UNIVERSITÉ DE MONTRÉAL

ÉVOLUTION DU PROTOCOLE GTP DANS UN CONTEXTE DE VPN ET DE MOBILITÉ-IP

YVES LEMIEUX
DÉPARTEMENT DE GÉNIE INFORMATIQUE
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES
(GÉNIE INFORMATIQUE)

AVRIL 2005

Your file Votre référence
ISBN: 978-0-494-48929-1
Our file Notre référence
ISBN: 978-0-494-48929-1

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé :

ÉVOLUTION DU PROTOCOLE GTP DANS UN CONTEXTE DE VPN ET DE MOBILITÉ-IP

présenté par : YVES LEMIEUX

en vue de l'obtention du diplôme de : Maîtrise ès sciences appliquées

a été dûment accepté par le jury d'examen constitué de :

MME BOUCHENEB Hanifa, Doctorat, présidente

M. PIERRE Samuel, Ph.D., membre et directeur de recherche

M. QUINTERO Alejandro, Doct., membre

REMERCIEMENTS

J'aimerais prendre l'opportunité de remercier Dr Samuel Pierre, mon directeur de recherche pour son soutien constant lors de la progression de cette recherche et de la rédaction du mémoire de maîtrise.

De plus, je remercie Ericsson Canada Inc. d'avoir supporté ce projet en collaboration avec le Laboratoire de Recherche en Réseautique et en Informatique Mobile (LARIM) de l'École Polytechnique de Montréal dans le cadre de la Chaire Polytechnique/CRSNG-ERICSSON.

Je remercie également M. Denis Monette, mon superviseur immédiat à Ericsson pour m'avoir donné l'occasion de poursuivre des études de 2^{e} cycle, à l'École Polytechnique de Montréal et M. Laurent Marchand, un collègue de travail, pour ses critiques constructives concernant l'évaluation de nouveaux mécanismes améliorant la performance de GTP. De plus, il serait opportun de souligner le support moral de M. Pierre Boucher, directeur du Département de Recherche et Innovation, à Ericsson.

En tout dernier lieu, je remercie ma femme Tracy et mes enfants Marie-Claire, Philippe et Victoria pour m'avoir accorder le temps voulu, pour mener à bien cette Maîtrise de Recherche en Télécommunications.

RÉSUMÉ

Les systèmes UMTS (Universal Mobile Telecommunications System) de 3e Génération satisfaisant les besoins technologiques des communications mobiles, ciblent la convergence de la téléphonie basée sur un paradigme IP, et aussi toute une panoplie de nouveaux services de façon à générer de nouvelles opportunités. En effet, la définition d'IMS (Internet Multimedia Sub-System) contenue à même les spécifications 3GPP, confirme cette tendance. UTMS vise aussi à satisfaire les limitations des réseaux mobiles d'opérateurs, en lui ajoutant une efficacité spectrale améliorée et un coût de transport réduit et ce, par l'intermédiaire de fondations tout-IP à partir de la Version 5 et des versions subséquentes.

Bien que la technologie UMTS existante soit adéquate pour les types de trafic de 2e et de 2.5e Générations, le protocole GTP (GPRS Tunelling Protocol) utilisé dans le but de supporter la mobilité macro avec une QdS soutenue, a été repris à partir de l'architecture GPRS (General Packet Radio Service) de base, et ne satisfait pas complètement les besoins des applications futures de l'Internet. Ceci est dû au fait qu'il dépend déjà de IP pour sa couche inférieure, alors que GTP lui-même encapsule les paquets IP en provenance des couches supérieures. Il devient alors primordial de considérer un sentier évolutif pour le protocole GTP, de manière à permettre au réseau d'accès sans fil tout-IP, d'être plus efficace, tout en étant basé sur MPLS (Multi Protocol Label Switching). Ceci permettra alors d'éliminer les limitations courantes existant sous GTP, principalement durant les relèves radio inter-SGSN.

L'objectif de ce mémoire demeure principalement de proposer une solution et de fournir une preuve de concept satisfaisant le sentier évolutif de GTP, basé sur MPLS. Un algorithme pour l'établissement des tunnels de GTP évolué se sert d'une multitude de LSP prédéfinis, à être utilisés lorsque des relèves inter-SGSN sont déclenchées par des messages RAU (Routing Area Update). De plus, nous proposons cet algorithme avec une infrastructure qui consiste en un arrangement VPN (Virtual Private Network) et support

de mobilité MIPv6, de manière à satisfaire le triplet de QdS-Sécurité-Mobilité durant le fonctionnement d'accès UMTS sans fil.

Dans cette optique, cette recherche se compose de quatre phases distinctes. Durant la première phase, nous présenterons des mesures prises lors des demandes RAU à l'interface G_b situé entre le BSC/RNC et le SGSN (Service GPRS Support Node). Ces mesures serviront de point de référence aux améliorations apportées. Dans une deuxième phase, un modèle de commutation par étiquettes sera fourni en utilisant l'outil de simulation *Modeler™ V10.5* d'Opnet. Le but sera d'obtenir une preuve de concept de l'établissement de tunnels statiques et dynamiques faisant partie du sentier évolutif au protocole GTP, tel que connu aujourd'hui. Dans la troisième phase, nous validerons la solution de GTP évolué en utilisant l'outil de vérificateur de modèle appelé *SPIN*. Ceci nous donnera une indication, à savoir si nous devons déplacer ou non, l'établissement des tunnels plus tôt dans le temps, en parallèle avec le déclenchement RAU. Finalement, durant la dernière phase, nous présenterons une approche analytique concernant le dimensionnement d'un réseau d'accès sans fil utilisant une multitude de flots de voix et de vidéo conférence, et ce, basé sur un routeur typique agissant comme GGSN (Gateway GPRS Support Node). Cette approche demeure basée sur un concept de bande passante efficace, qui est aussi intégré à l'algorithme mentionné auparavant servant à l'allocation de bande passante des LSP (Label Switched Paths) prédéfinis.

Les résultats obtenus confirment définitivement l'utilisation de chemins basés sur des LSP prédéfinis, démontrant des temps d'établissement améliorés. Notre objectif demeure depuis le début, d'obtenir des temps de relèves inter-SGSN inférieurs à ½ seconde. Nous sommes d'emblée capables de commuter entre deux chemins, avec un temps réponse moindre que 0.025 seconde (typiquement de l'ordre de 1 msec) et ce, dans un mode d'opération DBBM (Dynamic Break Before Make), et finalement, un temps réponse proche de 0 seconde, dans un mode d'opération SMBB (Static Make Before Break). Il demeure toutefois recommandé d'utiliser la solution DBBM pour les relèves demandant une évolutivité accrue, comme par exemple pour les trafics de voix, et d'utiliser le SMBB pour les trafics hautement prioritaires tels ceux des appels 911. La

validation basée sur *SPIN* nous a amené à considérer des spécifications d'horloges et de mémoire tampon, et a également identifié le besoin de commencer l'établissement de tunnels GTP évolué plus tôt, c'est-à-dire en parallèle avec le déclenchement RAU, de manière à éviter les états conflictuels d'un message context_ack arrivant avant l'établissement complet d'un LSP. Finalement, le dimensionnement analytique a démontré que les ressources requises pour une population de 1.4 millions d'usagers mobiles, dans une région métropolitaine typique, demandent environ la moitié de la capacité d'un routeur GGSN utilisant un châssis à débit binaire de 20 Gbps. Finalement, un cadre future est présenté, dans le but de supporter les fonctionnalités VPN et MIPv6 qui assureront qu'un réseau d'accès sans fil puisse favoriser un équilibre opérationnel de QdS-Sécurité-Mobilité.

ABSTRACT

The third Generation UMTS (Universal Mobile Telecommunications System) for mobile communication technology, targets the convergence of telephony, based on an IP paradigm, and also a suite of new services in order to generate new opportunities. In fact, the IMS (Internet Multimedia Sub-System) definition under 3GPP confirms this trend. Moreover, UMTS aims to satisfy the Mobile Network Operators' limitations with an improved spectral efficiency and a lower transport cost, by using the all-IP foundations promoted by Releases 5 and up.

Although the existing UMTS technology is adequate for the traffic types of the 2G and 2.5G, the GTP (GPRS Tunelling Protocol) sub-part being used to support macro mobility with a sustained QoS, has been originally derived from the GPRS (General Packet Radio Service) architecture, and does not fulfill completely the needs of all future Web-Applications. This is due to the fact that it already depends on IP as lower layer, while GTP itself encapsulates IP packets from higher layers. It would thus become crucial to consider an evolution path for GTP, in order to allow a Wireless all-IP Network Access, to be more efficient when based on MPLS (Multi Protocol Label Switching). This would therefore alleviate the current limitations experienced by the GTP protocol mainly during inter-SGSN Hand-Off operations.

The objective of this thesis is mainly to propose and provide a proof-of-concept for an evolution path for GTP, based on MPLS. An algorithm for the establishment of GTP-evolved tunnels promotes the use of a plurality of pre-defined LSP, to be used when needed especially when inter-SGSN Hand-Offs are triggered by an RAU. More-over, the intent is to support this new algorithm with an infrastructure that consists of a VPN (Virtual Private Network) arrangement and MIPv6 mobility support, in order to satisfy the triplet of QoS-Security-Mobility during the operation of UMTS Wireless Accesses.

For that purpose, this research is composed of four phases. During the first phase, we will present measurements taken during RAU (Routing Area Updates) at the G_b interface between the BSC/RNC and the SGSN (Service GPRS Support Node). These

measurements will serve as benchmark for Hand-Offs. In the second phase, a label-switching model is provided on the *Modeler*^{TM} *V10.5* simulation tool from Opnet with the aim of getting a proof-of-concept for static and dynamic pre-defined tunnel set-ups as an evolution path to the classical GTP. In the third phase, we will validate the GTP evolved solution by using the *SPIN Model-Checker* tool. We may have to consider for that very specific reason, the parallel establishment of tunnels at the time of RAU triggers. Finally, in the last phase, we present an analytical approach for the dimensioning of a voice and videoconference based Wireless-Access Network using a typical router supporting the GGSN (Gateway GPRS Support Node) functions. This approach is based on a concept of effective-bandwidth, which is also integrated in the bandwidth allocation algorithm used for the definition of pre-defined LSP (Label Switched Paths).

