (IPv4) est un protocole réseau fondamental qui définit les informations d'adressage et des informations de contrôle permettant aux paquets de données d'être routés à travers le réseau. Les informations d'adressage consiste en un identifiant unique de 32-bits appelé adresse IP. Chaque client sur l'Internet est identifié par une adresse IP unique. Les informations de routage contiennent l'adresse IP de destination et des informations supplémentaires employés pour trouver une route entre la source et la destination. Afin qu'un client reste identifiable pour d'autres hôtes, il doit conserver son adresse IP. Cependant, si l'adresse IP est conservée, le routage jusqu'au client est aussi conservé et le cheminement des paquets reste toujours le même vers le client. Ceci empêche le client de se déplacer et d'être capable de recevoir des paquets en utilisant le protocole IPv4.

Parmi les protocoles de mobilité les plus employés, on trouve les protocoles de couche réseau Mobile IP, le protocole de couche transport Mobile Stream Control Transmission Protocol (mSCTP) et le protocole de couche application Session Initiation Protocol (SIP).
Ces protocoles sont explicités et comparés par la suite.

Mobile IP

Mobile IP version 4 (Mobile IPv4) est un protocole de couche réseau supplémentaire qui résout le problème inhérent à IPv4 vis-à-vis de la mobilité. Mobile IPv4 permet au client d'utiliser deux adresse IP, une comme identifiant unique et l'autre pour le routage spécifique à sa localisation dans le réseau. L'adresse de domicile est une adresse statique employée comme identifiant unique, et la care-of adresse est une adresse qui change lorsque le client passe d'un domaine à un autre. Les couches de transport et supérieures n'emploient que l'adresse de domicile, ce qui laisse les protocoles de ces couches ignorants d'une quelconque mobilité du client, et de ce fait, les connexions TCP restent actives.
Mobile IPv4 fournit donc la transparence pour les couches supérieures en fournissant la mobilité intégrée au moyen de care-of adresses. [15][16][17]
Quand le client est dans son réseau domicile, les mécanismes de routage IP standards acheminent les paquets entrants et sortants depuis et vers le client au moyen de l'adresse de domicile.
Si le client change de localisation durant la communication de son réseau domicile à un autre réseau appelé réseau étranger, les mécanismes de routage IP standards ne suffisent pas, vu que le client n'est plus joignable par son adresse domicile. Il doit alors obtenir et enregistrer une care-of adresse au moyen de Mobile IPv4 pour continuer sans interruption sa communication. [15]

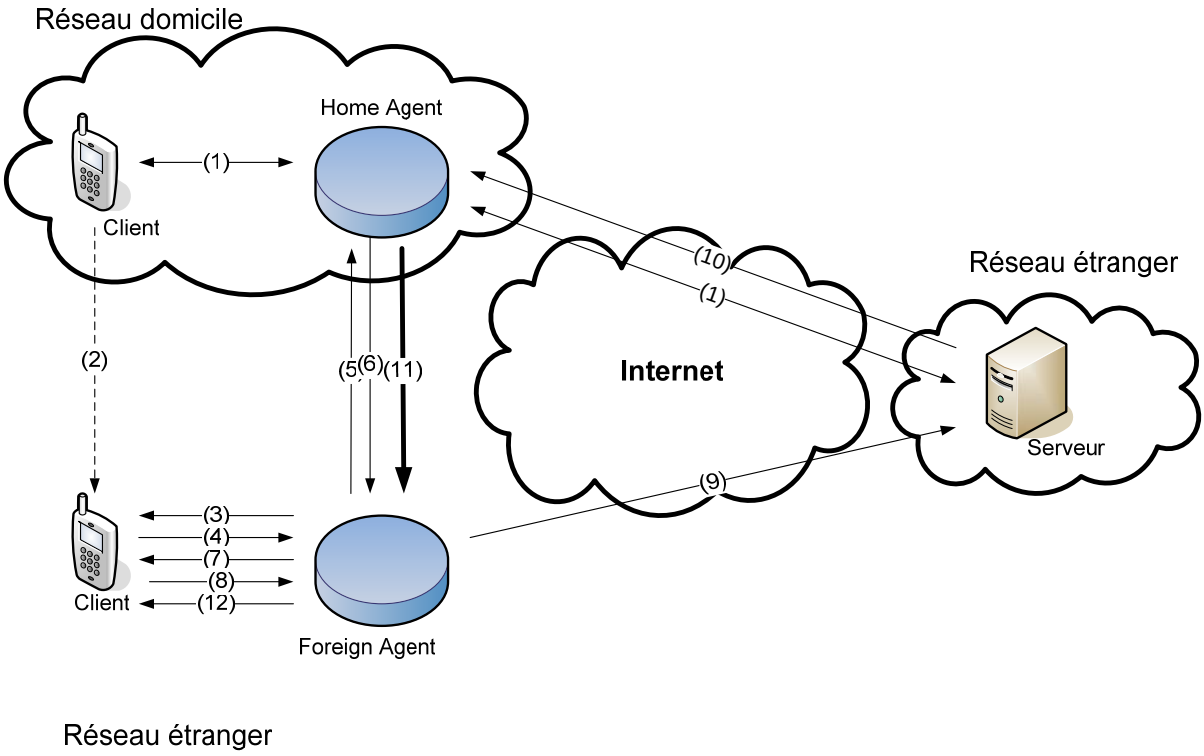

Figure 8 – Gestion de mobilité Mobile IPv4 [15]

Le client est tout d'abord localisé dans son réseau domicile à une position A où il a établi la communication avec un serveur dans un réseau étranger à travers l'Internet (1). Le client change alors de position de A dans le réseau domicile vers B dans un réseau étranger (2).

Afin de conserver la communication, Mobile IPv4 entre alors en action.

Mobile IPv4 emploie deux nouvelles entités réseau : un home agent dans le réseau domicile et un foreign agent dans le réseau étranger. Ces deux agents sont des nouvelles entités réseau dans le sens où les routeurs existant nécessitent le support du protocole Mobile IPv4 et non seulement de IPv4.

Les foreign agents tout comme les home agents annoncent leur disponibilité en implantant une extension spécifique à la fonction d'annonce du routeur. Les annonces sont généralement diffusées à intervalles réguliers sur le réseau. Alternativement, le client peut envoyer une sollicitation au routeur pour lui demander une annonce.

Quand le client reçoit une annonce d'agent, il détermine s'il se situe dans son réseau domicile ou dans un réseau étranger (3). S'il découvre qu'il est dans un réseau étranger, le client enregistre sa nouvelle care-of adresse, trouvée dans l'annonce d'agent, avec son home agent à travers le foreign agent (4)(5). A chaque fois que le client se déplace dans un réseau étranger, il enregistre sa nouvelle care-of adresse.

Le home agent répond à la requête du client en acceptant la requête, en mettant à jour sa table de routage avec la nouvelle care-of adresse et finalement en retournant une réponse d'enregistrement au client via le foreign agent (6)(7).

La réponse contient une durée d'enregistrement, spécifiant la durée de validité de la care-of adresse. Le home agent associe l'adresse domicile du client avec la care-of adresse jusqu'à ce que la durée d'enregistrement expire. Le triplet adresse domicile, care-of adresse et durée d'enregistrement s'appelle un « binding » du client.

Une requête d'enregistrement est dès lors considérée comme une mise à jour du binding envoyée par le client et la réponse d'enregistrement comme un acquittement du binding.

Après son enregistrement réussi, la communication entre le client et le serveur peut continuer de façon ininterrompue. Quand le client envoie des paquets destinés au serveur, il les envoie au foreign agent, qui les transmet directement au serveur (8)(9).
Dans la direction opposée, le serveur envoie toujours les paquets destinés au client au réseau domicile. Le home agent intercepte les paquets destinés au client et les encapsule en ajoutant à chaque paquet une nouvelle en-tête IP (10). Ce procédé est appelé encapsulation IP-in-IP. Le nouvel en-tête IP contient entre autres l'adresse de destination, c-à-d la care-of adresse.
Le home agent tunnelle alors les paquets encapsulés vers le foreign agent au moyen de la care-of adresse (11). Le foreign agent reçoit les paquets, les désencapsule, et les transmet enfin au client (12).
La façon asymétrique de router les paquets depuis/vers le client porte le nom de routage triangulaire. [15][16][17]

Mobile IPv6

Mobile IPv4 a été initialement défini comme un ajout à IPv4. Pour le protocole IPv6, le support de la mobilité (Mobile IPv6) a été envisagé d'emblée. Dès lors, certains problèmes de Mobile IPv4 ont été résolus dans Mobile IPv6. Les problèmes majeurs de Mobile IPv4 sont le déploiement, le routage triangulaire, l'overhead de tunnelling et la sécurité. [16][18]
Chacun de ces problèmes est décrit par après de même que l'approche Mobile IPv6 résolvant ces problèmes.

Le déploiement de Mobile IPv4 nécessite l'implémentation de foreign agents dans chaque réseau étranger potentiel. Cette implémentation suggère une reconfiguration étendue du réseau. Mobile IPv6 traite ce problème en éliminant totalement les foreign agents. Il conserve les idées de réseau domicile, home agent et l'usage de l'encapsulation pour acheminer les paquets depuis le réseau domicile jusqu'au client.

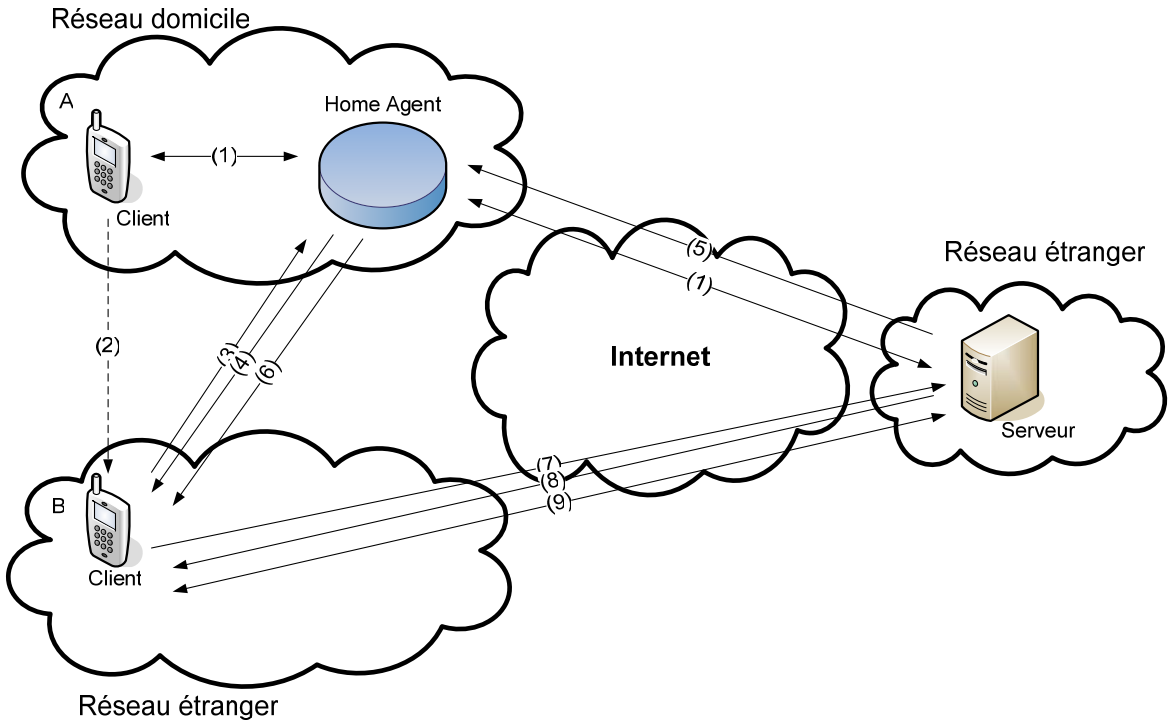

Figure 9 – Gestion de mobilité Mobile IPv6 [16]

Le scénario pour Mobile IPv6 est similaire au scénario Mobile IPv4. Le client est premièrement localisé dans son réseau domicile à une position A où il a établie la

communication avec un serveur, dans un réseau étranger à travers l'Internet, au moyen des mécanismes de routage IP standards (1).
Le client change alors de position de A dans le réseau domicile vers B dans un réseau étranger (2).

Enregistrement

Au lieu d'écouter les annonces de disponibilité des foreign agents, le client écoute les annonces de routeur. Les annonces de routeur dans IPv6 ont été étendues avec plusieurs bits. L'information de préfixe réseau IPv6 permet au routeur d'annoncer son adresse IPv6 globale au lieu de son adresse de liaison locale. Le client peut déterminer s'il se trouve dans son réseau domicile ou dans un réseau étranger à l'aide du préfixe réseau contenu dans l'annonce de routeur. Si le préfixe réseau correspond au préfixe réseau de l'adresse domicile du client, le client se trouve dans son réseau domicile. Si le client découvre qu'il est dans un réseau étranger, il obtient un care-of adresse et l'enregistre avec son home agent. Le client obtient une care-of adresse soit en contactant un serveur DHCPv6 dans le réseau étranger, soit en extrayant le préfixe réseau de l'annonce de routeur et en ajoutant un identifiant d'interface unique.
Quand le client a obtenu une care-of adresse, il envoie une mise à jour de binding à son home agent (3). Le home agent répond avec un acquittement de binding (4).
Le processus d'enregistrement de Mobile IPv6 diffère donc essentiellement par l'absence de foreign agent. [16][18][19]

Routage triangulaire

Le routage triangulaire implique que tous les paquets envoyés au client sont routés via le home agent, ajoutant un délai de transfert vers le client. Ce problème est résolu dans Mobile IPv6 en implémentant l'optimisation de route. L'optimisation de route a été initialement spécifiée comme une extension pour Mobile IPv4 et est présente d'origine dans Mobile IPv6.
Pour l'optimisation de route, le client enregistre d'abord sa care-of adresse avec le home agent comme décrit ci-dessus. Il envoie alors une mise à jour de binding directement au serveur pour lui signaler sa nouvelle care-of adresse (7). Le serveur répond avec un acquittement de binding. Le client et le serveur peuvent poursuivre leur communication de manière ininterrompue (8)(9).
Le home agent peut aussi recevoir des paquets du serveur avant que le client n'ait enregistré sa care-of adresse avec le serveur (5). Dans ce cas, le home agent reçoit les paquets du serveur, les encapsule et les transmet au client (6).
Quand le client reçoit le premier paquet encapsulé du home agent, il envoie une mise à jour de binding au serveur, qui répond au client par un acquittement de binding (7)(8).
Après cette étape, le serveur et le client poursuivent la communication sans interaction du home agent. En supprimant le home agent comme nœud intermédiaire, le délai supplémentaire dans la direction serveur-client est éliminé. [16][18][19]

Tunnelling

Quand le serveur envoie des paquets au client, les paquets transitent par le home agent qui intercepte les paquets et les encapsule. Il tunnelle ensuite les paquets encapsulés vers le foreign agent. Le tunnelling consiste typiquement en un overhead de 20-bytes ajoutés à chaque paquet (encapsulation IP-in-IP).
Mobile IPv6 résout le problème d'overhead en supprimant simplement la fonction de tunnelling. [16][18][19]

Sécurité

Enfin, il y'a des problèmes de sécurité. Quand le client enregistre une care-of adresse avec son home agent, le home agent doit être certain que la requête provient du client et non d'un nœud prétendant être le client. Un tel nœud pourrait entraîner le home agent à modifier sa table de routage de telle manière que le client ne soit plus joignable, et dans le pire des cas que les communications soient redirigées vers le nœud.
Mobile IPv4 emploie une association de sécurité entre le home agent et le client au moyen de l'algorithme Message Digest 5. Cet algorithme à clé de 128-bits crée des signatures digitales pour les demandes d'enregistrement. Mobile IPv4 ne requiert cependant pas l'authentification des foreign agents envers le client ou le home agent.
Mobile IPv6 implémente quant à lui des fonctions d'authentification et de cryptage puissantes dans tous les nœuds au moyen d'IP Security (IPSec). [16][18][19]
Au vu de ces améliorations, il semble évident de préférer l'usage de Mobile IPv6 à Mobile IPv4. Dans le futur, IPv6 est appelé à remplacer IPv4 sur l'Internet, cependant, ce n'est pas encore le cas. Cette recherche se focalisera sur Mobile IPv6 comme solution d'avenir tout en n'excluant pas Mobile IPv4.

Mobile Stream Control Transmission Protocol (mSCTP)

La mobilité de couche transport est proposée comme alternative à la mobilité de couche réseau pour le support de la mobilité intégrée. La gestion de la mobilité dans la couche transport est exclusivement effectuée par Stream Control Transmission Protocol (SCTP) et son extension Dynamic Address Reconfiguration (DAR). SCTP étendu avec DAR constitue Mobile SCTP (mSCTP).
mSCTP est un protocole de couche transport similaire à Transmission Control Protocol (TCP). Il fournit la communication point à point orientée connexion entre applications fonctionnant sur des hôtes différents. La différence majeure avec TCP est le multi-homing
mSCTP permet par le multi-homing de gérer plusieurs adresses IP aux nœuds terminaux en conservant la connexion point à point intacte. Ces adresses sont considérées comme des chemins logiques différents entre les nœuds terminaux. Durant l'initiation de la connexion, les listes d'adresses sont échangées entre les nœuds terminaux. Les nœuds terminaux doivent être capables de recevoir des messages de toute adresse IP associée au nœud terminal opposé. Une adresse est choisi comme adresse primaire et est utilisée comme adresse de destination pour la transmission normale. Les autres adresses sont employées uniquement pour les retransmissions. L'extension DAR permet aux nœuds terminaux d'ajouter, supprimer et changer les adresses IP durant une session SCTP sans affecter la connexion établie en utilisant des messages de configuration d'adresse.

Le scénario mSCTP débute avec un client localisé dans son réseau domicile à une position A où il établit la communication avec un serveur dans un réseau étranger à travers l'Internet.

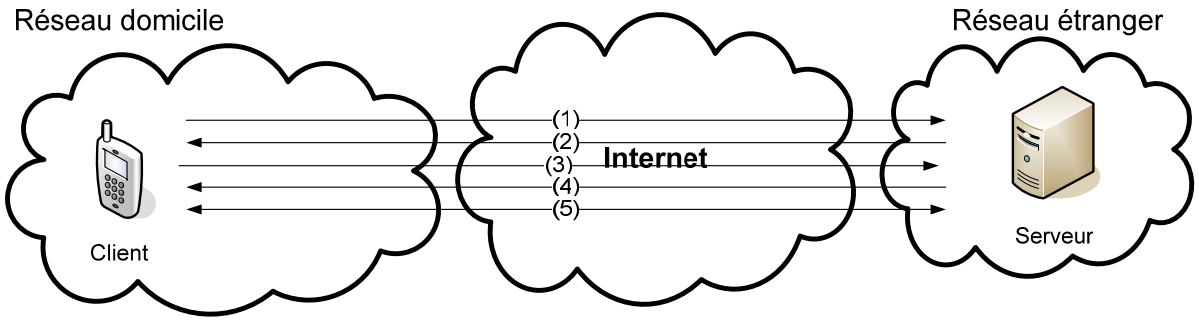

Figure 10 – Initialisation de connexion mSCTP [22]

Pour établir une connexion de couche transport avec le serveur, le client envoie d'abord une requête d'initialisation au serveur incluant une liste d'adresses IP et un numéro de port qui seront utilisés par le client (1). Le serveur répond avec un acquittement d'initialisation incluant un marqueur d'état, sa liste d'adresses IP et un numéro de port qui seront utilisées par le serveur s'il accepte la requête (2). Le client doit ensuite renvoyer le marqueur d'état au serveur pour confirmer sa réception (3). Le serveur passe alors en état d'établissement de connexion et répond avec un acquittement de marqueur (4). Enfin, le client passe en état d'établissement de connexion (5). La connexion client-serveur est alors établie et la communication peut débuter. Identiquement à TCP, mSCTP génère automatiquement un acquittement entre chaque séquence de messages. [24]

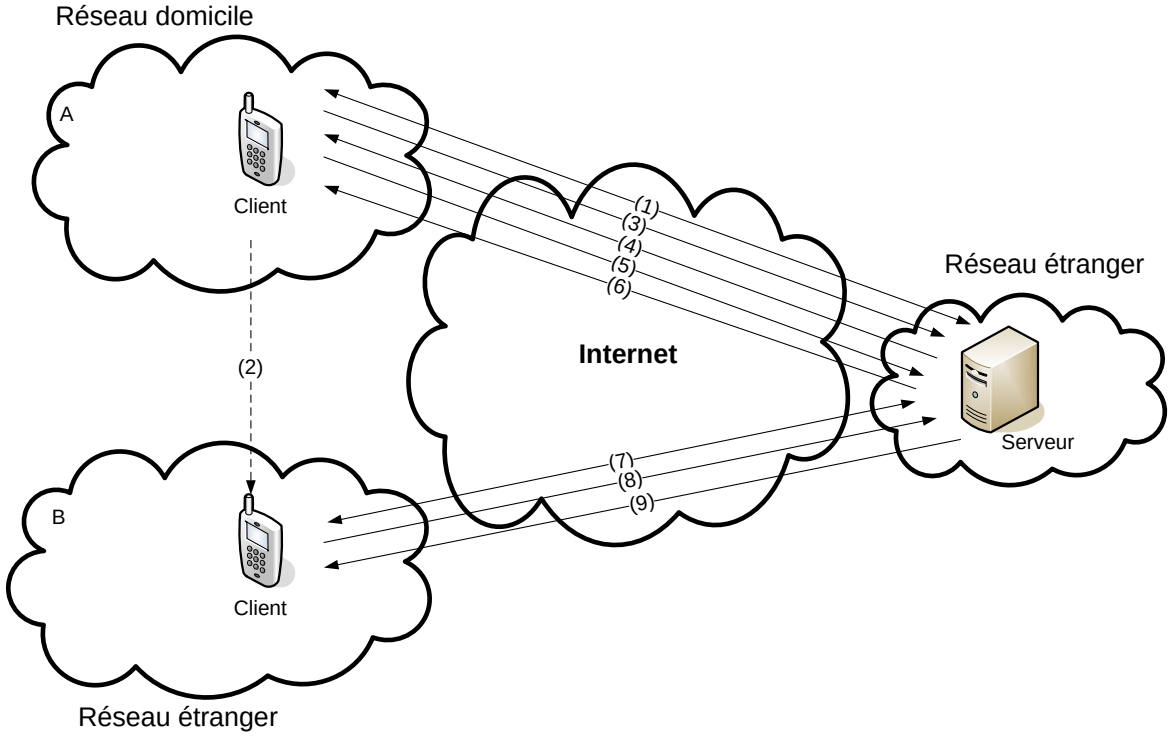

Figure 11 – Gestion de mobilité mSCTP [22]

Durant la communication (1), le client change alors de position de A dans son réseau domicile vers B dans un réseau étranger (2). Alors que le client passe dans la zone de couverture du réseau étranger, il reçoit une adresse IP du réseau étranger, soit en contactant un DHCP ou par la configuration automatique d'adresse IPv6. Le client est à présent capable d'établir une liaison avec le serveur grâce à cette deuxième adresse IP et être accessible via le réseau domicile et le réseau étranger. Le client envoie alors au serveur via son réseau domicile sa deuxième adresse IP (3). Il ajoute donc la nouvelle adresse IP à l'association identifiant la connexion avec le serveur. Le serveur lui répond par un acquittement (4). Quand le client quitte la zone de couverture de son réseau domicile, le client avertit le serveur d'assigner la nouvelle adresse IP comme adresse IP primaire (5), ce que le serveur approuve par un acquittement (6). La nouvelle adresse IP primaire devient à présent l'adresse de destination client pour la communication. Le serveur envoie à partir de ce moment tous ses messages vers la nouvelle adresse IP du client (7). Finalement, le client informe le serveur de supprimer la première adresse IP de l'association (8), ce que le serveur confirme par un acquittement (9). [21][22][24]

Session Initiation Protocol (SIP)

Le protocole de couche application SIP est un protocole de signalisation développé par l'IETF. SIP permet l'établissement, la modification et la fermeture de sessions multimédia consistant en flux de données média, unicast ou multicast. Les flux de données média incluent l'audio/vidéo, les applications partagées, etc.
Les utilisateurs SIP sont identifiés par des adresses semblables aux adresses e-mail (user@host). L'adresse SIP ne change pas lorsque l'utilisateur change de localisation. De ce fait, lors de déplacement, les appels à destination de l'utilisateur sont redirigés vers sa position courante.