The results obtained definitely re-enforce the idea of using pre-defined LSP paths with a better set-up time. Our objective was to perform faster than ½ second during an inter SGSN Hand-Off. We are able to switch between paths within less than 0.025 second (more like 1 msec) in a DBBM (Dynamic Break Before Make) mode of operation, and close to 0.0 second in a SMBB (Static Make Before Break) mode of operation. It is however recommended to use DBBM for the Hand-Offs requiring large scalability such as for voice traffic and the SMBB mode of operation for high priority traffic such as 911 calls. The *SPIN* validation provided us with the requirements for timers and buffers, and also with the need to start the tunnel establishment in parallel with the RAU trigger, in order to avoid conflicting operational states of the context_ack happening before the complete LSP establishment. Finally, the analytical dimensioning demonstrated that the required resources for a population of 1.4 million users in a typical metropolitan area, requires about half of the capacity of a 20 Gbps backplane based GGSN router. Finally, some future framework is presented to support VPN and MIPv6 in a QoS-Security-Mobility capable Wireless-Access-Network.

TABLE DES MATIÈRES

LISTE DES TABLEAUX

LISTE DES FIGURES

SIGLES ET ABRÉVIATIONS

AAA	Authentication, Authorisation and Accounting
APN	Access Point Name
AR	Access Router
ATM	Asynchronous Transfer Mode
BG	Border Gateway (GPRS)
BSS	Base SubSystem (GSM access network)
CDR	Call Detail Record
CE	Customer Equipment
CGF	Charging Gateway Functionality
CH	Correspondent Host (même chose que "Correspondent Node")
CN	Core Network
CoA	Care Of Address
CR-LDP	Constrained Routing – Label Distribution Protocol
DS	Differentiated Services
DBBM	Dynamic Break Before Make
FA	Foreign Agent
FACoA	Foreign Agent Care-Of Address
GGSN	Gateway GPRS Support Node
GPRS	General Packet Radio Service
GTP	GPRS Tunelling Protocol
HA	Home Agent
HLR	Home Location Register
HO	Hand-Off
IETF	Internet Engineering Task Force
IGSN	Internet GPRS Support Node
IWU	Inter-Working Unit
LA	Local Area

LAC	Location Area Code
LDP	Label Distribution Protocol
LER	Label Edge Router
LLC	Logical Link Control
LSP	Label Switched Path
LSR	Label Switch Router
MAP	Mobility Anchor Point
ME	Mobile Equipment
MIP	Mobile IP
MM	Mobility Management
MPLS	Multi-Protocol Label Switching
MT	Mobile Termination
NAI	Network Access Identifier
NAS	Network Access Server
N-PDU	Network layer PDU (utilisé dans GPRS pour identifier un PDU transporté par un contenant GTP)
OSPF	Open Shortest Path First
P	Provider
PE	Provider Edge
PHB	Per Hop Behavior
PLMN	Public Land Mobile Network
P-TMSI	Packet TMSI
QoS	Quality of Service
RA	Routing Area
RAI	Routing Area Identifier
RAN	Radio Access Network
RAU	Routing Area Update
RFC	Request For Comments
RNC	Radio Network Controller

RNTI	Radio Network Temporary Identifier
RSVP	Resource ReSerVation Protocol
SGSN	Service GPRS Support Node
SBBM	Static Break Before Make
SMBB	Static Make Before Break
SRNC	Serving RNC
TE	Terminal Equipment
TE-CMS	Traffic Engineering - Configuration Management System
TLLI	Temporary Logical Link Identifier
TMSI	Temporary Mobile Subscriber Identifier
UDP	User Datagram Protocol
UE	User Equipment
UMTS	Universal Mobile Telecommunications System
URA	UTRAN Registration Area
USIM	User Services Identity Module
UTRAN	UMTS Terrestrial Radio Access Network
VHE	Virtual Home Environment
VLR	Visitor Location on Register
VPN	Virtual Private Network
VRF	VPN Routing and Forwarding

LISTE DES ANNEXES

CHAPITRE I
INTRODUCTION

Les réseaux sans fil contemporains tendent de plus en plus vers l'accès à des services Internet multimédias. Ces services requièrent davantage de ressources adaptées aux besoins de temps réel, de regroupement d'individus visant un but commun, et de mobilité accrue. UMTS a cherché à construire et à étendre le potentiel des systèmes mobiles sans fil, et aussi de technologie satellite, tout en procurant une capacité accrue, une performance de transfert de données améliorée et un spectre plus large de services utilisant un accès radio plutôt progressif, tout en proposant de nouveaux mécanismes sur le réseau cœur. L'UMTS de 3^{e} génération (3G) permettra un débit binaire de 384 Kbps pour les applications en mouvement, et de 2 Mbps pour les applications plus sédentaires mais portables. Ces applications peuvent facilement passer du domaine de la voix, à celui du monde du commerce électronique, tout en touchant le trafic contraignant de la télé-médecine, ou de la vidéo-conférence.

L'UMTS de 3G, mise sur pieds pour la technologie des communications mobiles, cible donc la convergence de la téléphonie, du paradigme IP, et aussi une panoplie de nouveaux services dans le but de générer de nouvelles opportunités. Il vise également à satisfaire les contraintes d'opérateurs de réseaux mobiles avec une meilleure efficacité spectrale et un coût de transport moins élevé en utilisant des technologie telles que ATM pour la Version 99 (Release 99) et le tout-IP pour la Version 5 (Release 5). Bien qu'adéquate pour les types de trafic des 2G et 2.5G, la technologie sous-jacente GTP permet de supporter la mobilité macro avec une QdS soutenue; elle a été dérivée de GPRS et ne répond pas à 100% aux besoins applicatifs du futur. Ceci est dû au fait qu'il utilise déjà IP comme couche sous-jacente nous forçant ainsi à encapsuler un paquet IP dans un paquet GTP, qui lui-même sera encapsulé dans un paquet IP. Somme toute, le protocole GTP induit un certain niveau de latence relativement élevé durant la relève de communication d'une station sans fil à une autre. Il serait donc opportun de considérer un sentier évolutif de GTP qui nous permettrait d'avoir un réseau d'accès tout-IP plus

efficace basé sur MPLS, ceci dans le but de pallier les limitations existantes du protocole GTP. C'est donc l'objet de ce mémoire. Dans ce chapitre d'introduction, nous présentons d'abord quelques définitions et concepts de base, résumons par la suite les éléments de la problématique, précisons ensuite les objectifs de recherche avant d'esquisser finalement le plan du mémoire.

1.1 Définitions et concepts de base

Avant même d'esquisser des solutions possibles d'évolution de GTP, il demeure important de faire une mise en contexte de la technologie considérée, soit celle de GTP et de ces diverses composantes. Celles-ci consistent essentiellement en un certain nombre de nœuds d'accès UMTS et l'infrastructure de support de la QdS.

1.1.1 Introduction au protocole GTP

L'acronyme GTP signifie "GPRS Tunneling Protocol" où GPRS représente "General Packet Radio Service". GTP définit ainsi le protocole entre les nœuds GSN ou encore "GPRS Support Nodes" à l'intérieur du réseau cœur UMTS. Les nœuds GSN sont plus précisément ceux de SGSN et de GGSN. Ceci inclut deux parties, soient la signalisation de contrôle et les procédures de transfert de données d'usager. Ces deux parties sont identifiées par GTP-C pour Contrôle et GTP-U pour Usager. Le protocole GTP se rattache aux points d'interface suivants:

- l'interface G_n entre les GSNs, dans un PLMN (Public Land Mobile Network);
- l'interface G_p entre GSNs, dans différents PLMNs.

Pour mieux comprendre la position de ces interfaces, il est bon de se référer à la Figure 1.1.

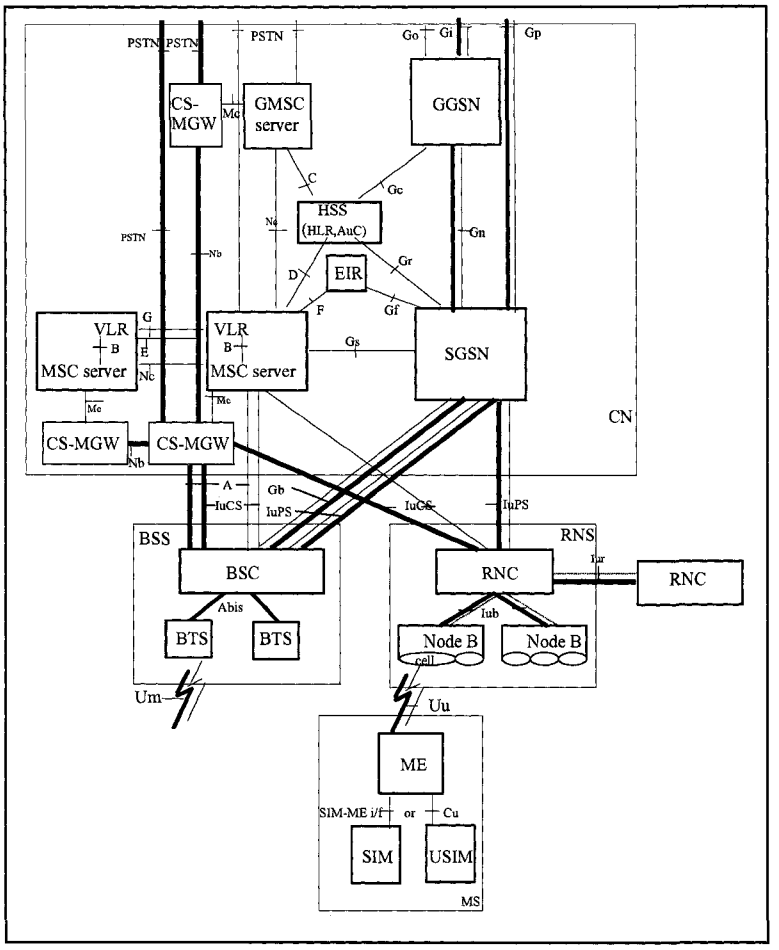

Figure 1.1 Spécification d'architecture du réseau UMTS selon 3GPP

Le but du protocole GTP est d'établir des tunnels pour transporter des paquets en provenance de multi-protocoles sur-jacents, tels que IP, trames primitives, etc., pour les faire passer au travers du réseau cœur UMTS, entre les nœuds SGSN et GGSN. Dans le plan de signalisation, GTP-C spécifie le contrôle de tunnel et le protocole de gestion permettant au SGSN de procurer des services basés sur GPRS, pour satisfaire aux besoins des stations mobiles d'usager communément appelées UE (User Equipment), selon la terminologie d'UMTS. En conclusion, la signalisation GTP-C permet donc de créer, modifier et retirer l'établissement de tunnels de paquets de données UMTS basés sur GPRS. Dans ce but, le protocole de couche 4, UDP (User Datagram Protocol), permet de transférer les messages de signalisation entre les nœuds GSN.

Pour ce qui est de l'entité protocolaire GTP-U, elle utilise le mécanisme de tunnelage pour transporter les paquets de données d'usager. GTP-U est donc le protocole du plan d'usager situé entre le RNC (Radio Network Controller), à l'intérieur de l'UTRAN (UMTS Terrestrial Radio Access Network) du système d'accès UMTS, et le nœud SGSN, en passant par l'interface I_u. À ce point d'interface, cependant, le plan de contrôle correspondant est celui de RANAP (Radio Access Network Application Protocol). Il va de soi que le modèle de réseau GPRS sert de fondation pour construire le système d'accès sans fil UMTS.

1.1.2 Garantie de QdS dans les réseaux actuels de GPRS

GPRS promet une capacité haute vitesse, toujours active (comparativement au connexions modem qui sont normalement établies lorsque requises), et un débit binaire élevé pour les usagers mobiles. Pour s'assurer que GPRS puisse procurer une QdS acceptable, les réseaux UMTS doivent être optimisés en conséquence. GPRS étend ses services courants de transfert de données en utilisant des classes de services appropriées. Suivant l'évolution des volumes de trafics UMTS et des applications en temps réel (voix, vidéo sur demande, etc.), le besoin d'assurer une QdS accrue devient de plus en plus important. Cette assurance requiert la spécification de paramètres de qualité pour les services de données à paquets commutés.

GPRS définit différents profils ou classes de QdS qui permettent au réseau d'établir une priorité des trafics supportés, de telle façon que l'usager puisse décider du niveau de QdS désiré. Le fait de garantir une QdS est plutôt difficile en GPRS ou en UMTS R99 (première version d'UMTS qui s'apparente grandement à celle du GPRS dont elle est dérivée, à part l'addition d'interface radio à plus grand débit telle que W-CDMA), puisque l'assignation de ressources pour chaque usager en bout de ligne dépend de la disponibilité des ressources radios dans un réseau mobile sans fil. La QdS consiste en quatre classes qui sont:

1. la classe de précédence: les trafics concernés sont associés à des priorités 1, 2 ou 3, 1 représentant la plus grande priorité ;

2. la classe de délai: cette classe garantit un délai moyen et un percentile de 95% ;
3. la classe de survavibilité: elle dépend de l'habilité des applications de soutenir des paquets corrompus ou dupliqués ;
4. la classe de débit traversant (throughput): ce débit de crête et moyen s'applique à des intervalles de courtes et de longues durées respectivement, pour le traitement de trafic hautement variable.

Les étapes d'établissement de flots avec une QdS adéquate sont décrites ci-après. Une station mobile requiert la QdS désirée durant une phase d'activation de contexte PDP. Le SGSN réplique avec la QdS que le réseau peut procurer. Bien que l'usager GPRS veuille obtenir une QdS donnée, plusieurs facteurs affectent la performance actuelle du réseau d'accès. Par exemple, GPRS dépend grandement d'une bonne qualité radio. GPRS est une technologie de commutation de paquets pour les services Internet traditionnels, qui a une tolérance réduite face aux paquets perdus ou corrompus. Même en utilisant la première classe de QdS, la perte de paquets au niveau radio peut forcer plusieurs retransmissions, ce qui se traduit par des débits de transfert, des latences et des gigues dégradées.

Finalement, puisque GPRS est un protocole de paquets de données, et que les terminaux d'usagers sont continuellement en déplacement, ceci force les paquets de données à subir des changements de latence et de haut niveaux de gigues. Ces variations seront corrigées par de nouvelles classes de services offertes par MPLS. Cette amélioration requiert plus de recherche et deviendra disponible sous l'égide de la Version 5 d'UMTS (UMTS Release 5). Cette identification de concept basée sur les accès UMTS nous amène maintenant aux éléments de la problématique.

1.2 Éléments de la problématique

Le protocole GTP a déjà été décrit brièvement dans les pages antérieures et sert essentiellement à établir des tunnels entre les nœuds du réseau de cœur de l'accès UMTS. Cependant, il est plus ou moins bien adapté en termes de temps réponse lors d'établissement de sessions, à cause du nombre et de la complexité des primitives (messages) qu'il comprend. Il devient donc d'une importance primordiale de favoriser une évolution appropriée du protocole GTP, tout en gardant à l'esprit la perspective des paradigmes de VPN (Virtual Private Network) et de mobilité IPv6. Une évolution de GTP est désirable, avec l'utilisation des ressources offertes par l'établissement de tunnels basés sur les LSPs (Label Switched Path) tel que décrits dans [1]. Pour bien comprendre le besoin d'amélioration du protocole GTP et de la problématique qui en découle, considérons d'abord certains aspects de macro-mobilité.

La spécification de base pour la Mobilité-IP, décrite dans RFC-2002, nous fournit des outils pour la gestion de mobilité macro. Cependant, cette Mobilité-IP n'est pas tout à fait appropriée pour soutenir la micro-mobilité [2]. Un exemple typique de micro-mobilité est celui de la relève d'un terminal mobile qui se déplace entre deux stations de base, couvrant tous deux une petite zone géographique.

Les réseaux GPRS sont divisés en deux parties: le réseau d'accès radio qui est appelé le sous-système des stations de base, et le réseau cœur qui, lui, consiste principalement en deux éléments de réseau qui sont le SGSN et le GGSN. D'un point de vue topologique, le SGSN a une connectivité directe au réseau d'accès radio. D'un autre côté, le GGSN a une connexion directe au réseau de dorsale IP (par exemple Internet et intranet). Le réseau cœur reflète en général des limitations de mobilité macro semblables à celle de la Mobilité-IP (RFC-2002).

À cet effet, voici quelques ressemblances et différences entre GPRS et Mobilité-IP:

- La méthode de tunnelage pour supporter la mobilité macro : Le support de mobilité pour les hôtes IP se trouve dans les deux cas (Mobilité-IP et GPRS) solutionné en appliquant le principe de tunnelage pour supporter l'envoi des

paquets sur les routeurs qui n'ont pas de mécanismes spécifiques de mobilité. Lorsque l'usager se retrouve loin de son accès initial domiciliaire, la Mobilité-IP utilise des protocoles de tunnelage pour cacher l'adresse IP du domicile, du nœud mobile des routeurs intervenant entre le réseau domiciliaire et l'endroit courant du hôte considéré [2]. Tous deux, les agents domiciliaires (Home Agent) et étrangers (Foreign Agents), doivent supporter le tunnelage de datagrammes, en utilisant IP dans une encapsulation IP. Dans GPRS, les données d'usager sont transférées de façon transparente entre la station mobile et le réseau externe de données, et ce, en utilisant le tunnelage offert par GTP. Plus précisément, GTP-U pour l'usager est transporté sur UDP/IP dans le cas de IP sur GPRS.

- Le point-en-bordure (end-point) pour le tunnelage : Sous la Mobilité-IP, le tunnel de l'agent domiciliaire se termine à l'*adresse-au-soin-de* (Care-of-Address) du nœud mobile. Le datagramme originel est enlevé du tunnel et délivré au nœud mobile, à cette adresse. Dans le cas d'une *adresse-au-soin-de* d'un agent étranger, le tunnel est terminé à l'agent étranger. Dans les environnements cellulaires, les ressources radios sont très limitées, ce qui nous force à adapter davantage ces ressources à l'environnement des tunnels, sans dédier des tunnels directement à des stations mobiles UE (User Equipment). Dans le cas de GTP, il est similaire à une *adresse-au-soin-de* d'un agent étranger, puisqu'un tunnel GTP n'est pas terminé directement sur une station mobile mais sur un SGSN. Dans GPRS, un tunnel est identifié par un identificateur de point-en-bordure de tunnel et en plus d'une adresse SGSN/GGSN.
- Connectivité aux réseaux extérieurs : Selon le groupe de travail (working group) de Mobilité-IP de l'IETF, l'agent étranger est un routeur sur un réseau visité par un hôte mobile. Ce routeur procure l'acheminement (routing) au nœud mobile lorsque enregistré. L'agent étranger est capable de dé-tunneler et de livrer les datagrammes au nœud mobile. Ces paquets avaient été tunnelés d'avance par l'agent domiciliaire du nœud mobile. Par contre, dans GPRS, tel que décrit auparavant, le SGSN est aussi capable de tunneler et d'encapsuler les données

d'usager. Ceci veut dire pratiquement que le SGSN n'est pas directement connecté aux réseaux externes (par exemple la dorsale Internet). En effet, les datagrammes d'usager sont toujours tunnelés au GGSN. Le GGSN encapsule toujours dans un tunnel les paquets IP vers le SGSN. Au contraire, dans la Mobilité-IP, l'agent domiciliaire est un routeur sur un réseau domiciliaire du nœud mobile et il n'a pas besoin de tunneler les paquets, qui sont destinés aux nœuds mobiles qui ont des connexions directes de couche 2.