Entités SIP

SIP définit quatre entités logiques, à savoir des agents utilisateur, des serveurs d'enregistrement, des serveurs de redirection et des serveurs proxy, et un service appelé service de localisation.
L'agent utilisateur a deux rôles : un agent utilisateur client émet des requêtes et reçoit des réponses ; un agent utilisateur serveur reçoit des requêtes et envoie des réponses d'approbation des requêtes.
Le serveur d'enregistrement conserve les informations d'accès des agents utilisateur basées sur les requêtes de modification de l'agent utilisateur. Ce serveur ne gère que des requêtes concernant des adresses SIP dans son domaine. Typiquement, ces requêtes concernent le changement de localisation de l'utilisateur.
Le serveur de redirection conserve la localisation de l'utilisateur et gère la redirection de contacts vers les agents utilisateurs situés hors du domaine du serveur d'enregistrement. Le serveur de redirection ne renvoie que la localisation de l'utilisateur, il ne transfère aucun message.
Le serveur proxy est responsable du transfert de messages entre agents utilisateur.
Enfin, le service de localisation est une base de données contenant les informations de localisation des agents utilisateur. Le service de localisation est employé par les serveurs proxy et de redirection pour localiser les agents utilisateur client et serveur.
Habituellement, un serveur SIP implémente un serveur de redirection et un serveur proxy avec l'information fournie par un serveur d'enregistrement intégré. Le service de localisation peut se situer soit sur le serveur SIP soit sur un serveur de localisation dédié. [27][28][29]

Scénario

Le scénario SIP débute avec un client localisé dans son réseau domicile en A où il établit la communication avec un serveur dans un réseau étranger à travers l'Internet.

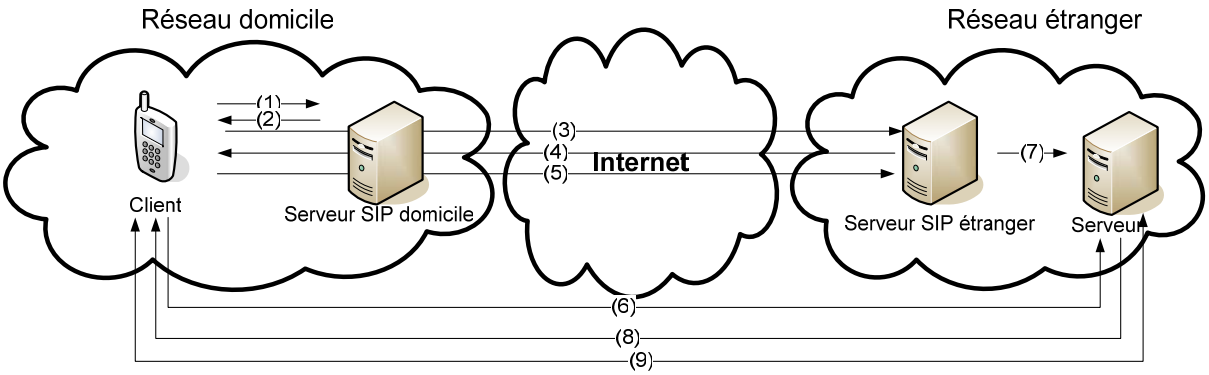

Figure 12 – Initialisation de connexion SIP [27]

Alors que le client est attaché au réseau domicile, son agent utilisateur envoie une mise à jour de localisation au serveur d'enregistrement dans le serveur SIP domicile (1). Le serveur d'enregistrement traite le message de mise à jour et le transmet au service de localisation, qui stocke l'information. Le serveur SIP domicile renvoie ensuite un acquittement (2).
Le client désire maintenant communiquer avec un serveur localisé dans un réseau étranger, son agent utilisateur envoie alors une requête d'invitation à son serveur SIP domicile. Le serveur SIP domicile reconnaît que la requête ne lui est pas adressée et la transfère au serveur SIP appartenant au réseau étranger (3). Le serveur de redirection dans le serveur SIP du réseau étranger reçoit la requête et consulte le service de localisation pour trouver la localisation du serveur. Le service de localisation renvoie l'adresse du serveur au serveur de redirection, qui à son tour l'envoie à l'agent utilisateur du client (4). L'agent utilisateur du client confirme la réception de l'adresse du serveur (5). L'agent utilisateur du client dispose à présent de l'adresse du serveur la plus récente et peut envoyer sa requête d'invitation à l'agent utilisateur du serveur (6).

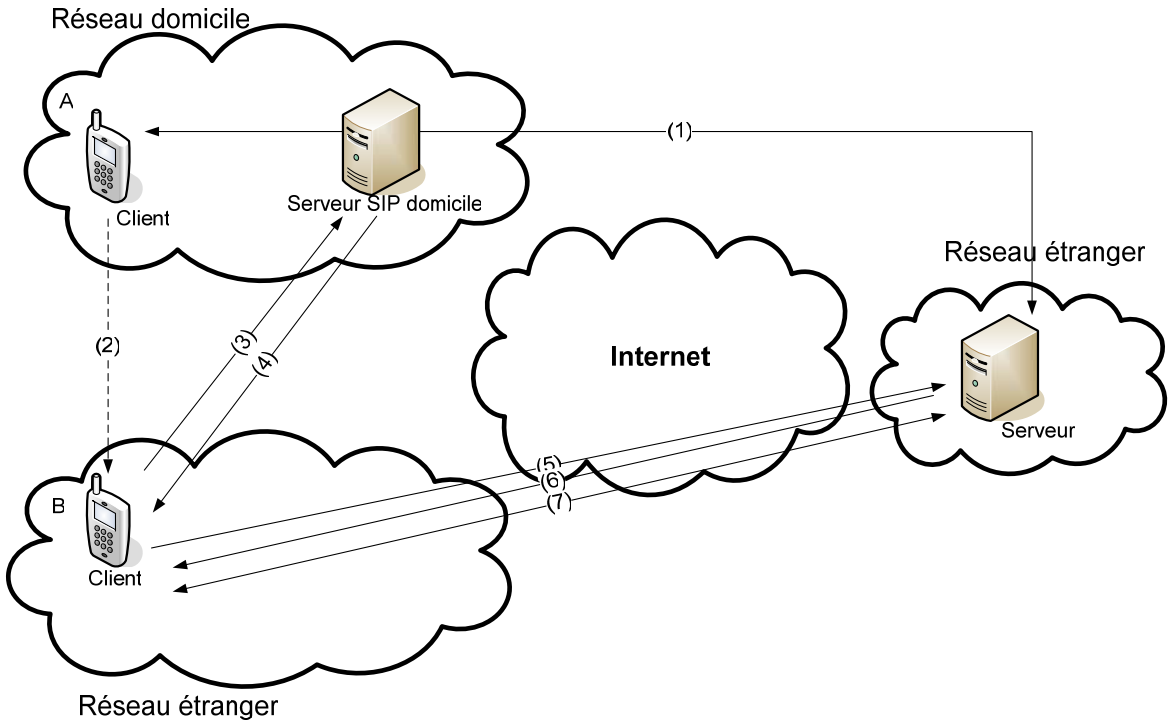

Figure 13 – Gestion de mobilité SIP [27]

Durant la communication (1), le client change à présent sa localisation de A dans le réseau domicile vers B dans un réseau étranger (2). Pour maintenir la connexion, SIP implémente une mise à jour de localisation. L'agent utilisateur du client envoie une mise à jour de localisation au serveur SIP domicile. Ainsi les nouvelles requêtes d'invitation peuvent être redirigées vers le client (3). Le serveur d'enregistrement traite le message de mise à jour et le transmet au service de localisation, qui stocke l'information. Le serveur SIP domicile répond avec un acquittement (4).
Le client envoie alors une nouvelle requête d'invitation à l'agent utilisateur du serveur en utilisant son identifiant (5). La requête contient la nouvelle adresse, informant l'agent utilisateur du serveur où il veut recevoir les futurs messages SIP. L'agent utilisateur du serveur acquitte la requête (6) et la communication continue (7) [27]

Comparaison des protocoles de mobilité

Mobile IP montre des limitations quand les applications multimédia sont sensibles au délai. Le routage triangulaire ajoute des délais de handover et l'overhead de tunnelling ajoute des bytes à l'en-tête des paquets. Mobile IP est davantage approprié pour des connexions TCP durables telles que telnet, ftp, etc.
Par opposition, SIP est plus approprié pour la communication temps réel par UDP. Il est cependant moins adéquat pour des applications basées TCP. SIP est dès lors souvent employé comme complément à Mobile IP.

Déploiement

Les protocoles de mobilité envisagés diffèrent tout d'abord par leur lourdeur de déploiement sur un réseau IP existant.
Mobile IPv4 est probablement le protocole de mobilité le plus lourd en terme de déploiement. En effet, pour assurer la mobilité du client, un home agent dans le réseau domicile et un foreign agent dans le réseau étranger sont nécessaires. Le support de Mobile IPv4 doit être assuré par les home agents et les foreign agents ainsi que par les hôtes mobiles (client / serveur).
Mobile IPv6 est similaire à Mobile IPv4 en étant moins lourd. Il nécessite un home agent dans le réseau domicile et le support de Mobile IPv6 est requis dans les home agents ainsi que dans les hôtes mobiles.
mSCTP, à contrario, ne nécessite aucune entité réseau additionnelle ou de modification des entités réseau existantes. Seul le support de mSCTP dans les hôtes mobiles est nécessaire.
Enfin, SIP nécessite la présence d'un serveur SIP additionnel dans le réseau domicile et le réseau étranger. L'implémentation du protocole est requise dans les hôtes mobiles également.
Le déploiement du protocole mSCTP ne nécessite qu'une intervention dans les hôtes mobiles ; les complexités de déploiement de SIP et Mobile IP sont comparables, ils nécessitent tous deux, la définition de nouvelles entités supportant leur protocole respectif.

Transparence

Les trois protocoles envisagés diffèrent également par transparence. La transparence est la capacité d'un protocole à ne pas altérer les couches supérieures à la sienne. Tous les protocoles sont transparents mais à différents niveaux.
Mobile IPv4 et Mobile IPv6 fournissent la transparence pour les couches transport et application. mSCTP fournit aussi la transparence pour la couche application. Enfin, SIP fournit la transparence à l'utilisateur.

Service de transport

Le concept de transparence est lié au service de transport. Les protocoles de couche réseau, à savoir Mobile IPv4 et Mobile IPv6, fournissent un service commun aussi bien pour TCP que UDP.
Le protocole de couche transport mSCTP diffère cependant en fournissant un nouveau service de transport SCTP. Au vu des implémentations software actuelles utilisant TCP ou UDP, le choix de SCTP engendrera des problèmes et nécessitera une reconfiguration des softwares.
Enfin, SIP s'appuie sur UDP comme service principal de transport. SIP n'est pas adapté pour des services basés TCP.

Conclusion

Comparatif qualitatif	Mobile IPv4	Mobile IPv6	mSCTP	SIP
Déploiement	- -	-	+ +	-
Transparence/Service transport	+ +	+ +	- -	+

Ce comparatif qualitatif basé sur deux paramètres de choix fait apparaître le protocole Mobile IPv6 comme protocole polyvalent, d'avenir et non contraignant quant au service de transport employé.

Dans la suite de ce travail, aucune comparaison ne sera faite entre ces protocoles de mobilité, de nombreux articles traitant déjà ce sujet, seul Mobile IPv6 sera envisagé pour fournir une architecture de développement.

Entité de handover

Les protocoles de mobilité permettent le support du handover entre les deux architectures. A présent, il est intéressant d'expliciter l'exécution à proprement parler du handover inter-système entre UMTS et 802.11. En effet, le protocole de mobilité autorise l'interaction de couche réseau ou supérieure entre les deux réseaux. Cependant, UMTS traite la mobilité par handover et resélection de cellule alors que 802.11 traite la mobilité par transitions.
Ce chapitre détaille les besoins pour traiter le handover inter-système ainsi que la procédure de handover.

Exigences du handover

Afin d'effectuer des handovers inter-système UMTS / 802.11, certaines exigences du terminal mobile et du réseau doivent être remplies.

Exigences du mobile

Le terminal mobile doit être un terminal bi-mode remplissant à la fois la fonction d'UE équipé d'une USIM et de STA équipée d'une carte d'accès sans fil 802.11. Ce terminal équipé des deux interfaces d'accès doit être capable d'opérer sur les deux réseaux et supporter le handover d'un réseau vers l'autre.

Exigences du réseau

L'interaction réseau implique que les réseaux UMTS et 802.11 sont interconnectés. Le réseau UMTS d'un opérateur peut servir de base pour y connecter un réseau 802.11.
Ainsi on définit trois types d'interconnexion possible : la configuration tight coupling, la configuration loose coupling et la configuration open coupling.
La configuration open coupling signifie qu'il n'y a aucune intégration entre les deux réseaux d'accès. Les réseaux 802.11 et UMTS sont considérés comme deux systèmes indépendants partageant un système de facturation entre eux. L'authentification d'un terminal mobile implique dès lors l'activation de la facturation sur une base de données commune. Cette configuration exclut le support du handover intégré.
La configuration loose coupling consiste à employer une base de données client commune AAA et une procédure d'authentification. La base de données AAA est chargée de la facturation et de l'authentification de clients UMTS et 802.11 et est liée au HLR UMTS.
Les deux réseaux sont dans cette configuration liés par le GGSN UMTS.
Enfin, la configuration tight coupling consiste à intégrer le réseau 802.11 au même niveau que le RNC UMTS. Les deux réseaux sont alors gérés tous deux par le HLR. [36]

Procédure de handover

La procédure de handover d'un terminal mobile se décompose en trois étapes. D'abord, certaines mesures doivent être effectuées et rassemblées dans un rapport de mesures. Ensuite, une décision de handover est prise en fonction du rapport. Enfin, le handover est exécuté si la décision de handover est positive.

Mesures

La première étape est la mesure de certains paramètres requis pour analyser le statut de la connexion existante entre le terminal et la cellule utilisée et le statut de la qualité d'autres cellules disponibles.
Les mesures peuvent être effectuées par le terminal ou le réseau. Pratiquement, le terminal participe toujours à la prise de mesures. [30]

Les mesures incluent à la fois des préférences statiques de l'utilisateur et des mesures dynamiques. Les préférences statiques réfèrent à une liste de services à laquelle l'utilisateur a souscrit, et une liste de préférence indiquant la priorité de services en cas de ressources faibles. Les paramètres dynamiques contiennent une liste de services supportés ou non par le réseau, une liste de services actifs ou suspendus et un indicateur sur la qualité de service délivré.
On s'intéresse principalement aux paramètres dynamiques dans le cadre de la modélisation. Ils comprennent la surveillance et analyse des paramètres d'accès réseau tels que la puissance de réception, le bit error rate, le block error rate et les informations de charge réseau obtenues en surveillant la charge en terminaux de la cellule courante et des cellules voisines. [31]

Quand les mesures sont effectuées, elles sont rassemblées dans un rapport de mesures et envoyées à l'entité de décision du handover.

Décision du handover

En fonction du rapport de mesures, l'entité de décision évalue si un handover est requis ou non. La décision du handover peut provenir d'une entité de décision du terminal (mode contrôlé terminal) ou d'une entité de décision du réseau (mode contrôlé réseau).
Dans le mode contrôlé terminal, le terminal mesure la puissance du signal de la station de base courante et des stations de base candidates. S'il perçoit un signal de puissance supérieure provenant d'une station de base candidate, le terminal initie le handover. Le réseau peut diffuser des paramètres pour influencer ce processus, cependant la décision du handover réside dans le terminal.
Le mode contrôlé terminal est un mode de décision décentralisé. L'avantage en est une architecture de handover simple, extensible et tolérante. De plus, dans un contexte de protocole de mobilité IP, ce mode de décision est particulièrement adéquat. En effet, le protocole de mobilité se charge du reroutage dynamique de paquets.
Comme mentionné dans le chapitre 802.11, l'entité de décision du handover dans les réseaux 802.11 est située dans la STA.

Dans le mode contrôlé réseau, le réseau mesure la puissance du signal du terminal et ordonne au terminal de se connecter à une cellule particulière si nécessaire. Ce mode conduit à une charge de signalisation importante sur le réseau vu que seul le réseau effectue des mesures.
De plus, dans le cadre du handover inter-système, le terminal est le seul élément conscient de la présence de plusieurs réseaux. Le mode contrôlé réseau est essentiellement employé dans les réseaux à commutation de circuits. L'avantage de ce mode est que le réseau surveille sa charge et peut éviter les saturations de ressources, ce que ne permet pas le mode contrôlé terminal. [30][32][33]

Enfin, il existe un mode de contrôle inspiré des deux modes précédents : le mode contrôlé réseau et assisté terminal. Dans ce mode, le réseau effectue les mesures de puissance de la même façon qu'en mode contrôlé réseau. Cependant, les mesures réseau sont accompagnées des mesures renvoyées par le terminal. Dès lors, le réseau contrôle la décision du handover en tenant compte des mesures du terminal. Ce mode de contrôle est employé par le réseau UMTS. Le RNC décide du handover en fonction de mesures effectuées sur le terminal et de mesures réceptionnées du terminal [34].

Après avoir récupéré les paramètres de mesure et leur changement au cours du temps, l'entité de décision décide du handover.

Les paramètres déclencheur du handover sont la puissance du signal, la mobilité plus ou moins élevée du terminal, la charge de la cellule et l'application utilisée par le terminal.
Si la puissance du signal de la cellule courante tombe sous une valeur seuil et que la puissance du signal d'une cellule voisine est supérieure, un handover peut être déclenché.
Si le client se déplace rapidement ou lentement, un handover peut également être déclenché.
Dans le cas d'un terminal communiquant avec le réseau 802.11 et se déplaçant rapidement, un handover vers le réseau UMTS (si le réseau est disponible) est déclenché vu la portée des APs 802.11.
Ensuite, la charge du réseau ou de la cellule courante peut déclencher le handover si une cellule voisine est moins chargée que la cellule courante. De cette façon, la charge du réseau est répartie sur les cellules.
Enfin, si la cellule courante ne supporte pas une certaine application en terme de bande passante requise ou de QoS, le réseau peut déclencher un handover vers un autre réseau assurant les conditions requises par l'application.[34]

Exécution du handover

Une fois la décision du handover prise, l'entité d'exécution, terminal ou réseau, est informée du handover à accomplir. En UMTS, l'exécution du handover est effectuée par le réseau, à savoir le DRNC alors qu'en 802.11, le handover est effectué par le terminal. [35]

Modélisation réseau

Après avoir introduit les architectures UMTS et 802.11, le fonctionnement des réseaux respectifs, les protocoles de mobilité assurant le support du handover inter-système et les exigences liées au handover entre UMTS et 802.11, on envisage la modélisation réseau du handover inter-système entre UMTS et 802.11.
La modélisation réseau est la définition d'une architecture rendant possible la simulation et l'analyse de performances des éléments constitutifs du réseau.
La modélisation réseau implique l'utilisation d'un simulateur logiciel capable de définir un réseau en terme de nœuds, liens et technologies.
Dans cette partie sont abordés le simulateur logiciel NS-2, son support technologique (UMTS, 802.11) et ses extensions. Les conflits entre le simulateur et les extensions ainsi que les conflits entre les extensions sont détaillés. Enfin, une solution de modélisation est proposée.