- Les enregistrements et l'établissement des tunnels : Dans la Mobilité-IP, il existe une fonction appelée l'enregistrement qui, lorsque le nœud mobile est hors du milieu domiciliaire, enregistre son *adresse-au-soin-de* avec son agent domiciliaire. Ceci est fait à partir d'échanges de messages "Registration Request" et "Registration Reply" qui sont envoyés avec UDP, en utilisant le numéro de port bien connu 434. Tout dépendant de la méthode d'attachement, le nœud mobile s'enregistra soit directement avec son agent domiciliaire ou à partir de son agent étranger, lequel enverra l'enregistrement à l'agent domiciliaire. L'*adresse-au-soin-de* peut être déterminée par des notifications de l'agent étranger (une *adresse-au-soin-de* de l'agent étranger) ou par des mécanismes dynamiques d'assignation d'adresses IP tel qu'un DHCP (une *adresse-au-soin-de* co-localisée). Selon les spécifications GPRS, il est possible à une station mobile d'être attachée au réseau GPRS même si elle n'a pas un tunnel GTP activé pour le transfert des données. Une fonction similaire à l'enregistrement de Mobilité-IP est présente dans GPRS, et fait partie de la procédure d'activation de la connexion de données. Durant cette procédure, un tunnel GTP est établi entre le SGSN et le GGSN. Cependant, dans le concept de Mobilité-IP (RFC-2002), le moyen d'établir un tunnel n'est pas explicitement décrit.

Suite à cette comparaison entre GPRS et Mobilité-IP, il devient plus clair qu'il existe des opportunités d'amélioration au traitement de la mobilité macro qui, si elle est bien définie dans GPRS tout en étant un peu lourde, est moins bien définie dans Mobilité-IP car trop simplifiée.

1.3 Objectifs de la recherche

L'objectif principal de ce mémoire est de proposer un certain nombre d'améliorations au protocole GTP, notamment en ce qui a trait au traitement de la mobilité macro et au processus de tunnelage. De manière plus spécifique, ce mémoire vise à :

- analyser le protocole GTP dans sa Version 99 afin d'en relever les principales faiblesses pouvant inspirer son évolution ;
- proposer un certain nombre d'améliorations à GTP, en utilisant la modélisation de systèmes avec support VPN et un début de Mobilité-IP ;
- évaluer les améliorations proposées au protocole GTP en regard des systèmes UMTS de spécifications 3GPP-R99.

1.4 Plan du mémoire

Ce mémoire comprend cinq chapitres. Ce premier chapitre d'introduction est suivi par le Chapitre II qui analyse le protocole GTP Version 99 dans le contexte d'une revue de littérature. Le Chapitre III présente un certain nombre d'améliorations à GTP en utilisant la modélisation de systèmes avec support VPN (définition de cadre) et un début de Mobilité-IP. Le Chapitre IV expose quelques détails d'implémentation et les résultats de validation obtenus par simulation. Le Chapitre V, en guise de conclusion, fait une synthèse des travaux réalisés dans ce mémoire en précisant les limitations, et esquisse des voies de recherches futures.

CHAPITRE II
ANALYSE DU PROTOCOLE GTP

Dans ce chapitre, nous analysons le protocole GTP en considérant une série de propositions pertinentes extraites de la littérature. Dans un premier temps, nous examinerons une solution de réseau mobile IP basé sur MPLS. Par la suite, nous passerons en revue la mobilité basée sur les technologies VPN. Nous terminerons le chapitre par un coup d'œil sur la gestion de mobilité tout-IP et son intégration dans un cadre VPN.

2.1 Réseau mobile IP basé sur MPLS

Un MN (Mobile Node équivalent à un UE d'UMTS) peut subir non seulement les effets d'une micro-mobilité (intra-GGSN), mais également ceux d'une macro-mobilité (inter-Routing Areas). Dans [3], il est décrit la manière de construire un réseau Mobilité-IP de grande envergure en utilisant un réseau MPLS comme fondation, un réseau Mobilité-IP de petite envergure pouvant être connecté à d'autres réseaux en passant par un réseau dorsal MPLS. Il y est également proposé une architecture de réseau MPLS hiérarchique pour supporter la Mobilité-IP de grande envergure. Plus spécifiquement, le protocole LDP (Label Distribution Protocol) peut être utilisé pour établir des tunnels LSP (Label Switched Path) entre les agents mobiles (ou encore entre FA et HA) basés sur une technologie IPv4. Un ou plusieurs LSP peuvent remplacer les tunnels IP-encapsulé-dans-IP, à partir d'un réseau MPLS.

Notre recherche se différencie déjà ici du fait que nous n'utiliserons pas LDP comme protocole de distribution d'étiquettes, mais plutôt RSVP-TE pour combler ce besoin et celui d'ingénierie de trafic. De plus, nous tendons davantage vers la technologie d'adressage IPv6 plutôt que celle d'IPv4 à cause d'améliorations marquées concernant le sujet de la mobilité, dont par exemple l'élimination de l'agent FA et aussi l'optimisation de route.

Selon [3] cependant, un réseau MPLS peut fournir le support nécessaire requis en procurant une solution dorsale à haute vitesse d'envoi et de type IP. MPLS doit supporter les Agents HA et FA dans le but de s'assurer que les services de mobilité IPv4 sont offerts [3]. Le routeur en bordure ou encore LER peut satisfaire cette condition, s'il utilise MPLS pour encapsuler les paquets IP jouissant de mobilité IPv4. Ce LER peut donc agir soit comme FA ou encore comme Nœud Correspondant. Il peut aussi fonctionner comme routeur d'adaptation (Gateway) pour se connecter au réseau de Mobilité-IP correspondant. Dans [3], une architecture hiérarchique se compose d'un nœud *Gateway* défini comme FA et également de nœuds FA régionaux, tous faisant partie intégrante d'un réseau MPLS. Pour ce qui est de la partie contrôle de cette proposition, les fonctions du protocole LDP peuvent, à la rigueur, servir à l'établissement de tunnels basés sur LSP entre agents mobiles, en passant par le réseau MPLS. Ces tunnels encapsulant IP dans IP pour Mobilité-IP sont faciles à obtenir par le mécanisme d'emboîtement de LSP offert par MPLS. De plus, lorsqu'un nœud mobile passe de son milieu d'origine à un milieu étranger, le sentier LSP déjà établi peut être étendu sans interruption de service. L'avantage d'utilisation d'un LSP est qu'il permet un chemin direct recalculé pour éviter l'utilisation de connexions cascadées.

Trois scénarios possibles sont décrits [3]:

1. appliquer le protocole Mobilité-IP de base à un réseau MPLS (strict minimum sans optimisation de délai d'établissement de LSP) ;
2. appliquer le protocole RO (Route Optimization) de Mobilité-IP au réseau basé MPLS (sans optimisation de délai d'établissement de LSP) ;
3. appliquer le protocole hiérarchique MPLS de Mobilité-IP.

Dans notre cas, le scénario 3 est celui d'intérêt qui sera considéré dans ce mémoire. Ce scénario est défini à partir de la solution de Choi [3] en considérant un réseau hiérarchique basé sur MPLS et utilisant une Mobilité-IP tel qu'illustré à la Figure 2.1. La méthode décrite par Choi [3] dépend d'un enregistrement régional de Mobilité-IP

et permet à un mobile d'effectuer ces enregistrements locaux en utilisant un nœud d'adaptation agent étranger ou encore "Gateway Foreign Agent", de manière à réduire sensiblement le nombre de messages de signalisation au réseau d'origine. Il en découle ainsi une réduction du délai de signalisation lorsqu'un nœud se déplace entre agents mobiles étrangers et, de ce fait, améliore la performance de ces relèves inter-domaines.

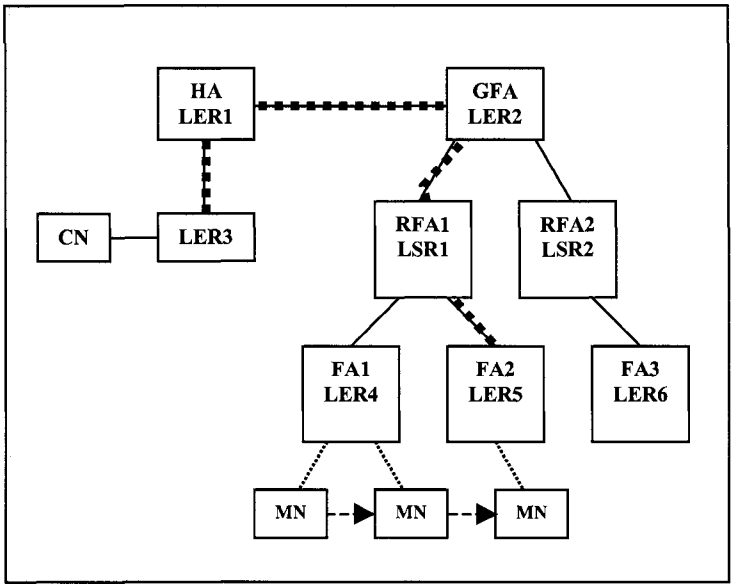

Figure 2.1 Définition d'un réseau d'accès mobilité-IP hiérarchique basé sur MPLS

Lorsqu'un mobile se déplace d'un nœud d'origine à un nœud adjacent sur un réseau établi selon le scénario de Mobilité-Hiérarchique-IP basé sur MPLS, ce mobile effectue un enregistrement régional et un re-établissement partiel de LSP. Dans ce cas, un LSP est établi entre le FA d'ancrage et le nouveau-FA, alors que le LSP déjà existant entre HA et FA demeure toujours. Ceci devrait réduire le délai d'établissement dû à l'intégration de Mobilité-IP et de MPLS.

Selon [4], nous devons utiliser l'encapsulation IP pour envoyer des paquets d'usagers caractérisés par MIPv4. Lorsque le mobile se trouve loin de son point d'origine, ce nœud doit être enregistré, à partir de mécanismes de découverte d'agent appropriés, au FA contenu dans le LER étranger. Cet enregistrement est aussi requis entre le HA et le FA correspondant au travers du réseau MPLS.

Lorsqu'un nœud essaie d'appeler un autre nœud mobile, un chemin d'envoi de données doit être établi entre le LER d'origine et les LERs destination, et ce, avec l'aide du HA. Les opérations provenant du protocole LDP permettent alors au LSP d'étiqueter les paquets IP et de les envoyer au nœud mobile de destination en passant par le réseau MPLS, tout comme n'importe quel réseau MPLS le permettrait.

Cette ébauche [3] demeure limitée aux besoins spécifiques requis par le protocole existant de Mobilité-IPv4. Ces requis sont ceux de découverte d'agents, d'enregistrement, et d'acheminement. Dans le cadre de ce mémoire, nous nous étendrons au réseau d'accès sans fil avec son propre mode d'enregistrement et ses mécanismes d'établissement de session. Bien que le but de notre méthode soit de réduire le temps d'établissement de LSP durant une relève inter-domaine, ce mémoire adopte l'hypothèse de mobilité IPv6 et considère que les LSP établis sont pré-définis à partir d'une entité TE-CMS d'optimisation de ressources.

2.2 Mobilité basée sur les technologies VPN

La Mobilité-IP n'est pas à court de contraintes comme celles d'évolutivité, de sécurité et de QdS. Ces aspects doivent être considérés les uns en combinaison avec les autres pour obtenir une perspective globale d'opération et de performance de MIPv4 ou MIPv6.

Bhagavathula et al. [5] proposent de considérer un réseau mobile équivalent à un site éloigné, et d'étudier les performances de différents types de technologies VPN pour répondre aux besoins d'évolutivité, de sécurité et de QdS. Deux types de trafics sont recommandés, soient ceux en temps réel et ceux du meilleur-effort (best-effort). MIPv4 est de mise pour cette étude de performance de bout-en-bout, mais encore une fois, lorsque nous étendons ce concept de mobilité à ceux d'évolutivité, de sécurité et de QdS (ou encore ESQ), l'établissement de tunnels IP-dans-IP pour assurer une gestion totale ESQ s'avère laborieux.

L'idée de Bhagavathula et al. d'utiliser une technologie VPN est bienvenue. Celle-ci a pour but de garantir une QdS à l'intérieur de tunnels, en plus de polices

d'admission pour supporter la sécurité, et finalement de l'agrégation de flots pour tendre vers une meilleure évolutivité. Bhagavathula et al. mentionnent que, contrairement à IPsec, la sécurité offerte par les VPNs de MPLS n'est pas de bout-en-bout, ce qui rend IPsec plus convenable. Dans notre cas, et à cause d'interceptes légaux requis à un niveau de granularité de flots, l'argument de ces chercheurs ne tient pas. La raison demeure qu'IPsec serait tenu de décrypter l'information à tout point d'intercepte demandé, ce qui l'empêcherait de toute manière d'effectuer un tunnelage IPsec de bout-en-bout. Les VPNs basés sur MPLS sont donc recommandés pour des questions de QdS et d'évolutivité. Trois scénarios sont alors considérés, soient :

- la performance d'un réseau typique de Mobilité IPv4, en utilisant des tunnels IP-dans-IP, sans sécurité et sans implémentation de QdS ;
- la performance d'un réseau semblable où les tunnels IP-dans-IP sont remplacés par un VPN basé sur IPsec ;
- finalement, la performance d'un VPN basé sur MPLS qui procure une connectivité adéquate entre le mobile et l'agent d'origine (HA).

Les types de trafics utilisés sont ceux de la voix et de données d'usager. Le banc de test consistait en deux réseaux d'accès rattachés par une dorsale pour effectuer le pontage. Un problème réside dans le fait que le FA ne peut pas diffuser toute sa table de routage au réseau d'origine, ce qui ne serait pas approprié puisqu'il appartient à un autre domaine administratif. Cette contrainte sera prise en compte dans ce mémoire puisque nous nous basons sur IPv6 pour la mobilité, qui se départit complètement du FA et utilise l'entité CoA d'extensions [5]. Les résultats qui en découlent sont:

- la qualité de la voix s'est détériorée dans le cas de VPN - IPsec comparativement au scénario classique de MIPv4, à cause des mécanismes additionnels d'encryptage et de décryptage des paquets ;
- la solution utilisant une technologie VPN - MPLS démontre une amélioration en termes de délai de bout-en-bout et de perte de paquets, lorsque comparée avec MIPv4 traditionnel et VPN basé sur IPsec.

Bhagavathula et al. [5] concluent en disant que IPsec requiert beaucoup de ressources de traitement pour assurer la sécurité, alors que la solution VPN-MPLS requiert assez de temps pour que MP-BGP puisse établir une paire avec d'autres routeurs PE. Cette contrainte est amoindrie dans notre recherche étant donné que les LSPs peuvent être prédéfinis et seront établis avant même qu'ils soient requis. De plus, l'aspect VPN demeurera au niveau de la dorsale seulement, sans déborder vers le réseau mobile d'accès.

2.3 Gestion de mobilité tout-IP

Chiussi et al. [6] sont d'avis que le tout-IP dans le réseau cœur UMTS constitue la tendance du futur puisqu'il permettra une meilleure convergence de transport, de multiples accès et finalement de différents services. En considérant la Figure 2.2, nous pouvons d'emblée identifier l'infrastructure courante d'un système UMTS.

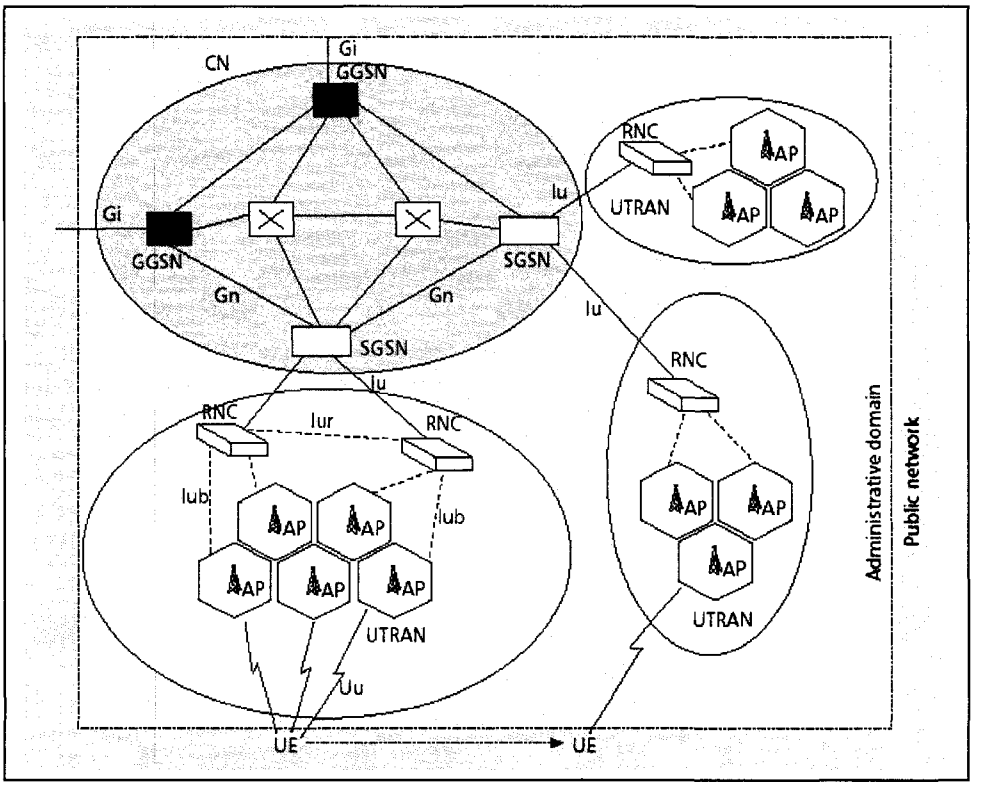

Figure 2.2 Représentation d'un réseau UMTS en mode transport

Nous y reconnaissons les interfaces Uu, Iub, Iur, Iu_ps, Gn et Gi. Cet arrangement est associé au mode d'opération transport puisqu'il utilise un transport GTP dans le réseau cœur. Le SGSN gère la mobilité inter-RNC, alors que le GGSN gère celle d'inter-SGSN. Le problème se présente lorsque la mobilité résulte d'un changement de point d'attachement GPRS initié par un nouveau SGSN. Les sessions GTP entre RNC et SGSN, et celle entre SGSN et GGSN doivent être rétablies. Tel que mentionné auparavant, le temps de réponse de GTP lors de relève RAU demeure inacceptable pour des types de trafic sensibles au délai.

Chiussi et al. [6] qui se concentrent plutôt sur la micro-mobilité à l'intérieur d'un RAU basé sur 3GPP, résument les différentes techniques de mobilité présentées durant les dernières années, et finalement proposent une gestion mobile hiérarchique basée sur MPLS. Cette solution consiste en l'utilisation d'un arrangement LEMA (Label Edge Mobility Agents) qui est évolutif, efficace et flexible. Il hérite alors des caractéristiques intrinsèques de MPLS concernant la QdS, l'ingénierie de trafic, les services IP et la restauration rapide.