Pour illustrer l'objectif de la modélisation, il est utile de le présenter à l'aide d'un scénario de simulation. Le scénario envisagé est un utilisateur équipé d'un terminal mobile bi-mode UMTS/802.11 connecté à la passerelle entre les deux réseaux d'accès. L'utilisateur est connecté initialement au réseau 802.11 ou au réseau UMTS. Durant la communication, l'utilisateur se déplace du réseau UMTS (802.11) vers un réseau 802.11 (UMTS). Les réseaux UMTS et 802.11 sont interconnectés en configuration tight coupling. Quand l'utilisateur change de localisation, l'entité de handover exécute un handover inter-système entre UMTS et 802.11. Pour assurer la continuité de la communication suite au handover, le protocole de mobilité Mobile IPv6 est mis en œuvre.

NS-2

Pour simuler ce scénario, l'outil NS-2 (Network Simulator v2) de simulation de réseau a été employé. NS-2 est bâti selon les idées de la conception par objets, de la réutilisation du code et de la modularité. Il est aujourd'hui un standard de référence dans le domaine de la simulation logicielle. Ce logiciel est dans le domaine public, son utilisation est gratuite.
Le projet VINT développant NS-2 est une collaboration entre USC/ISI, Xerox parc, LBNL et UCB. Il a pour objectif la construction d'un simulateur multi-protocoles pour permettre l'étude d'interaction entre les protocoles et le comportement d'un réseau. Le projet contient des bibliothèques pour la génération de topologies réseaux, des trafics ainsi que des outils de visualisation tel que l'animateur réseau NAM (Network ANimator). [37][38]

Introduction

Le simulateur NS-2 est particulièrement adapté à l'étude de réseau à commutation de paquets et à la réalisation de simulations. Il contient les fonctionnalités nécessaires à l'étude des algorithmes de routage unicast ou multicast, des protocoles de transport, de session, de réservation, des services intégrés, des protocoles d'application comme http. De plus, le simulateur possède une palette de systèmes de transmission, d'ordonnanceurs et de politiques de gestion de file d'attente pour effectuer des études de contrôle de congestion. Les principaux composants de NS-2 sont :

Application	Web, ftp, telnet, générateur de trafic (CBR,…)
Transport	TCP,UDP,RTP,SRM
Routage	Statique, dynamique, multicast
Gestion de file d'attente	RED, DropTail, Token bucket
Discipline de service	CBQ, SFQ, DRR, Fair queueing
Système de transmission	CSMA/CD, CSMA/CA, point à point

Utilisation du simulateur

Du point de vue utilisateur, la mise en œuvre de NS-2 se fait via une étape de programmation en langage Tcl qui décrit la topologie du réseau et le comportement de ses composants, vient ensuite l'étape de simulation et enfin l'interprétation des résultats. Ces différentes étapes sont détaillées.
Avant d'éclaircir le scénario général, on s'intéresse à un scénario simple n'impliquant qu'une topologie simple de quatre machines fixes communiquant l'une avec l'autre via trois liaisons.
Dans la terminologie NS, une machine s'appelle un nœud. Un nœud peut contenir des agents (TCP, UDP,…), ces agents pouvant supporter un type d'application (FTP, CBR, …). [37]
Cette simulation s'effectue via le script Tcl suivant : Exemple.tcl [37][38]

```
#création d'un simulateur
set ns [new Simulator]

#création du fichier de tracage des paquets
set trace [open out.tr w]
$ns trace-all $trace

#création du fichier de tracage des paquets pour le visualisateur NAM
set namf [open out.nam w]
$ns namtrace-all $namf

#quand la simulation est terminée, la procédure finish est appelée,
l'exécution de nam permet la #visualisation de la topologie et des paquets
transitant
proc finish {} {
      global ns trace
      $ns flush-trace
      close $trace
      close $namf
      exec nam out.nam &
      exit 0
}

#création de 4 nœuds
set n0 [$ns node]
set n1 [$ns node]
set n2 [$ns node]
set n3 [$ns node]

#création de lignes de communication full duplex entre nœuds
$ns duplex-link $n0 $n2 1Mb 10ms DropTail
$ns duplex-link $n1 $n2 1Mb 10ms DropTail
$ns duplex-link $n3 $n2 1Mb 10ms DropTail
```

```
#création d'agents UDP, les données dans NS sont transmises entre agents
set udp0 [new Agent/UDP]
$ns attach-agent $n0 $udp0
set udp1 [new Agent/UDP]
$ns attach-agent $n1 $udp1

#création d'application génératrice de paquets à vitesse constante
#paquets de 500 bytes générés toutes les 5ms
#l'agent cbr0 est implanté sur le nœud n0 et cbr1 sur le nœud n1
set cbr0 [new Application/Traffic/CBR]
$cbr0 attach-agent $udp0
$cbr0 set packetSize_ 500
$cbr0 set interval_ 0.005
set cbr1 [new Application/Traffic/CBR]
$cbr1 attach-agent $udp1
$cbr1 set packetSize_ 500
$cbr1 set interval_ 0.005

#création d'un agent vide, destiné à recevoir les paquets, implanté dans n1
set null0 [new Agent/Null]
$ns attach-agent $n3 $null0

#routage des trafics
$ns connect $cbr0 $null0
$ns connect $cbr1 $null0

#début et fin de génération de paquets par cbr0 et cbr1
$ns at 0.5 « $cbr0 start »
$ns at 1.0 « $cbr1 start »
$ns at 4.0 « $cbr1 stop »
$ns at 4.5 « $cbr0 stop »

#simulation durant 5 secondes avec appel de procédure finish
$ns at 5.0 « finish »

#début de la simulation
$ns run
```

L'exécution de NAM dans la procédure « finish » permet la visualisation dynamique des paquets circulant entre les nœuds :

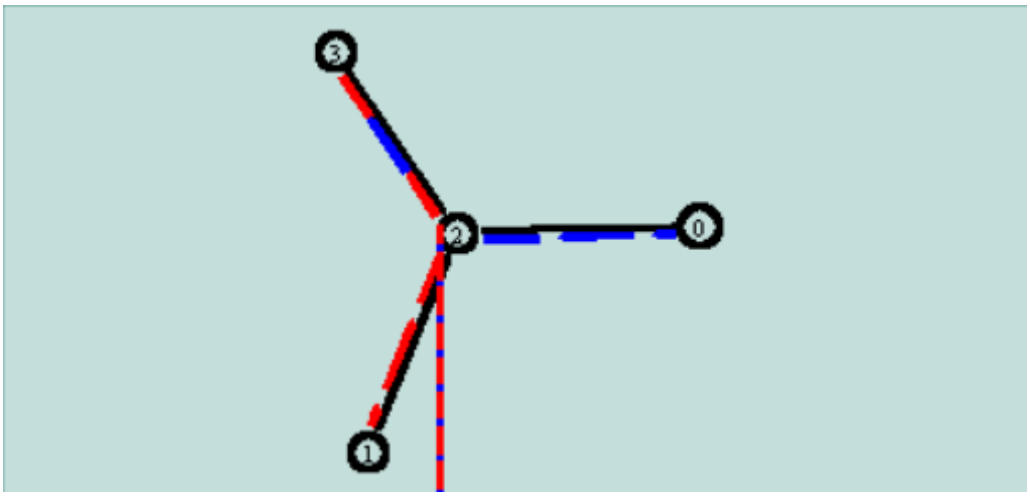

Figure 14 – Visualisation Nam de exemple.tcl [38]

Logiquement, on observe une perte de paquets dans le nœud n2 due à la liaison n2 – n3 de 1Mb/s alors que n0 et n1 envoient leurs paquets à un débit de 800kb/s.

Le fichier « out.tr » contient quant à lui des informations brutes sur les paquets :

Evénement	Temps	Depuis le nœud	Vers le nœud	Type de paquet	Taille Pkt	Flags	ID flux	Adress source	Adress dest	Num Seq	ID pkt

```
+ 1.764 2 3 cbr 500 ------- 0 1.0 3.0 150 400
r 1.764 0 2 cbr 500 ------- 0 0.0 3.0 250 401
+ 1.764 2 3 cbr 500 ------- 0 0.0 3.0 250 401
d 1.764 2 3 cbr 500 ------- 0 0.0 3.0 250 401
r 1.764 2 3 cbr 500 ------- 0 1.0 3.0 109 318
+ 1.765 1 2 cbr 500 ------- 0 1.0 3.0 153 406
- 1.765 1 2 cbr 500 ------- 0 1.0 3.0 153 406
+ 1.765 0 2 cbr 500 ------- 0 0.0 3.0 253 407
- 1.765 0 2 cbr 500 ------- 0 0.0 3.0 253 407
- 1.766 2 3 cbr 500 ------- 0 1.0 3.0 112 324
r 1.768 2 3 cbr 500 ------- 0 0.0 3.0 209 319
r 1.769 1 2 cbr 500 ------- 0 1.0 3.0 151 402
+ 1.769 2 3 cbr 500 ------- 0 1.0 3.0 151 402
r 1.769 0 2 cbr 500 ------- 0 0.0 3.0 251 403
+ 1.769 2 3 cbr 500 ------- 0 0.0 3.0 251 403
d 1.769 2 3 cbr 500 ------- 0 0.0 3.0 251 403
```

```
r : réception du paquet au nœud « vers le nœud »
+ : mise en file d'attente du paquet
- : sortie de file d'attente du paquet
d : suppression du paquet dans la file

Adress source : n° nœud.port
Adress dest : n° nœud.port
```

Développement de nouveaux composants

NS est en réalité un programme relativement complexe écrit en C++ et interfacé via Tcl. Pour modifier le comportement d'objets existants ou en concevoir de nouveaux, il est donc nécessaire de passer par une étape d'implémentation en C++.
Cette étape de développement pour la simulation du scénario de handover est détaillée au chapitre suivant. [38][39][40]

802.11 dans NS-2

Tout comme la classe Node employée sur le scénario simple ci-dessus, existe la classe MobileNode. Un MobileNode est un Node auquel sont ajoutées des fonctionnalités sans fil et mobile telle que la possibilité de se déplacer dans une topologie et de transmettre et recevoir sur un canal, les noeuds filaires (Node) transmettant sur une liaison câblée (Link). [3][5]
La définition des nœuds étant à la base de tout scénario, l'architecture des nœuds mobiles sur NS est explicitée.

MobileNode 802.11

Un nœud mobile 802.11 (STA ou AP) consiste dans NS-2 en un objet MobileNode configuré par un protocole de routage, une pile réseau, un canal, une topographie, un modèle de propagation. Le nœud mobile caractérisant un AP se distingue de la STA par l'activation du routage filaire. L'AP peut alors être connecté à des nœuds fixes d'une part et à des STAs d'autre part. [41][42]

La pile réseau du nœud mobile consiste en une couche liaison (LL), une file d'attente d'interface basée sur la priorité (IFq), une couche mac (MAC) et une interface réseau (netIF) caractérisée par une antenne, le tout connecté à un canal. L'interface réseau emploie un modèle de propagation. [41]

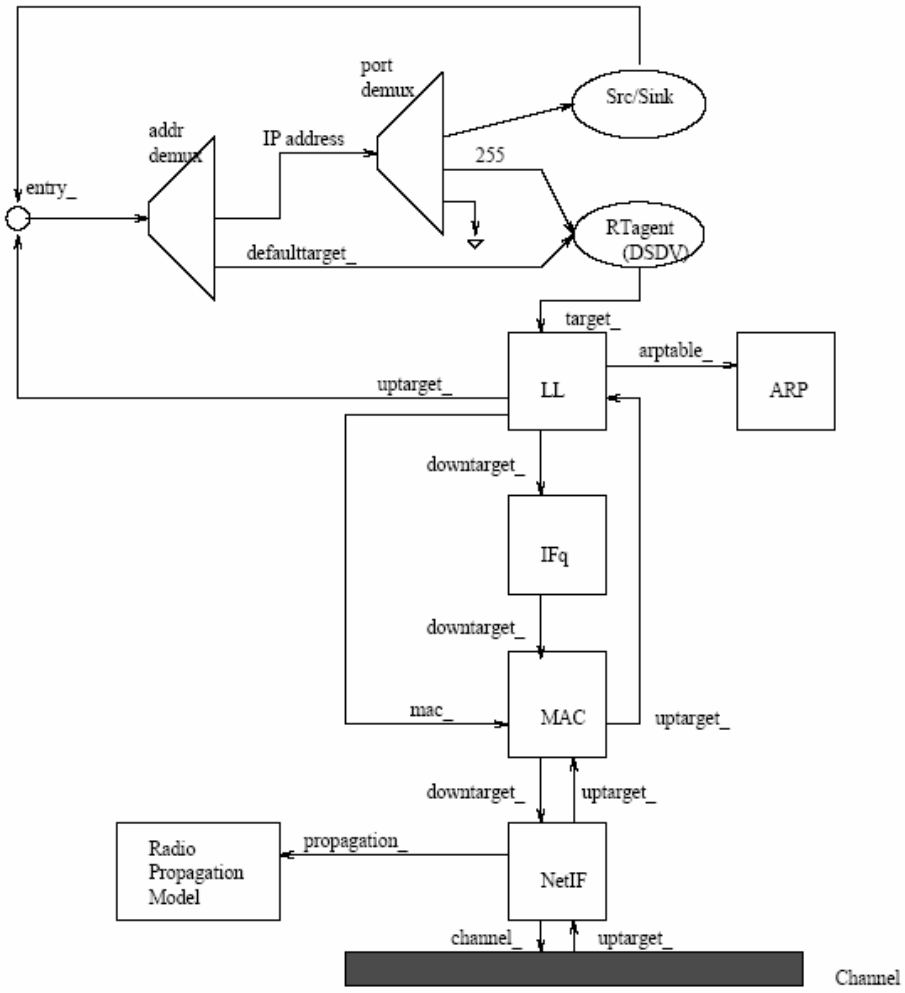

Figure 15 – Schéma d'un nœud mobile dans NS-2 de [41]

L'émission et la réception finale des paquets sur une STA ou un AP s'effectue entre agents (Src/Sink). Ces agents sont les constructeurs et consommateurs de paquets IP. Lors de l'émission et la réception d'un paquet par un agent, le paquet est transféré au point d'entrée (entry_) du nœud mobile. Le classificateur addr demux examine les champs du paquet (habituellement sa destination). Ce classificateur agit comme la table de routage du nœud. Cette table est remplie par l'agent de routage (RTagent).

Si le nœud correspond à l'adresse de destination du paquet, le paquet est transmis au port dmux. Le port dmux se charge de délivrer le paquet reçu à l'agent associé au port. Si le nœud n'est pas destinataire du paquet, addr dmux vérifie que l'adresse de destination est contenue dans sa table d'adresses. Si c'est le cas, le paquet est transféré à l'agent de routage qui assigne le prochain hop pour le paquet et transfère ensuite le paquet à la couche liaison.

Si ce n'est pas le cas, le paquet est toujours transféré à l'agent de routage mais via la variable defaulttarget_ indiquant à l'agent de routage qu'il doit envoyer des requêtes de routage pour obtenir la route jusqu'à la destination. Le paquet est mis en tampon durant le processus de recherche de la route. Si le nombre de paquet en attente dépasse la taille du tampon, les paquets sont perdus. Après avoir récupéré la route, l'agent de routage transfère le paquet à la couche liaison.

La couche liaison (LL) est connectée à un module ARP (Address Resolution Protocol) qui traduit les adresses IP en adresses matérielles MAC. En principe pour tous les paquets sortants (vers le canal), les paquets sont transmis à LL par l'agent de routage. La LL transmet à son tour les paquets vers la file d'attente d'interface. Pour tous les paquets entrants (depuis le canal), la couche MAC passe les paquets à la LL qui les remet ensuite au point d'entrée entry_.

Le module ARP reçoit des requêtes de la couche liaison. Si ARP dispose de l'adresse matérielle de destination, il l'écrit dans l'en-tête MAC du paquet. Sinon il diffuse une requête ARP, et met le paquet temporairement en attente. Pour chaque adresse matérielle de destination inconnue, il existe un buffer pour un seul paquet. Si plusieurs paquets vers la même destination sont envoyés à ARP, le paquet précédemment mis en tampon est supprimé. Une fois que l'adresse matérielle de destination est connue, le paquet est inséré dans la file d'attente de l'interface.

La file d'attente d'interface et en particulier la file « PriQueue » est une file d'attente basée sur la priorité. Elle donne priorité aux paquets de protocole de routage, en les insérant à la tête de la file.

La couche MAC 802.11 est implémente la fonction DCF (Distributed Coordination Function). Elle utilise un modèle RTS/CTS/DATA/ACK pour tous les paquets unicast et envoie simplement DATA pour les paquets de diffusion. Les mécanismes RTS/CTS et d'acquittement positif sont donc supportés. Les fonctions de détection de porteuses physique et virtuelle sont également implémentées. Les transitions BSS ne sont en revanche pas supportées.

L'interface réseau (netIF) consiste en l'interface matérielle (la couche physique) permettant au nœud mobile d'accéder au canal. L'interface sans fil partagée est implémentée par la couche physique « WirelessPhy ». Cette interface est sujette aux collisions et le modèle de propagation radio reçoit les paquets transmis par d'autres interfaces sur le canal. L'interface marque chaque paquet transmis avec des méta-données concernant l'interface de transmission telles que la puissance de transmission, la longueur d'onde, etc. Ces méta-données sont utilisées par le modèle de propagation de l'interface réseau réceptrice pour déterminer si le paquet a la puissance minimale pour être reçu et/ou détecté par le nœud récepteur.
La couche physique décrite par WirelessPhy n'inclut pas les mécanismes et principes des couches physiques 802.11. Les principes FHSS, DS, HR-DS et OFDM ne sont pas implémentés.

Enfin, le modèle de propagation radio permet pour modéliser le type d'environnement géographique. L'antenne est par défaut une antenne omni-directionnelle de gain unitaire.

NS-2 permet de surveiller et tracer les paquets réceptionnés, supprimés et envoyés par des agents, couche MAC ou file d'attente d'interface pour des analyses [37].

UMTS dans NS-2

Les extensions EURANE (Enhanced UMTS Radio Access Network) développée dans le cadre du projet SEACORN pour Ericsson Telecommunicatie B.V. et l'extension « Strathclyde University » UMTS développée par Pablo Martin et Paula Ballester, améliorée par Olivier Hynderick et Sven Raes des FUNDP ont été considérées pour ce projet.

EURANE ajoute 3 nœuds à savoir : le RNC, la BS et l'UE UMTS autorisant le support des canaux FACH, RACH, DCH et HS-DSCH. Les canaux FACH et RACH sont des canaux communs aux UEs d'une même cellule alors que le canal DCH est un canal dédié pour un UE. Le canal HS-DSCH permet le transport de données utilisateur à haut débit grâce à la technique HSDPA.
Ces nœuds implémentés sont caractérisés principalement par les services de segmentation et retransmission de données. Le handover entre canaux RACH/FACH et DCH est supporté.
Les nœuds ajoutés par EURANE sont des nœuds non mobiles (Node), ne supportant donc pas de se déplacer sur une topologie. Cependant, la modélisation de l'interface radio est effectuée par un canal, identique au canal utilisé dans les communications 802.11 de NS, associé à un modèle d'erreur. [43]

Interconnexion 802.11/UMTS dans NS-2

Le scénario d'interconnexion consiste à disposer d'une STA communiquant avec un nœud fixe d'interconnexion par l'intermédiaire d'un AP, le nœud d'interconnexion étant connecté à l'ensemble du réseau UMTS modélisé par EURANE. L'UE doit également pouvoir communiquer avec le nœud d'interconnexion.

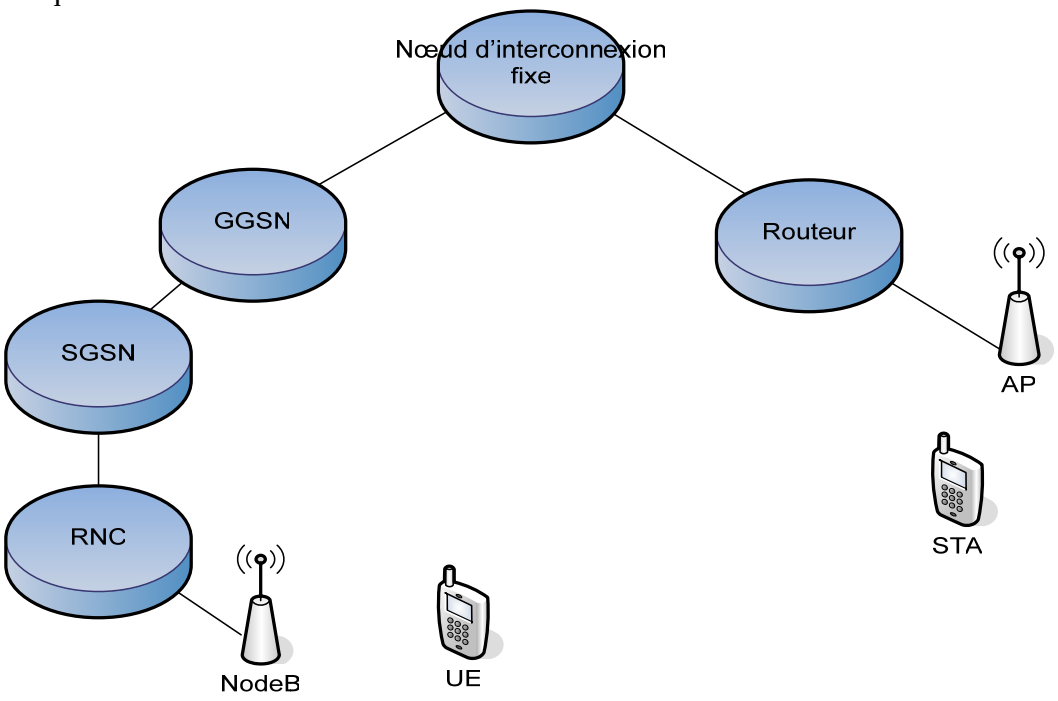

Figure 16 – Scénario d'interconnexion

Lors de la simulation de ce scénario, apparaissent des anomalies de routage au départ de la STA. En simulation d'interconnexion de réseau UMTS avec un réseau fixe, les paquets émis de l'UE et reçus par l'UE vers/depuis le nœud d'interconnexion suivent la route GGSN-SGSN-RNC-NodeB-UE.
Les nœuds fixes extérieurs à l'UTRAN et au CN peuvent communiquer avec l'UE. L'UE pour pouvoir communiquer avec des nœuds fixes extérieurs nécessite l'implantation d'une

passerelle, réalisée dans le SGSN par la fonction add-gateway. La fonction add-gateway n'est pas appliquée au GGSN car le SGSN est le premier nœud fixe non défini par EURANE. Cette fonction consiste à transmettre les paquets provenant du RNC à l'agent de routage via defaulttarget. La simulation de ce scénario a mis en évidence la nécessité de définir un routage statique entre les nœuds excepté l'UE, le NodeB et le RNC. En effet, en utilisant un routage AODV, DSDV ou DSR, la STA envoie ses paquets vers le nœud d'interconnexion via l'UE.

Mobile IPv6 dans NS-2

Pour que la communication soit maintenue entre le terminal mobile et la machine fixe, on a recours au protocole Mobile IPv6.
NS-2 inclut une implémentation du protocole Mobile IPv4. L'extension MobiWan développée par MOTOROLA Labs Paris en collaboration avec l'équipe INRIA PLANETE est basée sur le protocole Mobile IPv4 inclus dans NS-2. Cette extension a été adaptée par WMC (Twente Institute for Wireless and Mobile Communications BV).

MobiWan

MobiWan permet le support pour des nœuds mobiles et fixes prédéfinis – routeur transit, routeur de frontière, routeur de site, station de base et terminaux mobiles – du protocole Mobile IPv6. Ces nœuds prédéfinis ne contiennent pas les nœuds définis par EURANE.

Le protocole de mobilité IPv6 implémenté dans MobiWan inclut la modification de l'en-tête IPv4 originale, l'ajout d'annonce de routeur et de sollicitations entre les stations de base et les terminaux mobiles. Les mécanismes du protocole Mobile IPv6 sont implémentés dans MobiWan excepté la suppression de la fonction de tunnelling et le support d'IPSec. MobiWan implémente un protocole de routage (Network) dynamique hiérarchique.[44]
On observe ici une limitation fondamentale. Mobile IPv6 est un protocole de routage dynamique alors qu'EURANE nécessite l'implantation d'un protocole de routage statique pour que la STA communique avec le nœud d'interconnexion.

MobileNode supportant Mobile IPv6

De façon analogue au MobileNode 802.11, on définit un MobileNode supportant Mobile IPv6. Ce nœud mobile diffère sensiblement du MobileNode défini dans 802.11 dans NS-2 par l'ajout de classificateurs et d'agents :

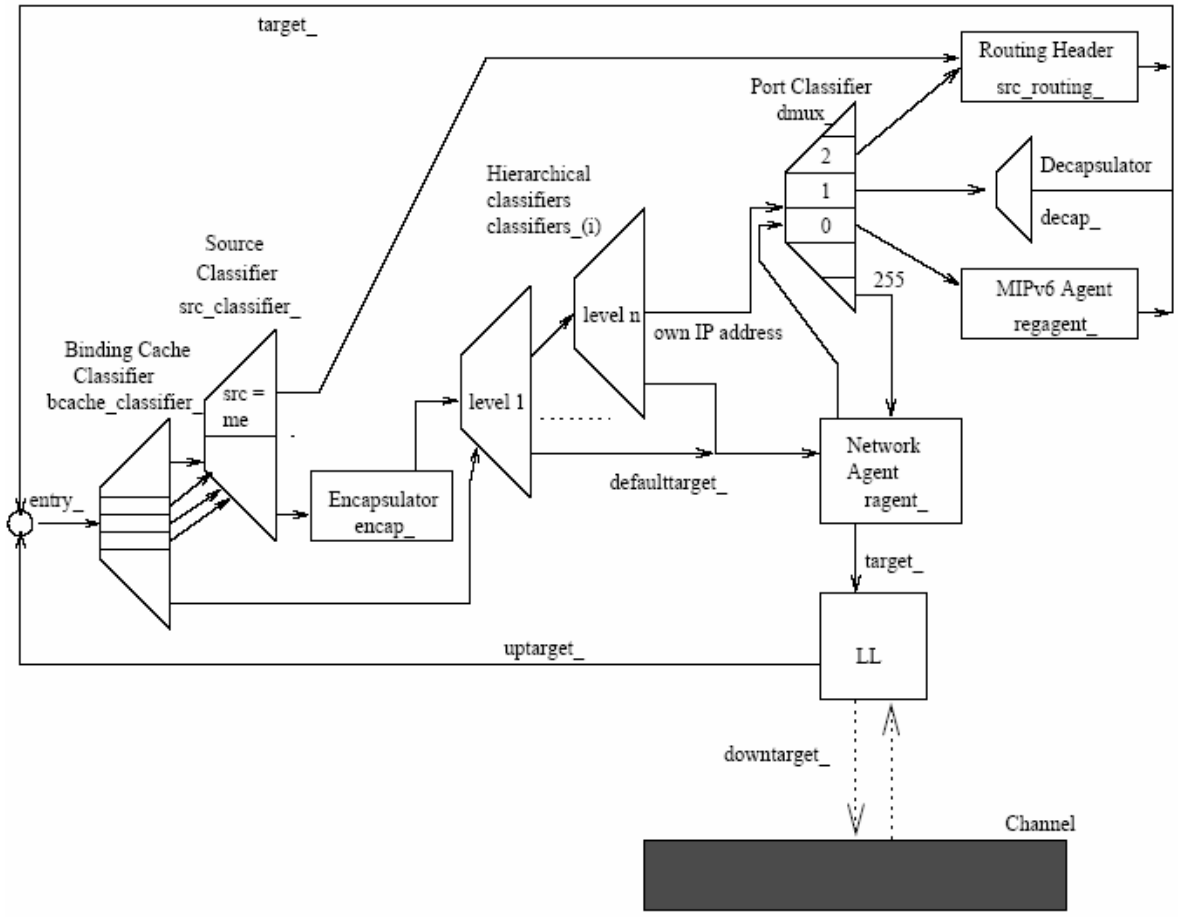

Figure 17 – Schéma d'un MobileNode supportant Mobile IPv6 [44]

Les mécanismes d'envoi et de réception de paquets restent semblables aux mécanismes mentionnés pour le MobileNode 802.11. Le classificateur Binding Cache est employé lorsqu'un nœud reçoit un message « Binding Update ». Une entrée est alors ajoutée dans son Binding Cache. Consécutivement, la table de routage du nœud doit être mise à jour pour rediriger les paquets vers la care-of adresse spécifiée.

Les classificateurs Hierarchical agissent en tant que table de routage. La différence avec le classificateur addr dmux est que l'adresse des nœuds est exprimée sur un certains nombres de niveaux hiérarchiques. Concrètement, le réseau est structuré sous forme d'arbre à plusieurs niveaux dans lequel chaque nœud est une branche ou feuille. Lors de l'utilisation d'adressage hiérarchique, l'agent de routage doit également être hiérarchique. En fonction de l'adresse de destination, le paquet est envoyé vers un nœud ascendant ou descendant dans l'arbre. Un exemple d'adresse hiérarchique à 3 niveaux est 1.0.3, l'espace d'adressage est divisé en 3 niveaux : domaine, cluster d'un domaine et nœud dans chaque cluster. [44]

Le problème majeur d'implémentation du scénario de handover avec l'extension EURANE apparaît lorsque l'adressage des nœuds doit être hiérarchique. En effet, EURANE emploie l'adressage plat. L'adressage plat consiste à assigner à chaque nœud un nombre. Chaque nœud doit connaître la route vers les autres nœuds avec ce type d'adressage.

A l'opposé, en employant un adressage hiérarchique, un nœud ne connaît la route que vers les nœuds appartenant à son cluster.

Une topologie de nœuds UMTS d'adresse plate, excepté le SGSN et le GGSN d'adresse hiérarchique a été envisagée, l'AP, la STA et le routeur étant caractérisés par une adresse hiérarchique. Elle a montré que les adressages hiérarchiques et plats ne peuvent coexister dans une même simulation.

L’agent Encapsulator et le classificateur Decapsulator assurent les fonctions respectives d’encapsulation et de désencapsulation mises en œuvre lors du tunnelling de paquet.

Enfin, l’agent MIPv6 est sous-classé en BSAgent, CNAgent et MNAgent. Le BSAgent définit les fonctions de home agent avec notamment la fonction d’envoi d’annonce de routeur.
Le CNAgent est l’agent Mobile IP attaché au nœud correspondant avec le terminal mobile.
Le MNAgent contient les fonctions de réception de care-of adresse, de réception des annonces routeur, d’envoie de sollicitation. Etant donné que les nœuds UMTS d’EURANE ne supportent pas l’adressage hiérarchique, ces nœuds n’acceptent pas d’agent MIPv6 dont les fonctions requièrent également une adresse hiérarchique.

Solution proposée

Pour modéliser le scénario de handover entre les réseaux UMTS et 802.11, au vu des incompatibilités d’adressage, de routage hiérarchique et de définition de nœuds entre MobiWan et EURANE, l’architecture proposée simule le fonctionnement du réseau UMTS par un réseau 802.11. Cette solution, explicitée par la suite, assure le support de l’adressage et du routage hiérarchiques sur tous les nœuds (l’extension « Strathclyde University » UMTS étant également fondée sur un adressage hiérarchique).

Architecture et implémentation

La solution d'architecture proposée repose sur l'AP et la STA 802.11 supportant Mobile IPv6. Le scénario de simulation envisagé est détaillé dans la section « Entité de handover ».

Comme mentionné au chapitre 7, le terminal doit supporter à la fois les couches de protocole UMTS et 802.11. Le terminal mobile est simulé par un nœud mobile (MobileNode). Ce nœud mobile est caractérisé lors de sa création par une pile de protocoles de liaison (LL), MAC 802.11 et physique (WirelessPhy), voir MobileNode 802.11 – chapitre 8.

Implémentation 1 : Ajout d'interface

La première étape d'implémentation consiste à ajouter une deuxième pile de protocoles au terminal mobile afin qu'il soit capable de communiquer avec l'AP simulant le NodeB et avec l'AP 802.11.
La fonction d'ajout d'interface « add-interface » de NS a été réécrite pour ce faire.

La pile de protocoles, simulant les protocoles UMTS, à ajouter par la fonction « add-interface » est limitée à la modélisation de la propagation radio UMTS. Lors de la tentative d'intégration des couches Mac/Umts ou Mac/Hsdpa ainsi que des protocoles UMTS/RLC/AM ou UMTS/RLC/AMHS d'EURANE [7] au lieu des couches MAC 802.11 et LL respectivement, un timer de NS ne peut être programmé. Sans autre indication de sortie, l'intégration de la couche MAC et du protocole RLC est hasardeuse dans les MobileNodes.

Paramétrisation « UMTS »

La couche physique WirelessPhy étant paramétrée pour un usage réseau 802.11, elle a été paramétrée pour le NodeB et l'interface UMTS du mobile par des valeurs issues de [[3] – Chapitre 8, Table 8.2 et Table 8.5].
Pt_(UE) = 0.25 W //puissance du signal transmis
Pt_(NodeB) = 3 W
Pt_consume(UE) = 0.125 W //consommation de puissance (moyenne) pour la transmission
Pt_consume(NodeB) = 1 W
P_idle(UE) = 0.005 W
P_idle(NodeB) = 0.5 W
Bandwidth_ = 384 kbps
Freq_ = 2 GHz
RXThresh_ = 1e-16 W //seuil de réception de puissance
CSThresh_ = 4.3e-18 W //seuil de détection de porteuse

L'antenne de l'interface UMTS est également caractérisée par rapport à l'antenne 802.11 :
Gt_(UE) = 2 dB //gain en transmission de l'antenne
Gt_(NodeB) = 18 dB
Gr_(UE) = 2 dB //gain en réception de l'antenne
Gt_(NodeB) = 18 dB

Enfin, la couche MAC est limitée à un dataRate_ = 384 kbps

Paramétrisation « 802.11 » de la couche physique WirelessPhy

Les paramètres de puissance de transmission et de réception en 802.11 sont également inadéquats dans le code de NS-2. Ces valeurs font référence à une interface radio Lucent WaveLan 914MHz. Le standard 802.11 spécifiant des fréquences comprises dans la bande 2.412-2.472GHz, ces valeurs ont été modifiées.
La référence pour la mise à jour de ces valeurs est l'interface Cisco Aironet 350. Dans 802.11, l'AP ne se distingue pas de la STA par ces valeurs.
Pt_ = 100 mW
Pt_consume = 100 mW
P_idle = 0.004 W
CSThresh_ = 1.559e-11 W
RXThresh_ = 3.652e-10 W
Freq_ = 2.412 GHz

La couche MAC accepte un débit de 54Mbps

Canal UMTS

Le canal radio employé pour les communications UMTS est simulé au moyen d'un canal partagé en accès DCF.

Implémentation 2 : Entité de handover

L'entité de handover a été décrite au chapitre 7.
Il convient d'abord de la déterminer où est située l'entité de handover. L'entité de handover se situe dans le RNC pour le réseau UMTS et dans la STA pour le réseau 802.11.
La collecte de mesures s'effectuant par le réseau en UMTS et par la STA en 802.11.
La solution adoptée est de placer l'entité de handover dans le terminal mobile. Le contrôle de handover est décentralisé.

La mesure de décision du handover est la puissance de réception du signal transmis par les NodeBs et APs à portée du terminal mobile. Ce paramètres tient compte de la charge d'utilisateurs communiquant avec le NodeB et l'AP ainsi que du type d'environnement géographique (indoor/outdoor, si outdoor, zone urbaine, suburbain ou rurale). Ce paramètre prend en compte la présence d'autres utilisateurs communiquant avec la station de base (NodeB ou AP) du terminal mobile car l'émission de chaque paquet nécessite une puissance de transmission de la station de base. L'influence du paramètre sur le type d'environnement se rapporte au modèle de propagation utilisé.

Consécutivement aux mesures effectuées par le terminal, l'entité de handover transfère les paquets soit vers l'interface UMTS soit vers l'interface 802.11.
Il convient alors de déterminer où se place l'entité de handover dans les couches de protocoles du terminal mobile.
L'interface physique est capable de mesurer la puissance de réception du signal.
Dans 802.11, cette fonction est réalisée par la détection physique de porteuse.
Dans UMTS, la fonction de contrôle de puissance est effectuée par les canaux physiques DPCCH et CPICH. [3]

Les fonctions de décision et d'exécution de l'entité de handover ne sont cependant par présentes dans la couche physique. En effet, suite à la prise de décision du handover, le terminal mobile doit informer le NodeB et l'AP de la nouvelle route que les paquets doivent suivre pour l'atteindre. L'entité de décision et d'exécution du handover doit donc se trouver au-dessus du protocole de mobilité afin qu'il se charge d'établir une route entre le réseau et le terminal mobile.

Implémentation de l'entité de handover

D'après les choix explicités ci-dessus, l'entité de mesure, décision et exécution est placée dans le terminal mobile bi-mode. L'implémentation « bi-mode » est détaillée par après.
Pour illustrer le fonctionnement de l'entité, on considère le scénario suivant :
Soit la topologie réseau :

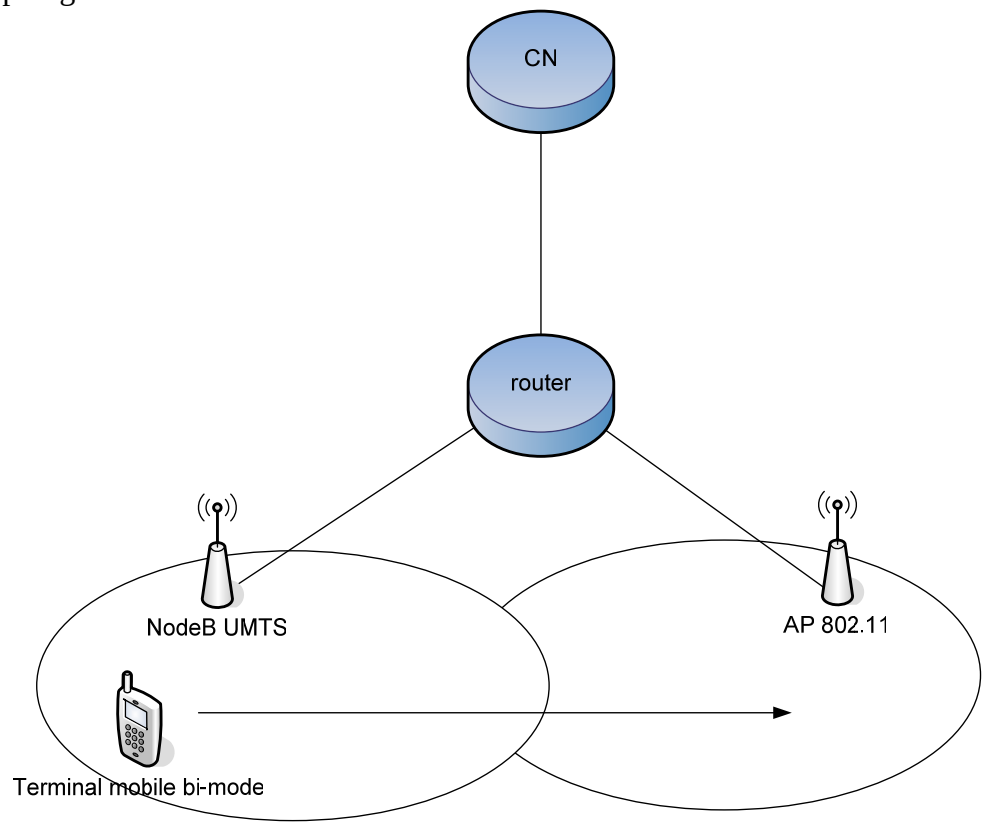

Figure 18 – Scénario de handover vertical

Le terminal mobile établit une connexion TCP avec le CN via le NodeB, le home agent du terminal, et débute un téléchargement FTP depuis le CN. Le terminal mobile se déplace dans la topologie géographique. Lorsqu'il arrive à portée de l'AP, l'AP annonce sa présence au terminal.
Le terminal enregistre alors sa nouvelle care-of adresse. Après réception de sa nouvelle care-of adresse, le terminal envoie une mise à jour de binding au CN et au home agent. Le CN apprend qu'il doit mettre à jour sa table de routage pour transmettre des paquets au terminal via sa nouvelle adresse. Le home agent, le NodeB, apprend qu'il ne doit plus transmettre de paquet vers l'ancienne adresse du terminal et que les paquets lui parvenant, à destination du terminal, doivent être routés par l'AP. Une fois la procédure de mise à jour de binding effectuée, le terminal poursuit la communication avec le CN.
L'entité de décision de handover surveille la puissance de réception au terminal mobile des paquets émis par le NodeB et l'AP. Périodiquement, le NodeB et l'AP diffusent leur annonce de routeur. En fonction de la puissance mesurée à la réception des annonces et des paquets de données, l'entité de handover prend sa décision. Si la puissance de réception du NodeB est

supérieure à la puissance de réception de l'AP et que le terminal mobile communique avec le CN via l'AP, l'entité de handover envoie une sollicitation au NodeB, ce qui marque l'exécution du handover.
Consécutivement, le terminal reçoit l'annonce de routeur depuis le NodeB puis transmet une mise à jour de binding au CN et au home agent.
Si la puissance de réception du NodeB est supérieure à la puissance de réception de l'AP et que le terminal mobile communique avec le CN via le NodeB, le handover n'est pas déclenché.

Implémentation 3 : Modèle de propagation

La puissance de réception mesurée par l'entité de handover est obtenue en fournissant à un modèle de propagation les caractéristiques de la couche physique (voir chapitre 8 – MobileNode 802.11).

Le modèle de propagation utilisé pour la simulation de communication entre le terminal mobile et l'AP est le modèle Shadowing. Ce modèle est basé sur l'équation de Friis corrigée par une fonction d'affaiblissement (path loss) dépendant d'un indice.

Le modèle de propagation utilisé pour la simulation de communication entre le terminal mobile et le NodeB est le modèle d'Okumura-Hata.
Le modèle d'Okumura-Hata est un modèle de propagation empirique. Il se base sur l'utilisation d'importantes quantités de mesures pour déterminer des propriétés statistiques générales.

Le modèle d'Okumura-Hata est un modèle classique pour la planification de réseaux cellulaires dans les zones urbaines. Il s'agit d'une approche empirique, qui décompose l'affaiblissement du signal radio en différentes fonctions. Ce modèle a été développé spécialement pour les applications cellulaires. Ce modèle a servi notamment au déploiement de réseau GSM à Tokyo et au déploiement de réseau UMTS à Espoo, Finlande. [3]
Ce modèle s'appuie sur des mesures effectuées à Tokyo, zone urbaine présentant peu de relief en utilisant une antenne émettrice à 200m de hauteur et un récepteur mobile à 3m.
Il est valable pour des fréquences de 100MHz à 3GHz et pour une distance émetteur-récepteur de 1km à 100km. [45]
Les équations caractérisant ce modèle sont :
Mesure de l'affaiblissement :

$$L_{50} = L_F + A_{MU}(f,d) - G(h_{te}) - G(h_{re}) - G_{AREA}$$

$$G(h_{te}) = 20\log_{10}\left(\frac{h_{te}}{200}\right) pour 10m < h_{re} < 1000m$$

$$G(h_{te}) = 10\log_{10}\left(\frac{h_{te}}{3}\right) pour h_{re} \leq 3m$$

$$G(h_{te}) = 20\log_{10}\left(\frac{h_{te}}{3}\right) pour 3m < h_{re} < 10m$$

Où L_{50} est la valeur médiane (valeur du 50ième percentile) de l'affaiblissement de propagation mesuré. L_F est l'affaiblissement en espace libre et $A_{MU}(f,d)$ est un terme de correction de la propagation libre (il tient intuitivement compte de la diffraction et des réflexions au sol). Intuitivement, car il est déduit des mesures.