La vision que proposent Chiussi et al. [6] en est une de diversité d'accès. Ils affirment que le réseau cœur devra en outre considérer, en plus des accès sans fil UMTS W-CDMA, ceux également de CDMA-2000, 802.11g. Nous aimerions y ajouter aussi la technologie d'accès WiMax décrite selon la norme 802.16. Chiussi et al. recherchent donc la QdS, la robustesse et la flexibilité de gestion d'un réseau cœur tout-IP pour faciliter les déploiements 3G utilisant plusieurs types d'accès sans fil. Deux types d'approches sont alors énumérés, soit celle du 3GPP (W-CDMA) et celle du 3GPP2 (CDMA-2000) incluant les accès sans fil 802.11. Le but visé est de supporter les services multimédias basés sur IP, la mobilité à grande échelle MIPv4 et MIPv6, la signalisation SIP pour l'établissement de sessions, et la gestion AAA (Authentication, Authorization, Accounting). Chiussi et al. [6] énumèrent trois raisons principales justifiant l'utilisation de MPLS à même l'infrastructure sans fil, et qui sont :

1. MPLS représente une technologie attrayante et légère de tunnelage qui se base sur l'établissement de LSP ;
2. MPLS s'avère bien rodé pour la redirection de paquets durant un événement de relève, en changeant tout simplement l'étiquette d'envoi ;
3. MPLS procure d'avance les avantages opérationnels énumérés plus haut qui sont la QdS, l'ingénierie de trafic, les services IP comme par exemple les réseaux privés virtuels, et finalement la restauration rapide.

Mais aujourd'hui, le mode d'opération utilisé en est un dit de transport où l'adresse IP du paquet transporté n'est pas utilisé pour en déterminer sa destination. En effet, ce même paquet IP est plutôt encapsulé dans un paquet GTP à l'intérieur du réseau cœur. Cette solution permet de préserver la compatibilité. Ce mode d'opération est représenté plus clairement à la Figure 2.3 [6].

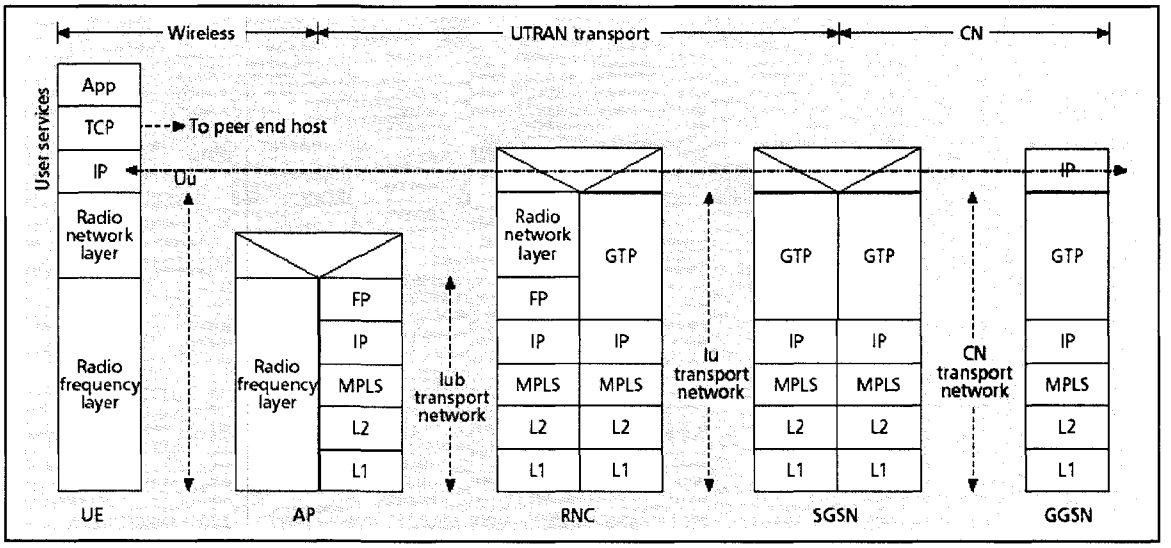

Figure 2.3 Représentation du mode d'opération transport

Selon cette figure, le mode transport existe entre les nœuds RNC, SGSN et GGSN et permet l'encapsulation de paquets IP à l'intérieur de paquets GTP. L'incongruence

technique qui se présente alors est celle de IP sur GTP sur IP qui s'avère selon tous, inefficace.

Une deuxième alternative, qui est celle du mode d'opération natif [6], dépend alors d'un mécanisme d'envoi de paquets utilisant l'adressage IP sans nécessairement impliquer d'autres couches intermédiaires. D'où l'amélioration immédiate de l'efficacité. IP devient alors la solution qui satisfait tous les types d'accès sans fil énumérés auparavant et rencontrés dans les réseaux hétérogènes. Cette alternative est illustrée à la Figure 2.4 [6].

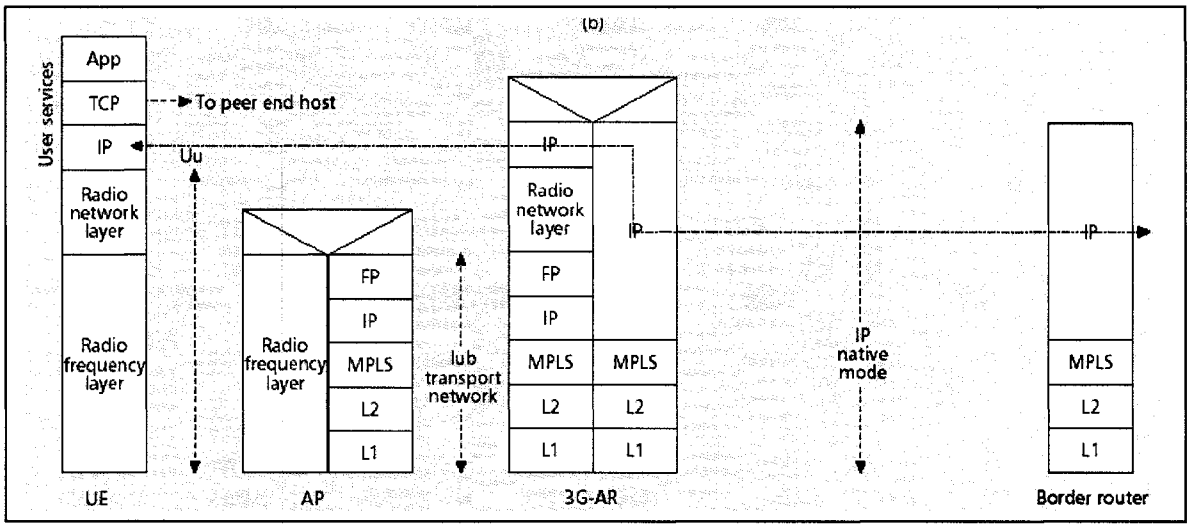

Figure 2.4 Représentation du mode d'opération natif

Dans cette figure, Chiussi et al. [6] proposent le mode natif dans le but d'utiliser IP/MPLS pour bénéficier de toutes les propriétés de IPv4 et de IPv6 pour la mobilité. Il est aussi possible de bénéficier des applications permises par MPLS, comme par exemple le transport rapide de paquets IP, la restauration améliorée, la garantie de QdS etc. Selon cette vue, les fonctions de Nœud_B, RNC, SGSN et GGSN sont tous co-localisées dans un nœud commun appelé le AR (Access Router). De plus, le reste du réseau cœur se compose de routeurs et de commutateurs typiques courants qui transportent les paquets IP selon les méthodes MPLS régulières. Quelques-uns des AR et des BR sont alors définis comme LER (Label Edge Routers) de façon à supporter l'établissement de LSPs.

La mobilité-IP basée sur IPv4 souffre de plusieurs limitations comme par exemple le routage triangulaire [6]. La mobilité-IPv6 y est plus de mise, en se débarrassant de l'agent étranger FA et en utilisant un MAP (Mobility Anchor Point). La solution finale proposée pour la micro-mobilité est celle d'un recouvrement hiérarchique appelé LEMA qui permet d'ajouter des fonctions avec sensibilité mobile. En conclusion, la micro-mobilité est assurée en changeant dynamiquement l'association d'adresse IP, au FEC à partir de signalisation spéciale.

Dans ce mémoire, nous ne considérons pas nécessairement une implémentation exhaustive de MPLS dans le réseau cœur, mais plutôt une conception de MPLS léger qui associe la fonctionnalité totale de MPLS de la Dorsale-IP, à celle allégée du réseau cœur. Nous encouragerons davantage l'idée de commutation rapide entre LSPs pour la QdS et la mobilité, et finalement considérerons la fonctionnalité VPN comme étant viable pour la sécurité. Somme toute, Chiussi et al. [6] considèrent les aspects de micro-mobilité, alors que nous considérons plutôt les aspects de macro-mobilité avec relève RAU inter-SGSN.

2.4 Concepts complémentaires de mobilité et d'intégration VPN

Afin de nous permettre de mieux comprendre les éléments importants constituant les fondements de mobilité tout-IP et d'intégration dans un cadre VPN, nous allons maintenant décrire quelques concepts additionnels qui s'y rattachent.