$G(h_{re})$ et $G(h_{te})$ sont des facteurs de gain complémentaires, tenant compte respectivement de la hauteur du récepteur et de l'émetteur. [45]
Enfin G_{AREA} est un facteur correctif qui tient compte du type d'environnement étudié (rural, suburbain, urbain).
Le modèle d'Okumura est encore aujourd'hui considéré comme parmi les plus simples et les meilleurs en terme de précision pour les systèmes cellulaires en zones urbaines ou rurales, présentant des zones spécifiques de propagation (forêts, mer, urbain, montagne…). Ce modèle est devenu le standard au Japon pour le déploiement des réseaux cellulaires.
L'inconvénient majeur de ce modèle est qu'il repose sur des valeurs caractéristiques moyennes de terrain, et s'adapte mal aux zones où les propriétés varient rapidement. Ainsi ce modèle est assez performant en zone urbaine et suburbaine, mais beaucoup moins bon pour les zones rurales. L'écart-type moyen entre les prédictions et les mesures avec ce modèle, tourne autour de 10 à 14dB. [45]
Mesure de l'affaiblissement :

$$L_{50}(urban) = 69.55 + 26.16 \cdot \log_{10} f_c - 13.82 \cdot \log_{10} h_{te} - a(h_{re}) + (44.9 - 6.55 \log_{10} h_{te}) \cdot \log_{10} d$$

$$L_{50}(suburban) = L_{50}(urban) - 2 \cdot [\log_{10}(f_c / 28)]^2 - 5.4$$

$$L_{50}(rural) = L_{50}(urban) - 4.78 \cdot [\log_{10} f_c]^2 - 18.33 \cdot \log_{10} f_c - 40.98$$

f_c est la fréquence centrale utilisée, en MHz, entre 150 et 1500MHz, h_{te} et h_{re} sont respectivement les hauteurs effectives de la station de base et du récepteur. La première est comprise entre 30m et 200m, et la 2ième entre 1m et 10m. d est la distance émetteur-récepteur.
La fonction $a(h_{re})$ qui est la fonction de correction d'antenne mobile, est le seul facteur qui s'adapte en fonction de la taille de la ville.
Pour les villes petites et moyennes, il est donné par :

$$a(h_{re}) = (1.1 \log_{10} f_c - 0.7) h_{re} - (1.56 \log_{10} f_c - 0.8)$$

Et pour les grandes villes :

$$a(h_{re}) = 8.29 \cdot (\log_{10} 1.54 h_{re})^2 - 1.1 \quad f_c \leq 300MHz$$

$$a(h_{re}) = 3.2 \cdot (\log_{10} 11.75 h_{re})^2 - 4.97 \quad f_c \geq 300MHz$$

Ce modèle référencé comme celui d'Okumura-Hata, s'est montré particulièrement efficace dans les zones urbaines et suburbaines, pour le déploiement des grandes cellules des systèmes mobiles, mais s'avère peu efficace pour les cellules de taille inférieure à 1km.
Pour cette raison, dans les scénarios de simulation, la cellule couverte par le NodeB a un diamètre de 2km. [45]

Implémentation 4 : Définition des nœuds

Après les modifications apportées aux éléments ci-dessus, les procédures d'instanciation de l'AP, du NodeB ainsi que du terminal bi-mode ont été implémentées. Elles permettent la définition rapide des éléments du réseau.

Architecture définie

Par ces implémentations, le terminal mobile est simulé par cette structure de protocoles et entité :

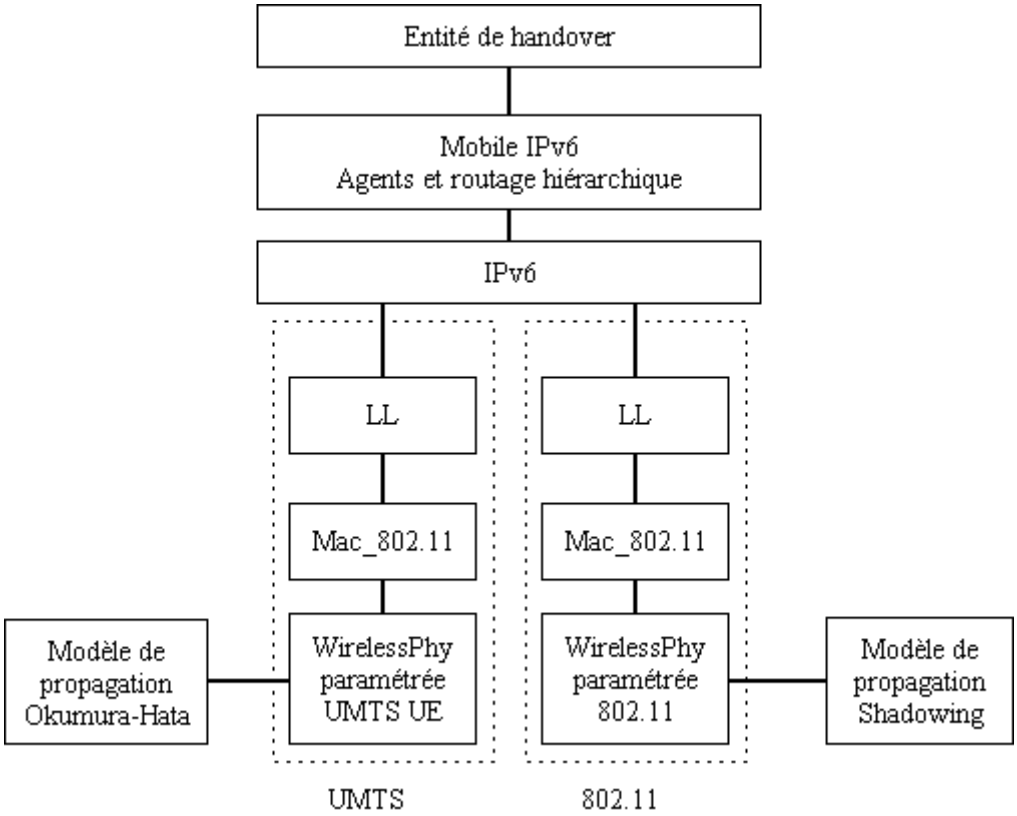

Figure 19 – Schéma de l'architecture du terminal mobile bi-mode dans NS-2

Le NodeB UMTS par cette structure de nœud :

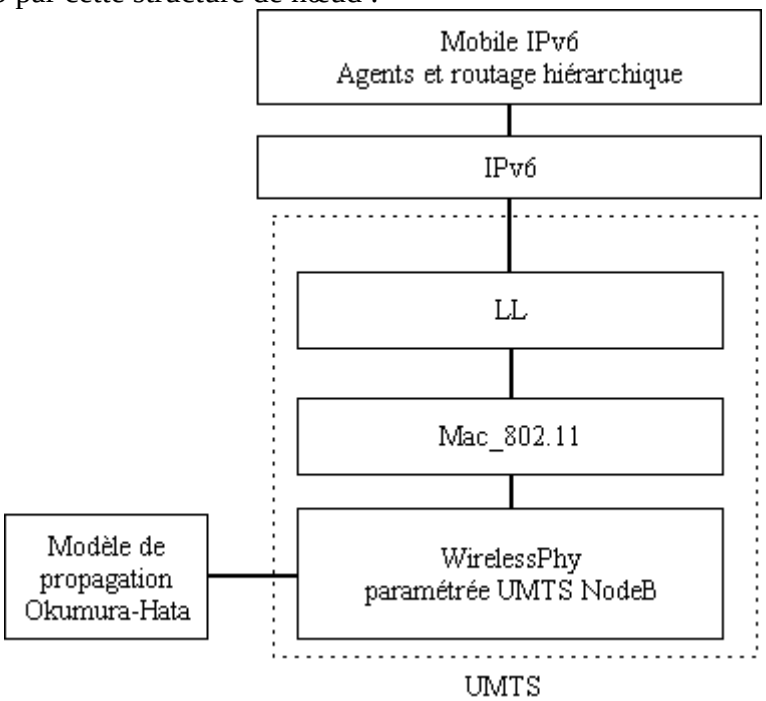

Figure 20 – Schéma d'architecture du NodeB dans NS-2

Enfin, l'AP 802.11 par cette structure de nœud :

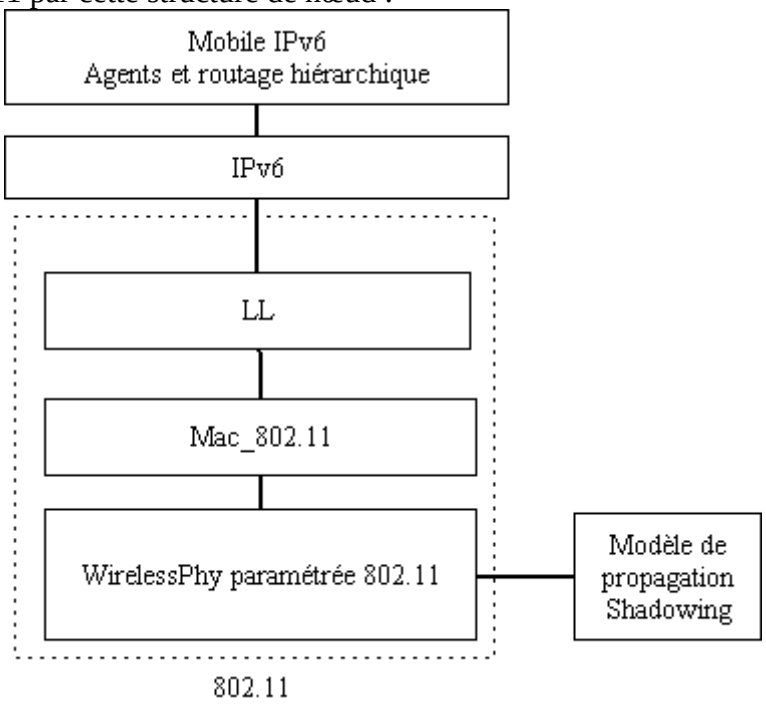

Figure 21 – Schéma d'architecture de l'AP dans NS-2

Scénario de simulation

Après avoir détaillé l'architecture proposée, on propose d'analyser le temps de basculement de l'AP au NodeB ainsi que la charge de signalisation introduite par Mobile IPv6. Le scénario envisagé emploie le protocole de transport en mode connecté TCP. Ce choix est motivé par le fait que TCP est plus exigeant que UDP en terme de perte de paquet. Or l'architecture mise en œuvre doit offrir ce service connecté tout en minimisant le temps de basculement d'un réseau à l'autre (point de vue utilisateur) et la charge de signalisation inhérente à l'utilisation d'un protocole de mobilité (point de vue opérateur).

Mesure du temps de basculement

Suite à la description du comportement de Mobile IPv6, on identifie les délais contribuant au temps de basculement de la connexion du terminal mobile lors du handover.
Les délais contribuant au temps de basculement sont :

- Le temps de détection de mouvement (td) : le temps requis par le mobile pour détecter et se rendre compte qu'il s'est déplacer vers une nouvelle station de base (découverte d'un nouveau routeur).
- Le temps de configuration de la care-of adresse (ta) : le temps entre la prise de conscience du déplacement du mobile et le temps nécessaire à la configuration de la care-of adresse dans le routeur.
- Le temps d'enregistrement de binding (tr) : le temps entre l'envoi de la mise à jour de binding vers le home agent et la réception de l'acquittement de binding.
- Le temps d'optimisation de route (to) : le temps nécessaire depuis l'enregistrement de la nouvelle care-of adresse avec le home agent à la fin de l'optimisation de routage du nœud correspondant (serveur).

Le temps de basculement = th = td + ta + tr + to.

Mesure de la charge de signalisation

La charge de signalisation du réseau engendrée par l'utilisation de Mobile IPv6 comprend :

- Les paquets d'annonce de routeur qu'ils soient réceptionnés ou non par le mobile et les paquets d'acquittement de la réception d'annonce par le mobile
- Les paquets de mise à jour de binding envoyés au home agent et les paquets d'acquittement de l'enregistrement de la nouvelle care-of adresse
- Les paquets de mise à jour de binding envoyés au nœud correspondant ainsi que les paquets d'acquittement du binding
- Les paquets de sollicitation envoyés par le mobile pour forcer l'annonce de routeur.

La charge de signalisation s'exprime donc :

$$C_S = (F_A \cdot T_{PA}) + (F_{HO} \cdot T_{SOL}) + 2 \cdot (F_{HO} \cdot T_{BU})$$

Avec Fa la fréquence d'annonce de routeur du NodeB et de l'AP, Tpa la taille des paquets d'annonce de routeur, Fho la fréquence de sollicitation de handover du terminal, Tsol la taille des paquets de sollicitation d'annonce + la taille des paquets d'annonce renvoyés au terminal et Tbu la taille des paquets (envoyés au nœud correspondant et au home agent) de mise à jour de binding + la taille des paquets d'acquittement de binding.

Scénario TCP

Dans ce scénario, le terminal mobile se déplace de façon aléatoire sur la topologie. L'environnement configuré est urbain pour le modèle de propagation Okumura-Hata tandis que le modèle de propagation Shadowing est d'indice n = 3.7 (situation indoor).
La cellule du NodeB couvre toute la topologie géographique définie en début de simulation.
La basic service area de l'AP est incluse dans la zone de couverture du NodeB.
Le terminal mobile est connecté au CN via l'AP par une connexion TCP et débute un téléchargement FTP depuis le CN.
On spécifie une taille de paquets de 1000-bits envoyés à un débit de 200kbps.
Les liaisons filaires entre CN, router, NodeB et AP telles que schématisées ci-dessous ne sont pas limitatives (liaisons 100Mbps).

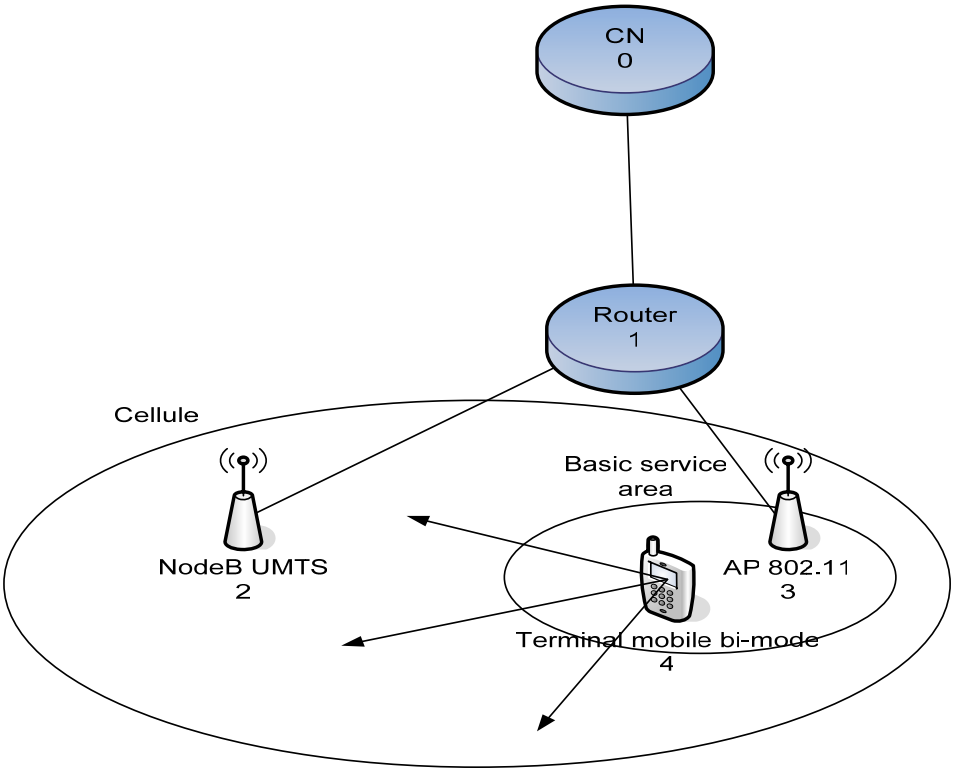

Figure 22 – Topologie de simulation

Mesure du temps de basculement

Après avoir répété la simulation de 1000 secondes à 3 reprises, une itération est présentée.
Les paramètres mesurés durant les deux autres itérations ont appuyé les mesures de l'itération présentée.
Le terminal mobile se déplaçant de façon aléatoire, le handover est détecté lorsque le terminal configure une nouvelle care-of adresse et qu'une entrée est ajoutée simultanément à la Binding Update List du terminal. [44]

Le dump de MobiWan est fourni ci-dessous :

```
1.00233 get_coa for BS 1.2.0:4198400
322.063 get_coa for BS 1.2.0:4198400
501.605 get_coa for BS 1.2.0:4198400


|Binding Cache for node 1.1.0 at 1000 ---------------------------------- |
|Node    COA   Type    Info    Flag    Last  Time     Life    Expire    Nb|
|1.1.1   1.2.4  7       MN      1       152  646.796   10       0      80 |


|Binding Cache for node 0.0.0 at 1000 ---------------------------------- |
|Node    COA   Type    Info    Flag    Last   Time    Life    Expire    Nb|
|1.1.1  1.2.4   7       MN      1       153  646.799   10       0      73 |

|Binding Update List for node 1.1.1 at 1000 ---------------------------- |
|Node    COA    Type    Info    Flag    Last  Time    Life    Expire    Nb|
|1.2.0   1.2.4   4       BS      1       -1   501.605  10     511.605   1 |
|0.0.0   1.2.4   5       CN      1       155  656.605  10     658.615  74 |
|1.1.0   1.2.4   3       HA      1       189  996.605  10    2.6e+08  115 |


|Base Station List for node 1.1.1 at 1000 ------------------------------ |
|Node    COA    Type    Info    Flag    Last  Time    Life    Expire    Nb|
|1.2.0   1.2.4   8       BS      1       -1  999.817   994               |


|History List for node 1.1.1 at 1000 ---------------------------------- |
|Node    COA    Type    Info    Flag    Last  Time    Life    Expire    Nb|
|1.2.0   1.2.4   4       BS      1       -1  322.063   10     332.063   1 |
```

Figure 23 – Dump MobiWan

La trace d'envoi des messages de signalisation du terminal mentionne l'envoi de la sollicitation du terminal au routeur (NodeB) :

```
s 501.400000000 _4_ AGT  --- 36761 ipv6_sol 48 [0 0 0 0] ------- [4196353:0
-1:0 32 0]
```

L'identifiant du terminal est 4 et son adresse initiale 1.1.1. Le type de paquet envoyé par le terminal est une sollicitation de routeur (ipv6_sol). Le terminal s'apprête donc à effectuer un handover. Le fichier de « Dump MobiWan » mentionne également un autre handover vers le NodeB en 322.063 enregistré dans la History List. La History List stocke les entrées de la Binding List d'un nœud dans le but de ne pas surcharger la Binding List en cas de simulations impliquant de nombreux handovers. [44]

Suite à l'envoi de la sollicitation, le terminal réceptionne l'annonce du NodeB et configure ensuite sa nouvelle care-of adresse : `501.605 get_coa for BS 1.2.0:4198400`
La BS d'adresse 1.2.0 est le NodeB alors que l'AP est d'adresse 1.1.0. Le terminal étant initialement attaché à l'AP, il obtient une care-of adresse pour pouvoir s'attacher au NodeB.

La Binding List du terminal indique que le terminal enregistre la care-of adresse (COA) 1.2.4. On remarque que la durée de validité de la care-of adresse est de 10 secondes, après ce délai, le terminal doit enregistrer une nouvelle care-of adresse s'il se trouve toujours dans un réseau étranger. Suite à l'enregistrement de la care-of adresse, le terminal envoie une mise à jour de binding à son home agent 1.1.0 ainsi qu'au nœud correspondant 0.0.0 (CN).

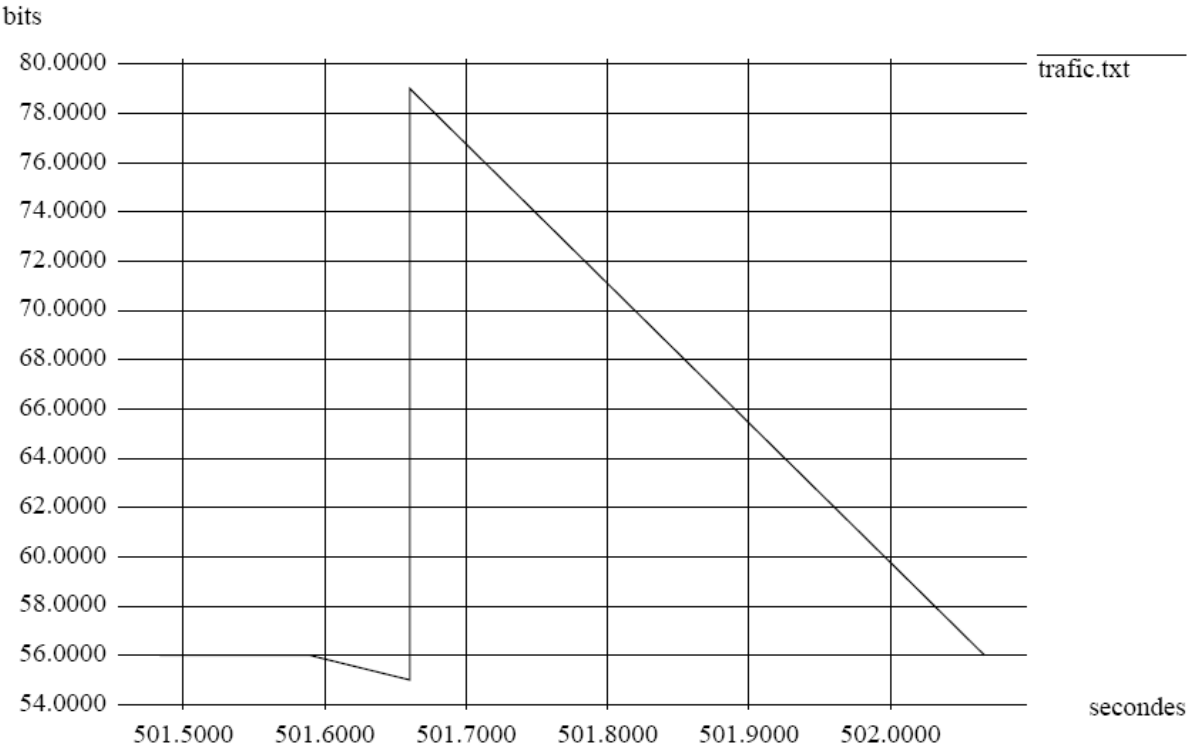

Figure 24 – Envoi de l'acquittement de binding du CN au terminal

Le terminal réceptionne ensuite l'acquittement de mise à jour de binding :

```
r 502.048655704 _4_ AGT  --- 36785 mipv6_back 79 [13a 2 1 800] -------
[4196352:0 4198404:2 28 4198404]
```

La réception de l'acquittement marque la fin de l'exécution du handover.

Le temps de basculement moyen pour les 3 itérations effectuées est de 648.5 ms.

Conclusion :

Un tel temps de basculement entre l'AP et le NodeB est inacceptable pour les applications soumises à des contraintes temps réel (QoS Conversational et Streaming UMTS). Cependant, dans un contexte de QoS UMTS Background correspondant à un service Best Effort 802.11 et un contexte de QoS Interactive correspondant à service Low Loss 802.11, le temps de basculement est acceptable.

Les exigences de QoS Interactive / Low Loss sont remplies vu le protocole de transport TCP utilisé.

QoS Mapping	
UMTS	*WLAN*
Conversational	Low-Latency & Low Jitter
Streaming	Low Latency
Interactive	Low Loss
Background	Best-Effort

Figure 25 – QoS Mapping UMTS – 802.11 [36]

Mesure de la charge de signalisation

Passons à présent à la mesure de la charge de signalisation supportée par le réseau. La charge de signalisation considérée ne concerne que les paquets de signalisation Mobile IPv6.
Pratiquement, on observe une charge de signalisation réseau (CN, router, NodeB et AP) et une charge de signalisation du terminal mobile.

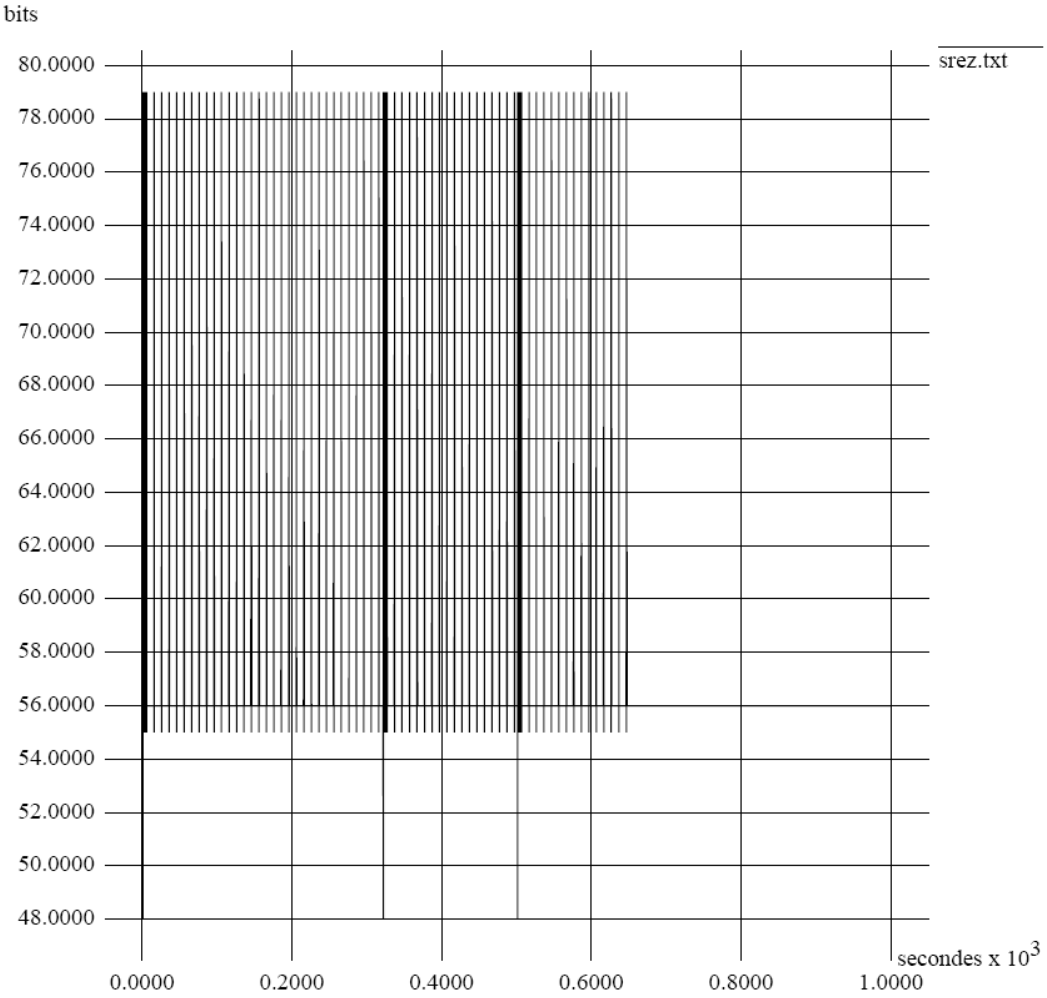

Figure 26 – Charge de signalisation réseau

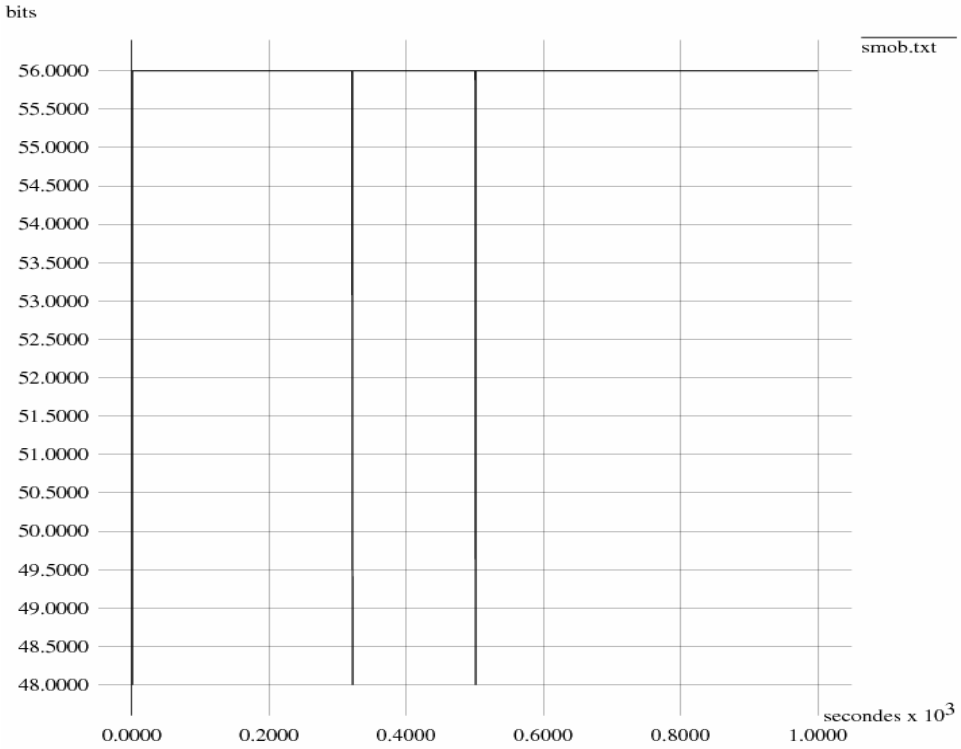

Figure 27 – Charge de signalisation terminal

On constate que la charge de signalisation du terminal est presque constamment à 56-bits, la charge de signalisation du réseau oscille elle entre 55-bits et 79-bits.

Conclusions :

La charge de signalisation du mobile est due à aux annonces de routeur qui pénalisent, de par leur fréquence d'émission, la réception de paquets IP augmentés de 56-bits. Lors des deux handovers, le terminal envoie une sollicitation de 48-bits.

Il est moins évident d'exprimer la charge de signalisation réseau vu ses oscillations. En l'exprimant par sa valeur moyenne pondérée, on détermine une charge de signalisation de 57.287-bits. Durant ce scénario se sont produits deux handovers de l'AP vers le NodeB et un handover du NodeB vers l'AP.

Le scénario de handover vertical envisagé est particularisé étant donné qu'il s'agit d'une simulation temporelle. Cependant, l'emploi de déplacements aléatoires du terminal assure une certaine robustesse des mesures.

Conclusion

Le projet développé ici a consisté à définir une architecture logicielle capable de supporter le handover vertical entre les réseaux UMTS et 802.11. Un comparatif qualitatif est établi entre les architectures réseau UMTS et 802.11. Un comparatif tend également à sélectionner un protocole de mobilité en fonction des critères de déploiement et de services supportés. Nous avons donc identifié les challenges technologiques entre les deux technologies réseau et fourni une solution au problème, déjà évoqué, de la mobilité entre les réseaux UMTS et 802.11.

En plus des challenges technologiques, se sont ajoutés des challenges pratiques liés au simulateur logiciel NS-2 tant pour l'interconnexion des réseaux UMTS et 802.11 que pour l'utilisation du protocole de mobilité Mobile IPv6. Consécutivement à l'analyse de ce challenges pratiques, une architecture de développement a été proposée pour pouvoir simuler un scénario de transfert de connexion TCP entre un point d'accès UMTS et un point d'accès 802.11.

L'importance d'une entité de décision pour effectuer le transfert de connexion a été mise en évidence. Cette entité de handover a été particularisée pour effectuer le transfert de connexion en fonction des puissances des stations de base 802.11 et UMTS reçues par le terminal mobile bi-mode.

Enfin, par l'étape de simulation du scénario de handover vertical, on a souligné les problématiques liées au temps de basculement du réseau 802.11 au réseau UMTS ainsi que la charge de signalisation résultant de l'utilisation du protocole Mobile IPv6.

D'une part, le scénario de handover vertical inter-système entre les réseaux UMTS et 802.11 modélisés met en avant les exigences de QoS acceptables lors du déplacement du terminal bi-mode, en autorisant, en terme de perte de paquet et délai, un handover intégré pour des applications de type Background telles que le browsing web, la consultation de mail ou le téléchargement FTP et de type Interactive telles que les jeux multi-joueurs de stratégie temps réel ou au tour par tour. En revanche, les exigences de QoS ne sont pas satisfaites pour des applications temps réel telles que la voix, la visioconférence ou le streaming vidéo.

D'autre part, on retire le problème de la charge de signalisation due au protocole de mobilité Mobile IPv6. Ce problème est susceptible d'engendrer une consommation de bande passante tant pour l'opérateur réseau que pour l'utilisateur du terminal mobile. L'utilisateur du terminal mobile doit s'attendre à une pénalité de bande passante due aux annonces de routeur fréquentes ; l'opérateur réseau subit, quant à lui, une pénalité importante en considérant un nombre élevé de stations de base UMTS et 802.11 supportant Mobile IPv6. La présence d'une entité de handover est importante afin de limiter cette pénalité. Elle entraîne des délais de mesure, décision et exécution du handover supplémentaires dans le temps de basculement mais permet de réduire considérablement les annonces de routeur, consommatrices principales de bande passante. Une diminution voire une suppression des annonces de routeur Mobile IPv6 non sollicitées est une solution à ce problème de charge au vu de l'entité de handover implantée dans le terminal mobile. Les exigences de QoS UMTS Background et Interactive demeureront respectées en limitant considérablement les charges de signalisation réseau et terminal

Travail futur

Les travaux de recherche future comprennent essentiellement l'implémentation de protocoles ou mécanismes rendant l'architecture modélisée davantage en accord avec les spécifications UMTS.

- Implémentation des protocoles d'interfaces terrestres UMTS appartenant au Control Plane. Ces protocoles sont susceptibles d'alourdir la charge de signalisation hors Mobile IPv6.

- Implémentation des protocoles d'interface radio Uu afin de tenir compte des mécanismes de retransmission spécifiques à RLC, de MAC UMTS notamment pour sa fonction de gestion de flux influant sur la QoS.

- Implémentation de l'interface physique WCDMA pour modéliser cette interface en accord avec les spécifications 3GPP, par opposition à la couche physique sans fil générique employée. Le même commentaire est valable pour l'interface physique HR-DS ou OFDM dont l'implémentation est absente de NS-2.

- Implémentation des nœuds définis dans les spécifications 3GPP (GGSN, SGSN, RNC) pour une simulation plus rigoureuse des délais de traversées du réseau UMTS. Ce point inclut également la possibilité de simuler le handover vertical dans les configurations d'interconnexion open coupling, loose coupling, tight coupling

- Permettre la simulation de handover de plus d'un terminal mobile simultanément – voir code de MobiWan.

- Envisager le protocole SIP ou Fast Mobile IP si un contexte de QoS Conversationnel ou Streaming est à considérer (usage du protocole de transport UDP).

- Envisager d'autres scénarios de simulations réalistes en considérant une entité de handover basée sur un taux d'erreur (bit error rate / block error rate) et en la confrontant, de même que Mobile IPv6 à d'autres facteurs d'entrée. Ces facteurs d'entrée peuvent être le type d'environnement géographique, le protocole de transport orienté connexion ou non, le type d'application plus ou moins sensible à un niveau de QoS élevé.

Bibliographie

Introduction :

[a] Miller, J.A. : « The State of Handhelds & VoIP », 2005

[b] Murray, K. & Pesch, D. : « Policy Based Access Management and Handover Control in Heterogeneous Wireless Networks », 2002

[c] Politis, C. *et al. :* « Hybrid multilayer mobility management with AAA context transfer capabilities for all-IP networks », IEEE WC, 2004

[d] Akyildiz I. *et al. :* « A survey of mobility management in next generation all-IP based wireless systems », IEEE WC, 2004

[e] Ma, L. *et al. :* « A new method to support UMTS/WLAN handover using SCTP », IEEE WC, 2004

UMTS :

[1] UMTS Protocols and Protocol Testing, http://www.iec.org

[2] Walke, B.: « Mobile radio networks », 2nd edition, John Wiley, 2001, p. 433-523.

[3] Holma, H. and Toskala, A.: « WCDMA for UMTS », 3rd edition, John Wiley, 2004.

[4] Walke, B., Seidenberg, P. and Althoff, M.P.: « UMTS – The Fundamentals », 1st edition, John Wiley, 2003.

[5] 3GPP TS 23.121 v3.6.0: « 3rd Generation Partnership Project; Technical Specification Group Services and Systems Aspects; Architectural Requirements for Release 1999 (Release 1999) », 3GPP, 1999.

[6] Wisely, D., Eardley, P. and Burness, L.: « IP for 3G – Networking Technologies for Mobile Communications », 1st edition, John Wiley, 2002.

802.11 :

[7] Crow, B. P. *et al.*: « IEEE 802.11 Wireless Local Area Networks », IEEE Communications Magazine, 1997, vol. 35, issue 9, p. 116-126.

[8] Gast, M. S.: « 802.11 Wireless Networks – The Definitive Guide », 1st edition, O'Reilly, 2002.

[9] IEEE Std 802.11-1999: « 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, 1999.

[10] http://www.umtsworld.com/technology/overview.htm

[11] http://www.intel.com/business/bss/infrastructure/wireless/solutions/technology.htm

[12] Jansen, P. and Nilsen, P.: « Next generation mobile communication infrastructure: UMTS and WLAN – who will succeed? », draft paper, Department of Informatics, University of Oslo, 2002.

[13] IEEE Std 802.11b-1999: « Supplement to 802.11-1999, Wireless LAN MAC and PHY specifications: Higher speed Physical Layer (PHY) extension in the 2.4 GHz band », IEEE, 1999.

[14] IEEE Std 802.11b-1999/Cor1-2001: « 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: Higher-speed Physical Layer (PHY) extension in the 2.4 GHz band – Corrigendum 1 », IEEE, 2001.

Protocoles de mobilité :
[15] Perkins, C.E.: « Mobile IP », IEEE Communications Magazine, 1997, vol. 35, issue 5, p. 84-99.
[16] Perkins, C.E.: « Mobile Networking Through Mobile IP », IEEE Internet Computing, 1998, vol. 2, issue 1, p. 58-69.
[17] Perkins, C.: « IP Mobility Support », RFC 2002, 1996, p. 1-79.
[18] Nokia: « Introducing Mobile IPv6 in 2G and 3G mobile networks », Nokia, 2001, p. 1-16.
[19] Johnson, D.: « Mobility Support in IPv6 », RFC 3775, 2004, p. 1-165.
[20] Foo, S. *et al.*: « Approaches for resolving dynamic IP addressing », Internet Research, 1997, vol. 7, issue 3, p. 208-216.
[21] Riegel, M. and Tüxen, M.: « Mobile SCTP: Transport Layer Mobility Management for the Internet », SoftCOM 2002, 2002.
[22] Ma, l. *et al.*: « A New Method to Support UMTS/WLAN Vertical Handover Using SCTP », IEEE Vehicular Technology Conference, 2003, vol. 3, p. 1788-1792.
[23] Ong, L. *et al.*: « An Introduction to the Stream Control Transmission Protocol (SCTP) », RFC 3286, 2002, p. 1-10.
[24] Stewart, R. *et al.*: « Stream Control Transmission Protocol », RFC 2960, 2000, p. 1-134.
[25] Stewart, R. *et al.*: « Stream Control Transmission Protocol (SCTP) Dynamic Address Reconfiguration », Internet draft, draft-ietf-tsvwg-addip-sctp-08.txt, 2003, p. 1-37.
[26] Koh, S.J. *et al.*: « Mobile SCTP for Transport Layer Mobility », Internet draft, draftsjkoh-sctp-mobility-03.txt, 2004, p. 1-14.
[27] Wedlund, E. and Schulzrinne, H.: « Mobility Support using SIP », ACM/IEEE International Conference on Wireless and Mobile Multimedia, 1999, p. 1-7.
[28] Schulzrinne, H. and Wedlund, E.: « Application–Layer Mobility Using SIP », IEEE Service Portability and Virtual Customer Environments, 2001, p. 1-9.
[29] Rosenberg, J. *et al.*: « SIP: Session Initiation Protocol », RFC 3261, 2002, p. 1-269.
Entité de handover :
[30] Pahlavan, K. *et al.*: « Handoff in Hybrid Mobile Data Networks », IEEE Personal Communications, 2000, vol. 7, issue 2, p. 34-47.
[31] Niedermeier, C. *et al.*: « Handover Management and Strategies for Reconfigurable Terminals », SDR Forum Document Number SDRF-02-I-0047-V0.00, 2002.
[32] Alsenmyr, G. *et al.*: « Handover between WCDMA and GSM », Ericsson Review, 2003, vol. 80, issue 1, p. 6-11.
[33] Kapoor, S.: « Mobile-Controlled Handoff for MBWA », IEEE 802.20 Working Group on Mobile Broadband Wireless Access, 2003.
[34] Mohyeldin, E. *et al.*: « Concepts and Scenarios for Intersystem Handover in Heterogeneous Environments ».
[35] Freedman, A. and Hadad, Z.: « Handoff Schemes Overview and Guidelines for Handoff Procedures in 802.16 », IEEE 802.16 Broadband Wireless Access Working Group, 2002.
[36] Samarasinghe, R. *et al.* : « Analysis of Intersystem Handover: UMTS FDD & WLAN » Centre for Telecommunications Research, King's College London.

Modélisation :
[37] Greis, M. : « Tutorial for the Network Simulator ns », http://www.isi.edu/nsnam/ns/tutorial/
[38] Anelli, P. & Horlait, E. : « NS-2: Principes de conception et d'utilisation », UPMC, 2001
[39] Chung, J., Claypool, M. : « NS by Example », Worcester Polytechnic Institute
[40] Haldar, P. & Chen, X. : « NS Tutorial 2002 », USC/ISI, 2002
[41] Fall, K. &Varadhan, K. : « The *ns* Manual », UC Berkeley, 2005
[42] Haldar, P. : « Wireless world in NS », USC/ISI, 2002
[43] « EURANE User Guide (Release 1.3) », SEACORN Project, 2004
[44] Thierry Ernst, « MobiWan User Guide », Motorola Labs Paris, 2001
[45] Gorce, J.M. : « Techniques avancées pour les radiocommunications », INSA Lyon
Scénario :
[46] Dunmore, M. & Pagtzis, T. : « Mobile IPv6 Handovers: Performance Analysis and Evaluation »
[47] Bocq, G. : « Public networks (ELEC321) », cours, ULB, 2004
[48] El Malki, K. *et al. :* « Fast Mobile IP Handoffs in Cellular Systems », 2000

Annexe UMTS

Protocoles d'interfaces terrestres UMTS

Les structures de protocole dans les interfaces terrestres UTRAN sont conçues selon le même modèle de protocoles. La structure est basée sur le principe que les couches et plans sont logiquement indépendants les uns des autres. Cette structure est composée de 2 couches : la couche de réseau radio (Radio Network Layer) composée de protocoles typiques à l'UTRAN et la couche de réseau transport (Transport Network Layer) composée de protocoles standards choisis pour être utilisé dans l'UTRAN. La structure est également divisée en plan de contrôle, plan utilisateur, plan de contrôle du réseau de transport et plan utilisateur du réseau de transport.