2.4.1 Améliorations de GTP issues de Mobilité-IP

Commençons tout d'abord par les améliorations de GTP provenant de mobilité-IP. Ceci nous aidera à aligner l'implémentation voulue vers un fondement tout-IP. Ces améliorations ciblées sont décrites comme suit:

- **Relève macro (Roaming) de l'hôte-IP et l'accès aux réseaux externes**: Sous l'égide de la Mobilité-IP, le NAI ("Network Access Identifier") est une méthode normalisée pour l'identification des usagers. Cette méthode est utilisée pour améliorer l'inter-opérabilité des services de relève Macro et de

tunnelage. En plus de l'addition de l'obtention d'un nom de réseau éloigné, le NAI inclut également le nom de l'abonné. Selon GPRS par contre, l'APN ("Access Point Name") est une identification de référence au GGSN à être utilisée pour s'interconnecter à la Dorsale-IP externe. De façon pratique, l'APN est utilisé pour décider à quelle adresse IP le tunnel **GTP** s'ancrera. Plus précisément, l'APN peut, à même le GGSN, identifier tout simplement la mise-en-correspondance du plus grand préfixe ("Longest Prefix Match") de la Dorsale-IP externe. L'APN permet donc de déterminer à quelle adresse IP un tunnel subséquent, originant d'un GGSN, sera terminé (exemple, L2TP ou IPsec). Un APN est composé de deux parties : une partie Identificateur-Réseau et une autre partie Identificateur-Opérateur. L'Identificateur-Réseau est une référence au service ou réseau externe, alors que l'Identificateur-Opérateur est une référence au réseau GPRS d'origine, dans un cas de relève inter-systèmes (roaming). Le NAI de Mobilité-IP et le APN de GPRS ont plusieurs différences mais ils supportent pratiquement le même besoin de mobilité macro : ce sont des outils pour procurer à l'abonné une façon de décrire à quel réseau externe IP il doit accéder pour atteindre le service voulu, mais en utilisant des adresses de type nom qui démontrent une dépendance géographique. Une limite importante de l'APN est qu'il n'est pas reconfiguré dans des conditions de relève inter-systèmes (roaming) alors que le NAI l'est, ce qui, dans le cas de l'APN, force l'utilisation d'une multiplicité d'allers-retours du chemin des données. Notre solution (GTP évolué) recherche donc une méthode d'adressage basée directement sur IP.

- **Support PPP:** Si on se limite au RFC-2002, il ne spécifie pas comment Mobilité-IP est utilisée quand les nœuds mobiles se connectent à leur fournisseur de service Internet (ISP) en passant par PPP (Point-to-Point-Protocol). Cependant, le RFC-2290 définit un support PPP dans un monde de Mobilité-IP. Un besoin de base imminent est la compatibilité avec la fonctionnalité d'authentification de l'ISP. Dans GPRS, il existe une

fonctionnalité similaire pour le tunnelage PPP transparent entre la station mobile et le GGSN. PPP peut être transporté par GTP. Une autre option est que IP (sans PPP) peut aussi être transporté sur GTP. Si l'option IP utilisé directement sur GTP s'applique, quelques-unes des fonctionnalités spécifiques à PPP sont perdues, telle que la sécurité. Par contre, une simple demande ou réponse d'authentification basée sur PPP peut être quand même supportée par GPRS, quand IP est transporté par GTP. Notre moyen de transport s'occupera donc de maintenir un certain niveau de sécurité en y associant une fonctionnalité potentielle VPN, pour les communications privées.

- **Allocation dynamique de l'adresse IP :** Dans le RFC-2002, il est assumé que le nœud mobile inclut son adresse d'origine permanente pour enregistrement. De plus, l'adresse de l'agent d'origine (Home Agent ou encore HA) est incluse dans le message d'enregistrement, sinon l'agent étranger (Foreign Agent ou encore FA) fait suivre le message au HA. L'extension NAI de Mobilité-IP a été proposée pour s'occuper de l'affectation temporaire de l'adresse d'origine. Une station mobile peut utiliser un NAI plutôt qu'une adresse d'origine dans la partie principale du 'Registration_Request' de Mobilité-IP. L'agent d'origine peut alors allouer une adresse IP à l'hôte mobile en utilisant le principe d'allocation dynamique d'adresse. Le 'Registration_Reply' sera alors envoyé du réseau d'origine à l'agent étranger (ou le MAP dans le cas de IPv6), qui extraira l'information voulue. Ce message inclut l'adresse IP allouée au hôte mobile [2]. L'allocation dynamique d'adresse est aussi une partie intégrale de GPRS. Durant la procédure d'activation de contexte, une station mobile GPRS peut indiquer si elle requiert une adresse IP statique ou dynamique. Le GGSN peut allouer l'adresse intérieurement ou encore utiliser un serveur AAA (Authentication, Authorization and Accounting). Il est aussi possible de négocier l'adresse IP après complétion de la procédure d'activation de contexte. Cette négociation prend place à partir de la station mobile avec le

réseau IP externe et devrait normalement utiliser PPP transporté sur GTP. Encore une fois, ce dernier est basé sur une adresse de type nom qui fait preuve de dépendance géographique. Le mécanisme d'évolution de GTP proposé utilisera un adressage de type direct IP. Ceci nous permettra une désensibilisation géographique et rendra possible l'acheminement optimisé non triangulaire.

- **Découplage d'avec les technologies de réseaux d'accès :** Avant que Mobilité-IP puisse être déployée dans les réseaux futurs, il existe un besoin réel qui est l'inter-opérabilité entre les protocoles existants dans les réseaux cellulaires. Pour permettre à Mobilité-IP de devenir une solution qui supporte plusieurs sortes de réseaux (ou technologies) d'accès différents, la fonctionnalité de Mobilité-IP doit être indépendante de la technologie des réseaux d'accès. Une séparation de la procédure d'authentification demeure motivée par le fait que les ressources radio sont réduites et que le nombre de messages au mobile doit être diminué. Également, l'opérateur du réseau d'accès peut ne pas vouloir permettre de la signalisation Mobilité-IP jusqu'à ce que le réseau d'accès sans fil ait accepté de procurer les ressources voulues à la station mobile. GTP évolué procure ce support de découplage de technologie d'accès en considérant un réseau cœur commun pour les différentes sortes d'accès.

2.4.2 État de Convergence dans les Normes GPRS/UMTS R99

Il existe déjà un certain niveau de convergence entre la Mobilité-IP et GTP qui est supporté par les spécifications UMTS R99. Optionnellement, une fonctionnalité d'agent étranger (ou MAP dans le cas de IPv6) peut être fournie à l'intérieur du GGSN. L'interface entre le GGSN et l'agent étranger, incluant la mise-en-correpondance entre l'*adresse-IP-au-soin-de* (care-of-IP-address) et les tunnels GTP, est assumée non-normalisée si le GGSN et l'agent étranger sont intégrés dans un seul nœud. En d'autres mots, ceci devient la préférence du manufacturier / opérateur. En principe, il est aussi

possible d'intégrer le SGSN et le GGSN en un seul nœud. Cependant, dans ce dernier cas, GTP devra toujours être supporté dans le but de permettre la relève inter-SGSNs. Les spécifications courantes de Mobilité-IP ne peuvent pas être utilisées pour assurer des relèves sans perte de paquets. Ceci sera supporté sous Mobilité-IP Hiérarchique [7] qui sera discutée dans la section 3.1. Notre recommandation de GTP évolué visera donc la relève sans perte de paquets.

2.4.3 Les besoins de MIPv6 requis par GTP évolué

Notre solution GTP-évolué supportera les spécifications de MIPv6 suivantes:

- Le protocole MIPv6 permet à un nœud mobile de se déplacer d'un lien physique à un autre sans changer l'adresse IP de ce nœud mobile. Ce dernier demeure toujours adressable à partir de son adresse d'origine (home address), qui est une adresse IP assignée au nœud mobile à même son préfixe de sous-réseau d'origine, sur son lien d'origine. L'aiguillage optimisé de paquets est alors rendu possible.
- Le MIPv6 Hiérarchique, communément appelé HMIPv6, introduit un nouveau nœud connu sous le nom de MAP (Mobility Anchor Point) et de nouvelles extensions mineures au Nœud Mobile et aux opérations de l'Agent d'Origine (Home Agent). Ce concept minimise la latence due aux relèves de terminaux entre les routeurs d'accès puisque c'est plus rapide d'associer des mises à jour à partir d'un MAP local plutôt que d'un Agent d'Origine distant.
- Également comme faisant partie de l'inter-opération entre RSVP et MIPv6, un mécanisme sous-jacent de support de mobilité est requis pour procurer une identité unique du flux de paquets d'un usager donné et ce, sans égard à la mobilité de cet usager, rendant ainsi la mobilité transparente. Ceci est possible à partir d'un objet de mobilité ajouté au protocole de signalisation RSVP. Ce mécanisme permet donc de diminuer le délai de signalisation et ainsi d'amoindrir les délais/pertes durant la relève de mobiles.

- Le bi-chemin (bi-casting) de trafic est un concept qui enlève l'ambiguïté due à des chemins asynchrones, lorsqu'il devient temps d'envoyer du trafic à partir d'un Nœud Mobile jusqu'à son nouveau point d'attachement, suite à une relève rapide (Fast-Handover). Le bi-sentier permet ainsi le découplage des relèves de couche 2 de celles de couche 3.
- Finalement, la QdS conditionnelle pour supporter les associations de mise à jour (Binding-Updates) est ici introduite. Le nœud de commutation, où les deux chemins du flux de données d'usager divergent, prend la décision finale s'il doit mettre à jour l'association d'adressage, dépendamment du résultat de mesure de la QdS. Ceci prendra place seulement si tous les nœuds le long du chemin entre le routeur d'accès et le routeur de commutation sont capables de rencontrer la demande en QoS.

Il devient évident que la portion mobilité HMIPv6 aidera à rendre plus performante, la mobilité dite micro et celle dite macro, dû au fait d'ajouts à la Version 6 de IP, mais encore une fois, l'utilisation de IPv4 ou de IPv6 dépend du choix de l'opérateur du réseau d'accès sans fil. Ces ajouts supportés par la solution GTP évoluée sont les suivants :

- **Optimisation de route** (en IPv6) permet un routage direct de n'importe lequel nœud correspondant, à n'importe laquelle station mobile, tout cela sans devoir passer par le réseau d'origine de la station mobile (identifiée MN sous IPv6) et d'être par la suite redirigé par sont agent d'origine. En d'autres mots, ceci écarte complètement le problème d'acheminement triangulaire présent sous IPv4. GTP a de la difficulté avec le routage direct tel que mentionné auparavant.
- Il y a également support pour permettre aux stations mobiles et à la Mobilité-IP de coexister efficacement avec des routeurs qui performent **'Filtration à l'Ingres'**. En effet, la station mobile maintenant utilise son *adresse-au-soin-de* comme Adresse-Source dans l'entête IP des paquets qu'elle envoie, permettant ainsi à ces paquets de passer normalement au travers des routeurs

avec Filtration à l'Ingres. L'adresse d'origine des nœuds mobiles est transportée dans chaque paquet, à l'intérieur d'une option d'adresse de destination, permettant ainsi l'utilisation d'une *adresse-au-soin-de* dans le paquet, transparent à la couche IP. La capacité de traiter l'option d'Adresse d'Origine dans un paquet reçu est requise dans tout nœud IPv6.