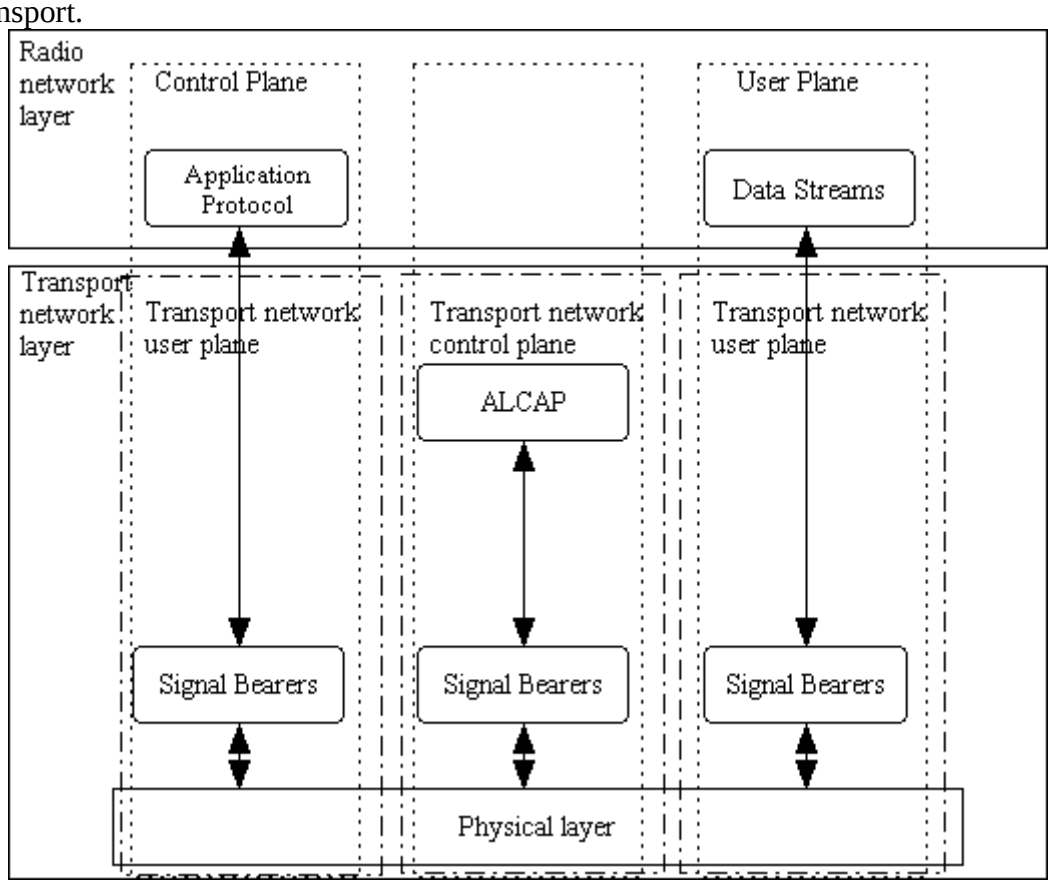

Figure 28 – Modèle de protocole général pour les interfaces terrestres UTRAN [3]

Le plan de contrôle (Control Plane) est utilisé pour toute la signalisation de contrôle spécifique à UMTS. Ceci inclut le protocole applicatif et les porteuses de signal pour le transport de message des protocoles applicatifs.

Le protocole applicatif est employé, entre autres, pour configurer les porteuses vers l'UE c-a-d la porteuse d'accès radio dans l'Iu et le lien radio dans l'Iur et l'Iub.

Pour l'interface Iu, le protocole utilisé est RANAP. Il inclut les fonctions suivantes :

- La gestion des porteuses d'accès radio, y compris leur établissement et libération
- La gestion des connexions Iu
- L'échange d'informations de localisation d'UE entre le RNC et le CN
- Les requetes d'appel du CN vers l'UE
- Le traitement de la surcharge et d'erreur

Pour l'interface Iur, le protocole utilisé est RNSAP. Il inclut les fonctions de :

- gestion des liens radio, liens physiques et des ressources des canaux communs
- appel
- réattribution du SRNC
- mesures de ressources

Pour l'interface Iub, le protocole utilisé est NBAP. Il inclut les fonctions de :

- gestion des canaux communs, des ressources communes, et liens radio
- gestion de configuration comme la configuration des cellules
- traitement et contrôle des mesures
- synchronisation
- report d'erreurs

Pour l'interface radio Uu, l'architecture des protocoles de l'interface radio est explicitée au chapitre suivant.

Le plan utilisateur (User Plane) sert au transport de toute l'information envoyée et reçue par l'utilisateur, comme la voix ou les paquets IP. Le plan inclut les flux de données (Data Streams) caractérisés par des protocoles de trame, et les porteuses de données.

Le plan de contrôle du réseau de transport (Transport Network Control Plane) est utilisé pour toute la signalisation de contrôle nécessaire à la configuration des porteuses de transport du User Plane. Le protocole ALCAP permet la collaboration du Control Plane et du User Plane, ces 2 plans étant indépendants l'un de l'autre.

A titre d'illustration, voici les protocoles employés pour les interfaces Iu, Iur et Iub :

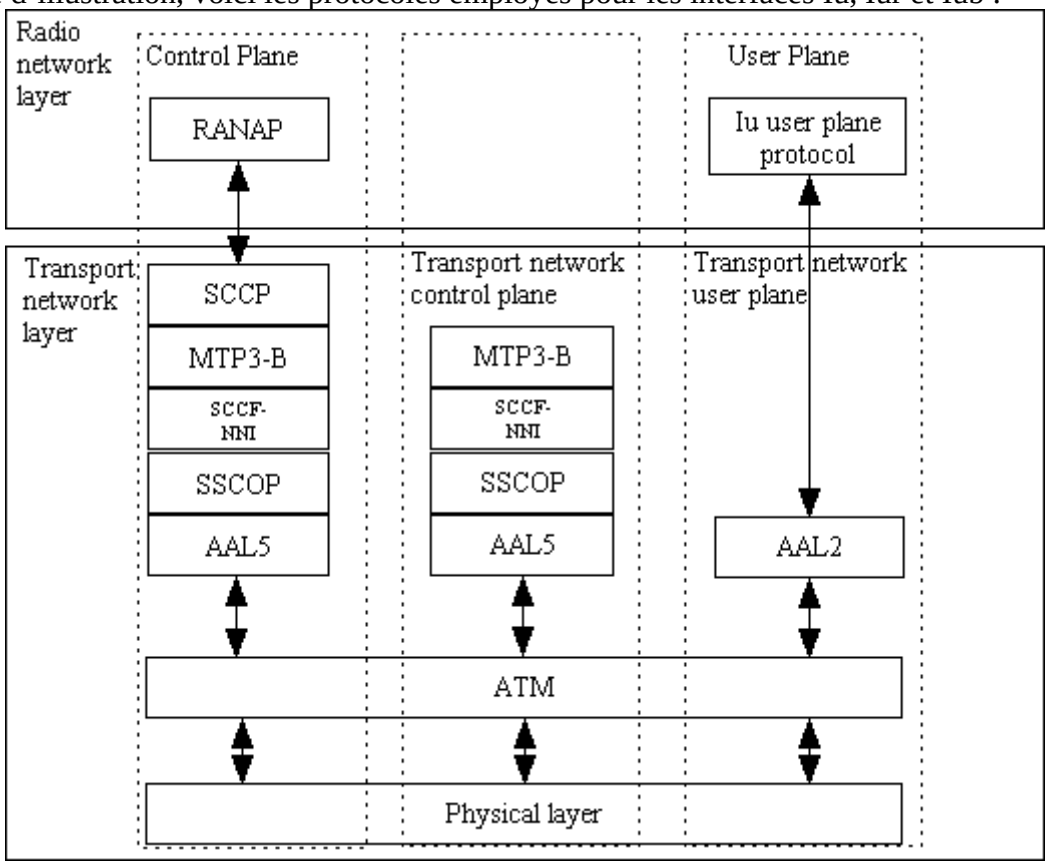

Figure 29 – Modèle de protocole de l'interface Iu [3]

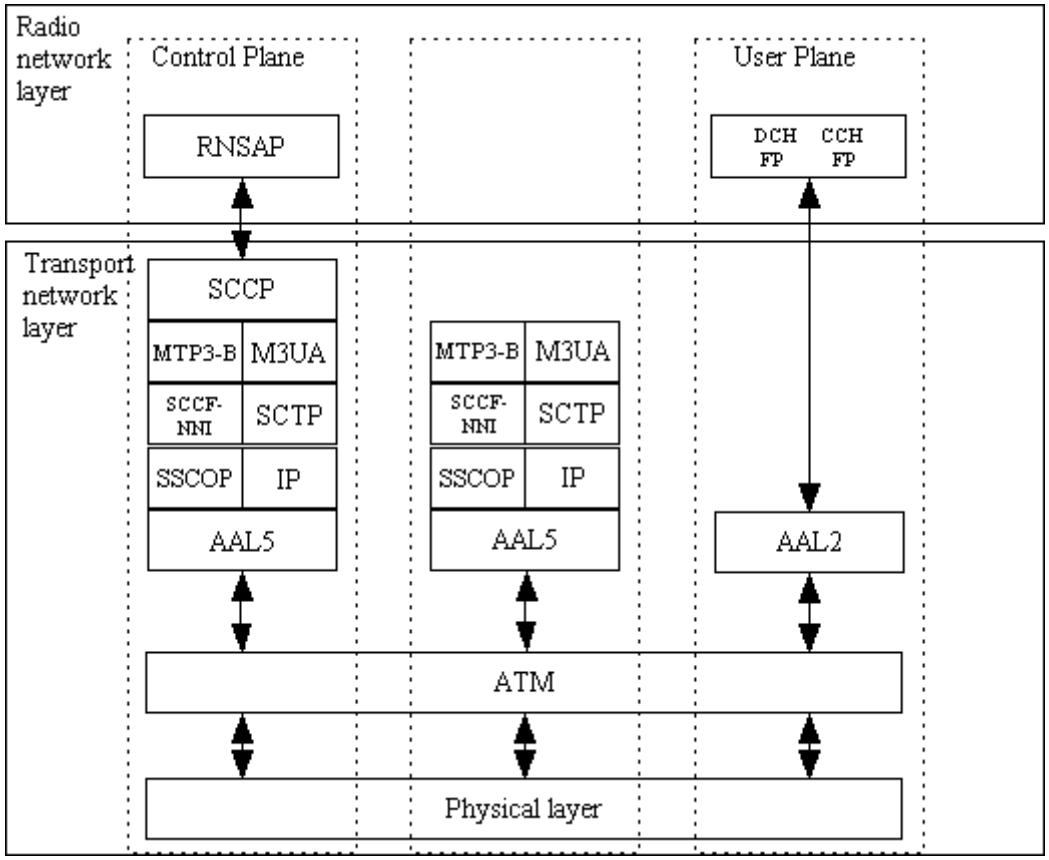

Figure 30 – Modèle de protocole de l'interface Iur [3]

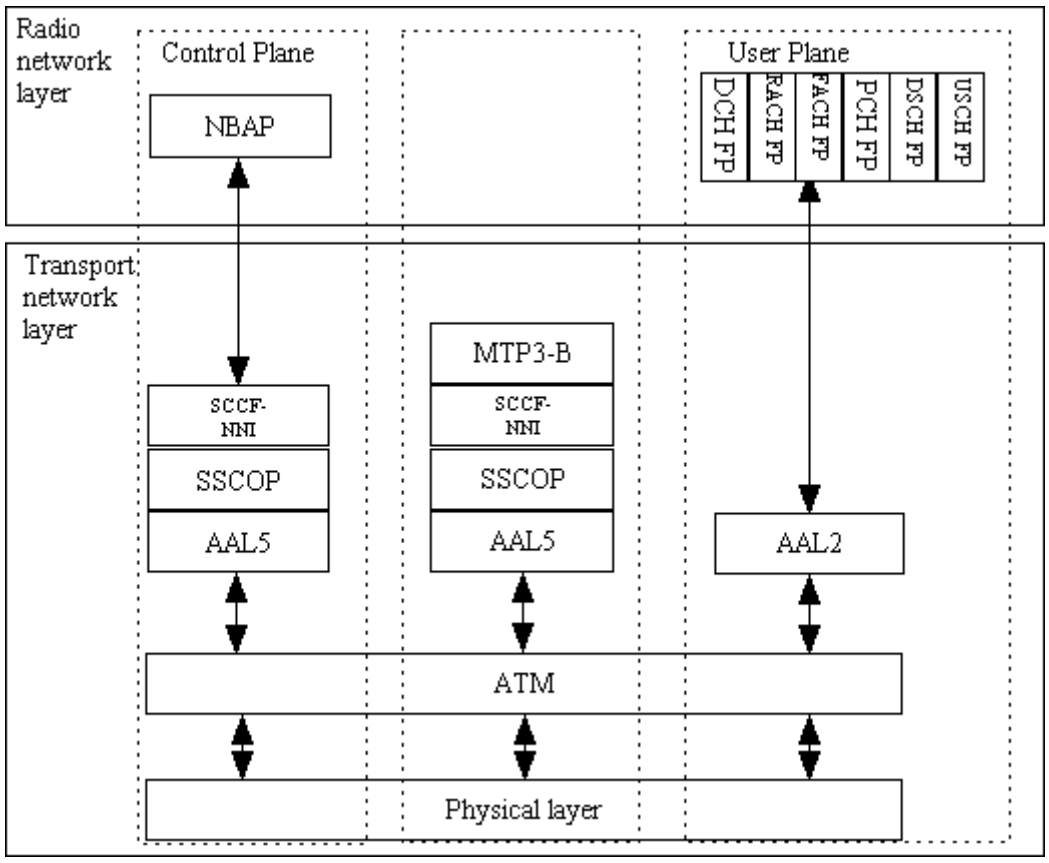

Figure 31 – Modèle de protocole de l'interface Iub [3]

Protocoles d'interface radio

L'architecture des protocoles d'interface radio (Uu) est organisée en couche réseau (Layer 3), couche liaison (Layer 2) subdivisée en protocole MAC (medium access control) et RLC (radio link control) et couche physique décrite dans le chapitre suivant.

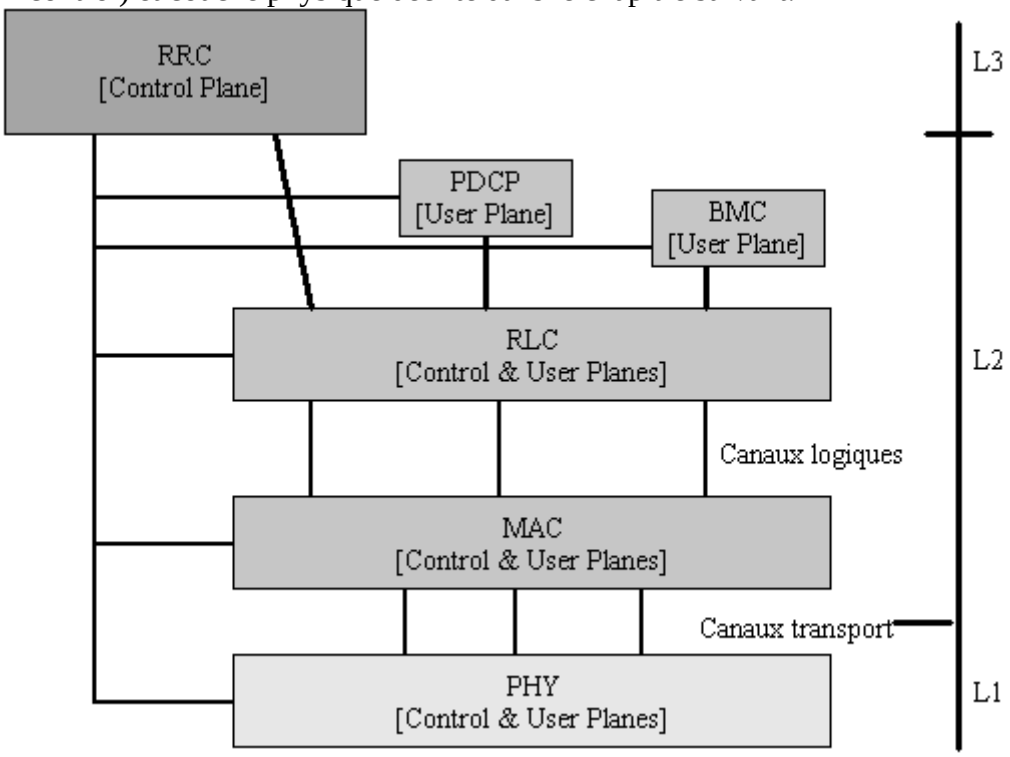

Figure 32 – Architecture des protocoles d'interface radio [3]

La couche PHY offre des services à la couche MAC via des canaux de transport. Ces canaux sont caractérisés en fonction des caractéristiques de données à transmettre et de la façon de les transmettre.

La couche MAC offre des services à la couche RLC au moyen de canaux logiques caractérisant le type de données à transmettre.

Protocole MAC

Dans la couche MAC, les canaux logiques sont mappés vers les canaux de transport. La couche MAC est aussi responsable de sélectionner le format de transport (TF) approprié pour chaque canal de transport en fonction du débit des canaux logiques.

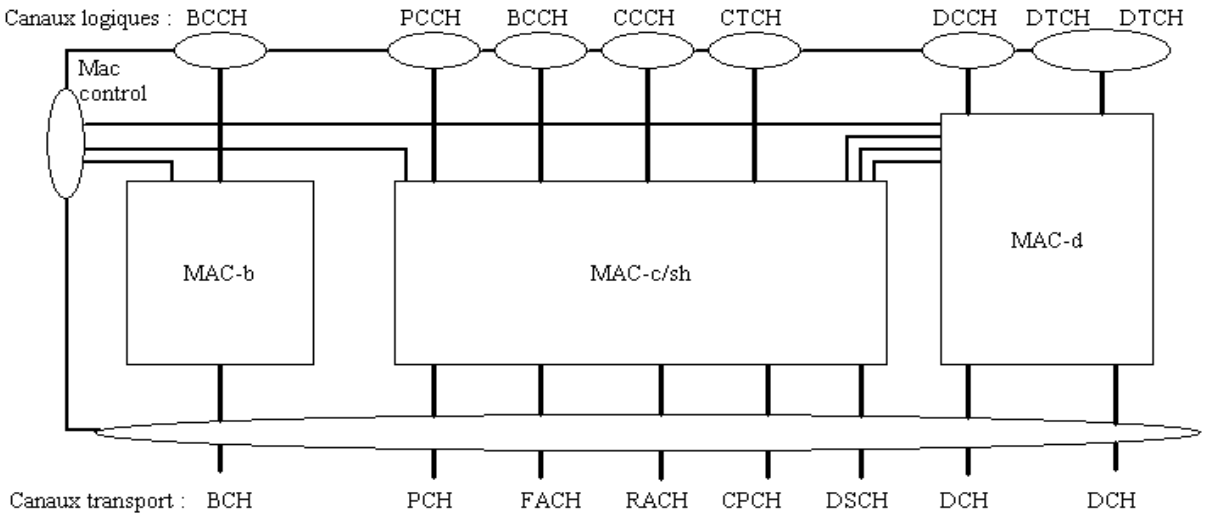

Figure 33 – Architecture de couche MAC [3]

La couche MAC consiste en 3 entités logiques :
MAC-b se charge du canal de broadcast (BCH). Il y'a une entité MAC-b dans chaque UE et une pour chaque cellule (localisée dans le nodeB).
MAC-c/sh se charge des canaux communs et des canaux partagés. Il y'a une entité MAC-c/sh dans chaque UE et une pour chaque cellule (localisée dans le RNC).
MAC-d se charge de gérer les canaux dédiés (DCH) alloués à un UE en mode connecté. Il y'a une entité MAC-d dans l'UE et une pour chaque UE (localisée dans le SRNC).

La couche MAC inclut aussi les fonctionnalités :

- Gestion de la priorité entre UEs et flux de données d'une UE
- Identification des UEs sur les canaux de transport communs
- Multiplexage/démultiplexage de PDUs issus de couches supérieures vers/depuis des blocs de transport délivrés à la couche physique sur les canaux de transport.
- Surveillance du volume de trafic
- Commutation dynamique entre canaux de transport dédiés et communs
- Cryptage

Les services de transfert de données de la couche MAC sont fournis sur des canaux logiques. Il existe des canaux logiques de contrôle et des canaux logiques de trafic.
Les canaux de contrôle sont :
BCCH (Broadcast Control Channel) : un canal descendant pour l'information de contrôle du système de diffusion
PCCH (Paging Control Channel) : un canal descendant transférant les informations d'appel
DCCH (Dedicated Control Channel) : un canal point à point bidirectionnel transmettant les informations de contrôle dédiées entre un UE et le RNC.
CCCH (Common Control Channel) : un canal bidirectionnel pour la transmission d'information de contrôle entre le réseau et les UEs.

Les canaux de trafic sont :
DTCH (Dedicated Traffic Channel) : un canal montant/descendant point à point dédié à une UE pour le transfert d'information utilisateur.
CTCH (Common Traffic Channel) : un canal point à multipoint descendant pour le transfert d'information utilisateur pour un groupe d'UEs.

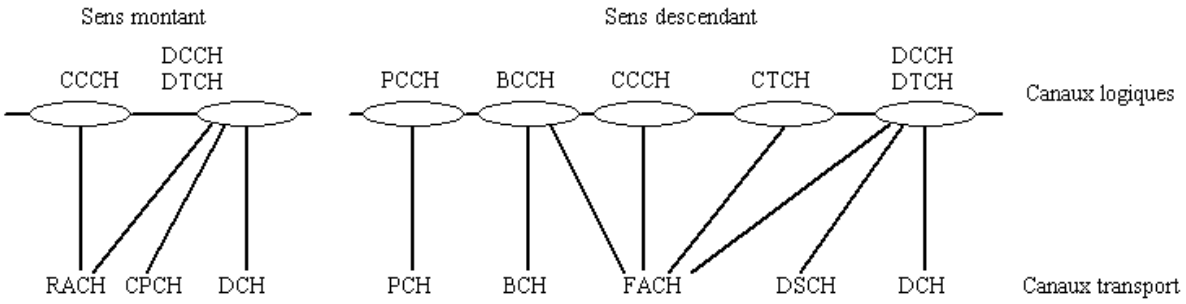

Figure 34 – Mapping entre canaux logiques et transport, sens montant et descendant [3]

Protocole RLC (Radio Link Control)

Le protocole RLC fournit les services de segmentation et retransmission pour les données utilisateur et les données de contrôle (les PDUs de longueur variable de couche supérieure). Chaque instance RLC est configurée par RRC pour opérer soit en mode transparent (Tr) soit en mode non acquitté (UM) ou encore en mode acquitté (AM).

En mode transparent, les services de segmentation et retransmission sont assurés de façon limitée. Aucun overhead n'est ajouté aux données de couche supérieure. Les PDUs erronés peuvent être supprimés ou marqués.

En mode non acquitté, aucun protocole de retransmission n'est employé et la réception de données transmises n'est pas garantie. La segmentation est fournie au moyen d'un header. La structure PDU inclut des numéros de séquence pour conserver l'intégrité des PDUs de couche supérieure.

En mode acquitté, un mécanisme de requête de répétition automatique (ARQ – Automatic Repeat reQuest) est employé pour la correction d'erreur. Si RLC ne peut délivrer les données correctement (nombre de retransmissions maximum atteint ou temps de transmission excessif), la couche supérieure en est informée et le SDU est abandonné. Le récepteur est aussi informé de l'abandon du SDU concerné.

Une entité RLC AM est bidirectionnelle et capable de « piggybacker » une indication sur le statut de la liaison dans les données utilisateur. RLC peut etre configuré pour une délivrance de PDU en séquence ou hors séquence. Dans le cas de délivrance en séquence, l'ordre des PDUs de couche supérieure est maintenu, alors que la délivrance hors séquence transfère les PDUs de couche supérieure dès qu'ils sont complètement reçus par RLC.

Protocole PDCP (Packet Data Convergence Protocol)

PDCP contient des méthodes de compression nécessaires pour obtenir une meilleure efficacité spectrale pour des services de transmission de paquets IP par radio. Entre autres, PDCP se charge de la compression d'information de contrôle redondante (e.g. les headers TCP/IP et UDP/IP) à l'entité de transmission et la décompression à l'entité de réception.

Protocole BMC (Broadcast/Multicast Control)

BMC est chargé de l'adaptation des services de diffusion et multicast provenant du domaine Broadcast sur l'interface radio. Le seul service utilisant ce protocole actuellement est le service SMS Cell Broadcast.

Protocole RRC (Radio Resource Control)

La majorité de la signalisation de contrôle entre l'UE et l'UTRAN consiste en messages RRC. Les messages RRC transportent tous les paramètres nécessaires à l'établissement, la modification et la libération d'entités protocolaires de Layer 2 et Layer 1. La mobilité de l'UE

en mode connecté est contrôlée par la signalisation RRC (mesures, handovers, mises à jour de cellule,…).

Etats de service RRC :

Les 2 modes opérationnels de base d'un UE sont le mode inactif (idle) et le mode connecté. Le mode connecté peut être divisé en états de service qui définissent quel type de canal physique est utilisé par l'UE.

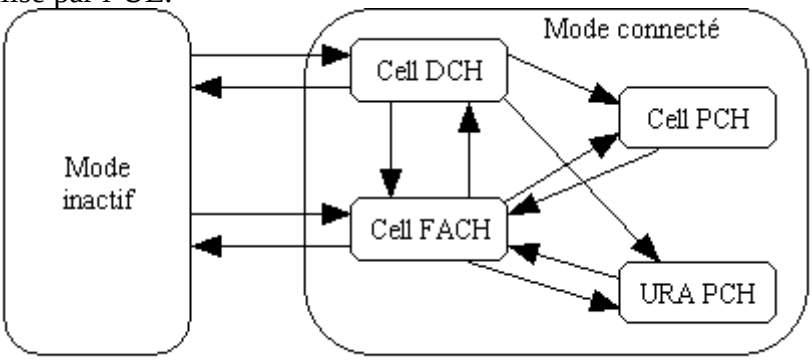

Figure 35 – Modes de l'UE et états RRC en mode connecté [3]

En mode inactif, l'UE est réglé sur le canal de contrôle d'une cellule du PLMN, en particulier une cellule couvrant l'UE. L'UE reste en mode inactif jusqu'à ce qu'il transmette une requête d'établissement de connexion RRC. A ce moment, il passe en mode connecté.

Dans l'état Cell_DCH, un canal physique dédié (DCH et éventuellement DSCH) est alloué à l'UE et l'UE est connu par son SRNC.

Dans l'état Cell_FACH, les canaux communs RACH et FACH sont utilisés pour transmettre des messages de signalisation et des données utilisateur. L'UE est également capable d'écouter le canal de diffusion (BCH) pour obtenir des informations système et le canal CPCH quand l'UTRAN demande une resélection de cellule.

Dans l'état Cell_PCH, l'UE est toujours connu par son SRNC mais n'est joignable que via le canal d'appel (PCH). Si l'UE effectue une resélection de cellule, l'état de l'UE passe à Cell_FACH pour exécuter la procédure « Cell Update » après quoi, l'UE repasse en état Cell_PCH si aucune activité n'est survenue durant la procédure.

L'état URA_PCH est similaire à l'état Cell_PCH excepté que l'UE n'exécute pas la procédure « Cell Update » après chaque resélection de cellule mais lit son identité URA (UTRAN Registration Area). La procédure « URA Update » est similaire à « Cell Update ».

L'UE quitte le mode connecté et retourne en mode inactif lorsque la connexion RRC est libérée ou lors d'un problème de connexion RRC. [3]

Couche physique

Dans l'UTRAN, les données générées par les couches supérieures sont acheminées par l'air avec des canaux de transport, qui sont mappés au niveau physique vers plusieurs canaux physiques. La couche physique est requise pour supporter les canaux de transport à débit variable assurant les services de bande passante à la demande et le multiplexage de services sur une connexion.
Les canaux de transport sont divisés en canaux dédiés et canaux communs. La principale différence entre ces 2 types de canaux est qu'un canal commun est une ressource divisée entre tous ou un groupe d'utilisateurs d'une cellule alors qu'un canal dédié est réservé pour un seul utilisateur.

Canal dédié de transport

Le canal dédié de transport (DCH) peut transporter des données utilisateur ou des informations de contrôle. Il est caractérisé par le contrôle de puissance rapide et le changement de débit trame par trame, chaque trame durant 10ms. DCH supporte le soft handover.

Canal commun de transport

Les canaux communs de transport ne supportent pas le soft handover mais certains supportent le contrôle de puissance rapide. Il existe 7 types de canal commun dans la Release 5 3GPP.

Broadcast Channel (BCH)

Le canal BCH est utilisé pour transmettre des informations spécifiques à l'UTRAN ou à une cellule. Il véhicule notamment les codes d'accès aléatoires et les slots d'accès d'une cellule.

Forward Access Channel (FACH)

Le canal FACH est un canal descendant (downlink) transportant des informations de contrôle vers les terminaux localisés dans une cellule.

Paging Channel (PCH)

Le canal PCH est un canal descendant transportant des données relatives à la procédure d'appel d'un terminal. Lors de l'appel d'un terminal, le réseau transmet le message d'appel à ce terminal via le PCH des cellules appartenant à la « location area » du terminal.

Random Access Channel (RACH)

Le canal RACH est un canal montant (uplink) destiné au transport d'informations de contrôle depuis le terminal comme les requetes d'établissement de connexion.

Common Packet Channel (CPCH)

Le canal CPCH est une extension du canal RACH pour transporter des données paquet utilisateur dans le sens montant.

Downlink Shared Channel (DSCH)

Le canal DSCH est un canal de transport descendant destiné au transport de données dédiées à un utilisateur et/ou d'informations de contrôle. Il est similaire au canal FACH excepté qu'il est toujours associé à un canal DCH descendant.

High Speed Downlink Shared Channel (HS-DSCH)

Le canal HS-DSCH, basé sur le canal DSCH, permet le transport de données utilisateur à haut débit grâce à la technique HSDPA.

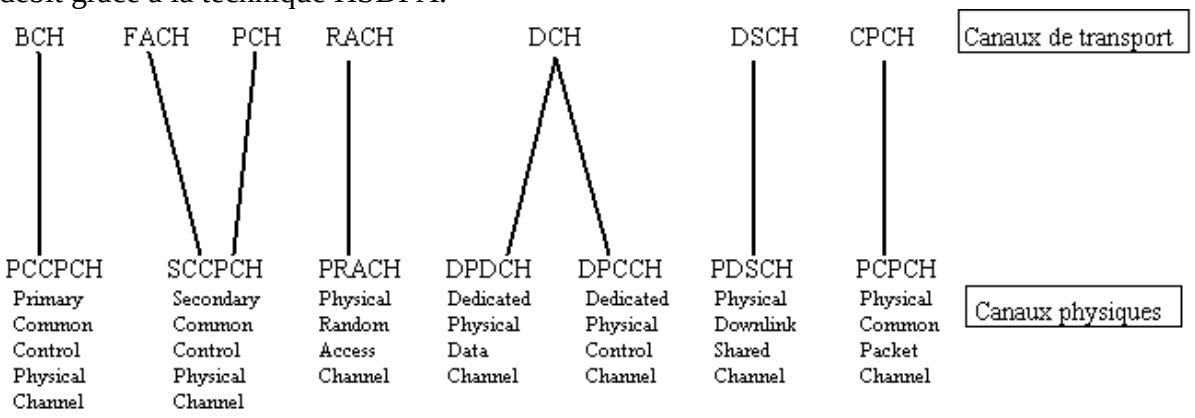

Figure 36 – Mapping canaux de transport et physiques [3]

On remarque que DCH est mappé vers 2 canaux physiques, DPDCH transportant l'information de la couche supérieure - entre autres les données utilisateur, DPCCH transportant l'information de contrôle de couche physique.

En plus des canaux physiques en correspondance directe avec des canaux de transport, il existe des canaux spécifiques au transport d'information relative aux procédures de la couche physique. Ces canaux ne sont pas visibles depuis les couches supérieures. [3]

WCDMA

Comme mentionné, les données sont transmises par l'air et plus exactement par l'interface air WCDMA (Wideband Code Division Multiple Access).

Le système CDMA emploie des codes d'étalement uniques basés sur la technique OVSF (Orthogonal Variable Spreading Factor) pour séparer les transmissions d'une source c-a-d séparer les canaux physiques uplink d'un terminal dans une cellule et séparer les canaux physiques downlink d'un NodeB dans une cellule [47].

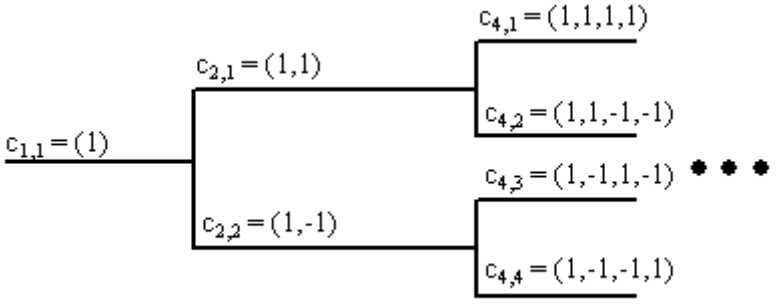

Figure 37 – Arbre de codes d'étalement [47]

Les bits de données à transmettre sont multipliés par le code d'étalement. Le nombre de « chips » (bits pseudo-alétoires) du code d'étalement constitue le facteur d'étalement. Un canal physique peut utiliser un certain code dans l'arbre si aucun autre canal physique n'emploie un code se situant sur la branche sous-jacente. De la même façon, un code à facteur d'étalement plus petit appartenant au chemin vers la racine de l'arbre ne peut être attribué. Ceci assure le maintien de l'orthogonalité entre différents codes.

L'interface WCDMA est caractérisée par un chiprate de 3.84Mcps et une longueur de trame de 10ms. Par l'usage de codes OVSF, le débit des canaux physique peut être ajusté toutes les 10ms en changeant le code d'étalement [3].

Associé au procédé d'étalement, un code de brouillage (scrambling) est appliqué. Il permet différencier les sources (terminaux et cellules) les unes des autres [3].
Les procédés d'étalement et de brouillage sont associés de cette manière :

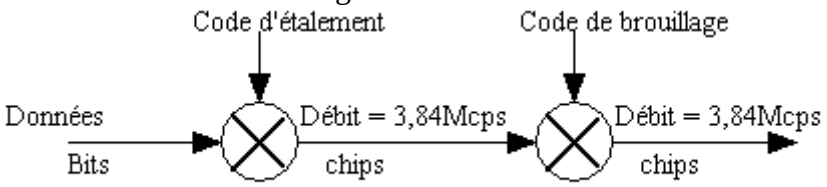

Figure 38 – Relation entre étalement et brouillage [47]

Annexe 802.11

Liaison radio

802.11 définit 3 couches physiques : 2 couches radio à spectre étalé et une couche infrarouge. La couche infrarouge ne sera pas évoquée vu qu'elle n'est que peu utilisée.
802.11a décrit une couche physique basée sur le multiplexage de fréquences orthogonales (OFDM).
802.11b décrit une couche physique basée sur DSSS : HR/DSSS.
802.11g définit une couche physique compatible avec 802.11b et OFDM opérant dans la bande de fréquence 2.4GHz ISM. [8]

Caractéristiques communes des couches radio

Les couches FHSS, DSSS et HR/DSSS emploient la bande de fréquence 2.4-GHz ISM disponible mondialement pour un usage sans licence.
La technologie de spectre étalé est à la base des bandes de fréquence ISM pour la transmission de données. Les communications radio traditionnelles tendent à restreindre le signal sur une bande aussi étroite que possible. Le spectre étalé fonctionne en utilisant des fonctions mathématiques afin de diffuser la puissance du signal sur une large gamme de fréquences. Quand le récepteur effectue l'opération inverse, le signal étalé est reconstitué comme un signal à bande étroite. De cette façon, le bruit généré par des signaux à bande étroite est localisé en fréquence et lors de l'opération inverse, le signal reconstitué n'est pas perturbé par ce bruit. L'usage de la technologie de spectre étalé est une obligation imposée par le régulateur (ETSI en Europe) pour les transmissions radio sans licence. [8]

Couche physique radio FHSS

La couche FHSS (frequency hopping spread spectrum) repose sur le principe de saut de fréquence selon une combinaison prédéterminée pseudo-aléatoire, en transmettant une courte rafale de données sur chaque sous-canal.
La bande de fréquences 2.402-2.479GHz est découpée en 77 sous-canaux de 1MHz. La transmission s'effectue en émettant successivement sur un sous-canal pendant une durée de 390ms puis sur un autre. La durée du saut d'un sous-canal à un autre est de 224µs au minimum. L'ordre d'utilisation des sous-canaux, déterminé par des fonctions mathématiques, constitue une combinaison de sauts. Les combinaisons de sauts sont orthogonales entre elles afin de minimiser le risque d'emploi simultané d'un sous-canal par plusieurs combinaisons.

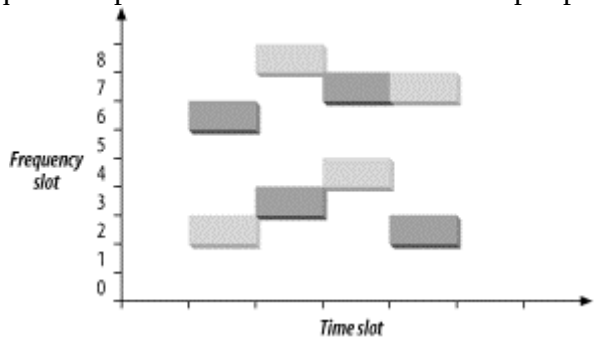

Figure 39 – 2 combinaisons de sauts orthogonales (gris clair & gris foncé) [8]

La couche FHSS utilise le codage GFSK pour la transmission radio. GFSK code les données par une série de changement de fréquence dans la porteuse. On distingue la technique 2GFSK et 4GFSK. 2GFSK emploie 2 fréquences différentes. Pour transmettre un 1, la fréquence de la porteuse est augmentée d'une certaine déviation. Un 0 est codé en diminuant la fréquence par

la même déviation. 4GFSK, basé sur le même principe de 2GFSK, emploie 4 fréquences différentes pour coder les symboles (00, 01, 10, et 11), chaque symbole correspondant à une fréquence. Théoriquement, la méthode GFSK peut être étendue afin d'émettre davantage de bits/s en recourant à des symboles de plusieurs bits. Pratiquement, la difficulté de distinguer de petits changements de fréquence par les composants radio limite le débit à 2Mbps. [8]

Couche physique radio DSSS

La couche DSSS (direct sequence spread spectrum) a pour principe l'étalement de la puissance sur une large bande de fréquence au moyen de fonctions de codage.
La bande de fréquences 2.412-2.472GHz est divisée en 13 canaux de 5 MHz.
La transmission DSSS est une technique alternative de l'étalement de spectre qui permet la transmission d'un signal sur une bande de fréquence plus large. Elle s'effectue sur un canal de 22 MHz. On observe donc un recouvrement partiel des canaux adjacents et seuls 3 canaux sur les 14 sont entièrement isolés les uns des autres. A la réception, le signal est recorrélé c-a-d que le récepteur inverse le processus d'étalement du canal de 22MHz vers un canal de 5MHz.

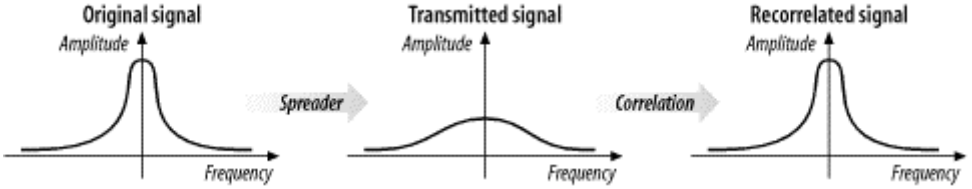

Figure 40 – Etalement DS de 5MHz vers 22MHz et corrélation [8]

Le bruit, caractérisé par des impulsions étroites, n'altère qu'une bande étroite du signal transmis sur le canal de 22MHz. Dès lors, la fonction de corrélation disperse le bruit sur le canal de 5MHz à la réception, altérant peu le signal transmis.

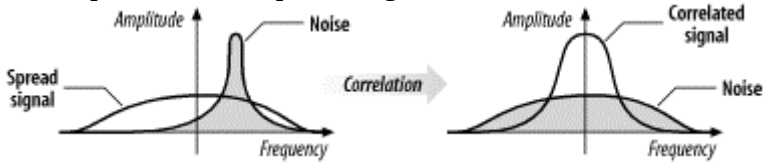

Figure 41 – Dispersion du bruit par corrélation [8]

La modulation DS consiste à appliquer une séquence de « chips » au flux de données à transmettre. Un chip est un digit binaire utilisé par le processus d'étalement. Pratiquement, il n'y a aucune différence entre chip et bit. Les bits sont des données de plus haut niveau alors que les chips sont des nombres binaires utilisés dans le processus de codage.
Pour générer les flux de chips, 802.11 DSSS adopte une séquence Barker de 11bits (10110111000). Chaque bit est codé en additionnant modulo-2 cette séquence Barker. Un bit 1 sera donc représenté par la séquence 01001000111 et 0 par 10110111000. La transmission de flux de chips s'effectue à un débit de 11Mcps.
Les flux de chips générés sont codés pour la transmission radio par le codage PSK (phase shift keying). Cette technique code les données par changement de phase du signal transmis. DBPSK (differential binary phase shift keying) consiste à utiliser 2 ondes porteuses dont une est l'onde de référence utilisée pour coder un 0 et l'autre déphasée de 180° utilisée pour coder un 1. DBPSK permet un débit de 1Mbps. DQPSK (differential quadrature phase shift keying), utilisant 4 porteuses déphasées de 90° les unes par rapport aux autres permet un débit de 2Mbps. [8]

Couche physique radio HR/DSSS

La couche HR/DSSS (high rate / direct sequence spread spectrum), définie par 802.11b, est basée sur la couche DSSS. Elle emploie les mêmes canaux que DSSS avec une technique de modulation différente. La transmission peut s'effectuer à un débit de 11Mbps.
Pour ce faire, les flux de chips sont dérivés partiellement des données à transmettre, la séquence Barker n'étant plus utilisée.
Au lieu d'utiliser le codage PSK, une méthode de codage alternative est employée : CCK. Le codage CCK (complementary code keying) consiste à diviser le flux de chips en séquences de 8chips. Chaque séquence encode 4bits (pour un débit de 5.5Mbps) ou 8bits de données (pour un débit de 11Mbps). [8]

Couche physique OFDM

Le principe d'OFDM (orthogonal frequency division multiplexing) est de diviser un canal disponible en plusieurs sous-canaux et de transmettre une partie du signal codé sur chaque sous-canal en parallèle. Les sous-canaux opèrent à des fréquences indépendantes (orthogonales) les unes des autres.
802.11a offre un vaste choix de techniques de modulation et de codage.
Les canaux spécifiés par 802.11a, larges de 20MHz, sont situés dans la bande de fréquences de 5.15-5.35GHz et 5.725-5.825GHz. 802.11a permet d'atteindre un débit de 54Mbps par canal. [8]