- L'utilisation des **options destination** de IPv6 permet à tout trafic de contrôle de type Mobilité-IPv6 d'être mis en 'PiggyBack' sur tous paquets IPv6 existants. Ceci diminue le besoin d'entêtes IP, en concaténant le contenu avec des paquets subséquents.
- Les entités **d'Agents Étrangers** tels qu'utilisés sous IPv4, **disparaissent** sous IPv6. Les fonctions de *Découverte de Voisins* (Neighbour Discovery) et d'Auto-configuration d'adresses sont utilisées pour opérer depuis n'importe quel endroit éloigné de l'origine.
- Le mécanisme de détection de mouvement, à l'intérieur de Mobilité IPv6, fournit une confirmation bidirectionnelle de l'habileté d'une station mobile de communiquer avec son routeur par défaut, à l'intérieur de son emplacement courant.

Lorsqu'une station mobile est éloignée de l'emplacement d'origine, son agent d'origine (HA) intercepte tous les paquets destinés à la station mobile qui arrivent au réseau d'origine, en utilisant la découverte des voisins basée sur IPv6 plutôt que d'utiliser ARP tel que dans les réseaux IPv4.

2.4.4 La spécification TR23.923 concernant l'enregistrement d'un mobile

Le support de mobilité basé sur IP ou encore Mobilité-IP, permet à un nœud mobile de maintenir la connectivité à l'Internet ou à un réseau corporatif tout en utilisant une seule adresse non-changeante (son adresse originale) même quand le point d'attachement du lien change.

Quand le nœud mobile se déplace du réseau d'origine au réseau étranger, il enregistre une adresse IP (voir Figure 2.5) avec son Agent d'Origine (Home Agent) sous

IPv4. Cette adresse IP, aussi appelée CoA (Care of Address) ou *adresse-au-soin-de*, pourra être utilisée pour tunneler les paquets associés au nœud mobile, en question. Le HA intercepte alors tout paquet adressé à l'adresse d'origine du nœud mobile et les tunnels au CoA. Aucune interaction avec les registres de localisation UMTS n'est requise.

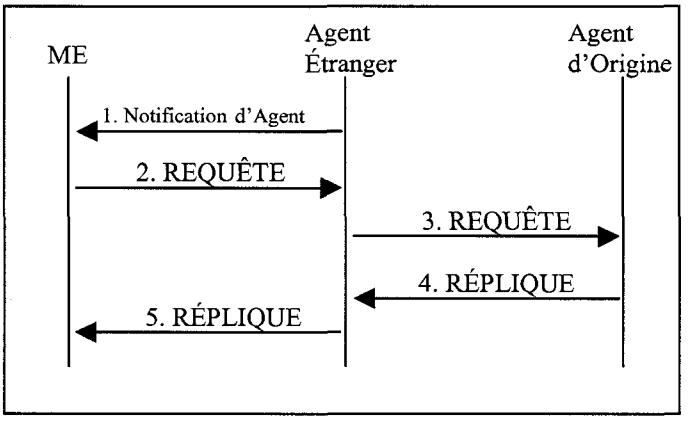

Figure 2.5 Enregistrement Mobilité-IP d'un ME visiteur

Le CoA est une adresse publicisée (ou dérivée d'une manière quelconque) par un Agent Étranger (FA ou encore Foreign Agent). Dans le cas de IPv6, le concept de FA disparaît et est remplacé par le MAP (Mobility Anchor Point) qui s'occupe d'associer l'adresse locale d'un mobile avec une adresse CoA, à même le paquet IPv6. Mais dans le cas de IPv4, c'est le CoA qui est dérivé du FA. Ce FA devient le point de terminaison du tunnel en question. Le FA extirpe les paquets du tunnel et les redirige au lien logique du RAN approprié, de manière à livrer ces paquets au nœud mobile ciblé. D'où le besoin de définir quelques interactions/mécanismes avec la couche 2 ou couche liaison du réseau d'accès.

Les tunnels inverses (ce qui signifie de l'Agent Étranger à l'Agent d'Origine) sont nécessaires pour la Mobilité-IP (ou du MAP à l'Agent d'Origine, dans le cas de IPv6), pour deux raisons: rendre sécure l'accès aux réseaux éloignés, et éviter le rejet de paquets à cause de filtration d'ingres. Un tunnel bi-directionnel de bout-en-bout peut résulter en

un acheminement non-optimal, mais il peut quand même être désirable de tunneler des paquets vers le réseau d'origine.

Le développement d'un réseau GPRS vers un réseau principal IP peut prendre place en trois étapes, toutes compatibles avec des versions antérieures de réseaux et terminaux capables d'opérer avec Mobilité-IP améliorée. Brièvement, ces étapes possibles sont :

- Étape 1 demeure une configuration minimale d'un opérateur, qui désire offrir de la Mobilité-IP améliorée (ou encore Mobilité accrue). La structure courante de GPRS est conservée et soutient la mobilité à l'intérieur du PLMN (circuit switched Public Land Mobile Network), alors que la Mobilité accrue permet aux usagers de se déplacer vers d'autres systèmes, tels que LANs, et des Accès UMTS sans perdre la continuité d'une session (par exemple TCP).
- À l'étape 2, des méthodes d'acheminement plus efficaces peuvent être obtenues suite à des relèves entre SGSNs en changeant de GGSN/FA ou encore de GGSN/MAP (sous IPv6), auquel l'Équipement Mobile (ME) est attaché pour la communication en cours. En maintenant pour une courte période de temps des tunnels du nouveau SGSN aux deux, l'ancien et le nouveau GGSN/FA/MAP, les problèmes de pertes de paquets sont minimisés. Pour les Équipements Mobiles qui transfèrent des données durant les relèves entre SGSN, le changement de GGSN/FA/MAP peut être fait après que le transfert de données a été complété.
- L'étape 3 consiste à combiner le SGSN et GGSN en un seul nœud, appelé le IGSN et de laisser la Mobilité-IP améliorée s'occuper des relèves entre IGSN, c'est-à-dire la mobilité à l'intérieur du Réseau-Cœur-PLMN et entre réseaux.

Il est important de noter ici que la troisième étape est celle considérée sous le concept utilisé dans ce mémoire de recherche qui s'appelle GTP évolué avec compatibilité VPN, dans le but d'une simplification d'implémentation à l'intérieur des routeurs IP. Cette troisième étape permet aux mécanismes de Mobilité-IP améliorée de s'occuper de la mobilité à l'intérieur du réseau de cœur, la mobilité PLMN, de même que

la mobilité entre les Aires de Desserte (Routing Areas) du domaine de paquets commutés (Packet Switched). La fonctionnalité du SGSN et celle du GGSN sont combinées en un nœud, tel que déjà mentionné, l'IGSN (Improved GPRS Support Node) et quelques autres fonctions sont ajoutés pour utiliser la Mobilité-IP dans le but d'opérer les déplacements entre IGSNs. Dans ce cas, le IGSN/FA/MAP sera le nœud qui marquera la terminaison de la partie d'UMTS spécifique au PLMN. Les fonctionnalités de base du IGSN sont donc:

- le support de la gestion de mobilité UMTS/GPRS à travers le UTRAN/BSS ;
- le support de la signalisation MAP (Mobile Application Part) qui est une entité complètement différente du MAP (Mobile Anchor Point) d'IPv6 ;
- l'interaction avec le HLR du PLMN, en passant par le FA/MAP avec une infrastructure AAA (Authorization, Authentication, Accounting) ;
- la capture de données de tarification et son formatage selon les spécifications UMTS/GSM; les spécifications de l'IETF peuvent être utilisées pour la comptabilité du FA ou du MAP ;
- le support de Mobilité-IP avec la fonctionnalité nécessaire pour être compatible avec le déploiement de Mobilité-IP dans des réseaux non-UMTS partout dans le monde ;
- le support des relèves inter-IGSN, fait à partir de mécanismes de Mobilité-IP ou GTP.

À partir du troisième scénario (qui consiste en l'intégration du SGSN avec le GGSN et ce, en un seul nœud appelé l'IGSN, et complémenté par la Mobilité-IP améliorée), le HA ou MAP (selon la version de IP utilisée) agira comme point d'attachement pour le trafic généré par l'Équipement Mobile (ME) si le tunnelage inversé est utilisé. Sinon, ce trafic sera acheminé directement vers le nœud correspondant. Si les mécanismes d'optimisation de route sont disponibles et déployés, le point d'ancrage sera utilisé principalement dans un but de contrôle, tandis que le trafic sera acheminé normalement le long des chemins tout en évitant les problèmes de cheminement triangulaire. Les MEs sans Mobilité-IP peuvent être supportés en laissant l'IGSN

enregistrer le mobile avec un HA du PLMN. Alternativement, les SGSN et les GGSN peuvent être déployés en parallèle avec des nœuds de Mobilité-IP et/ou le IGSN peut aussi agir comme SGSN.

Il est important de noter ici que l'Agent Étranger avec *adresse-au-soin-de* demeure l'adresse temporaire du réseau visité à laquelle le réseau d'origine envoie les paquets qui lui sont destinés. Encore une fois, les paquets destinés au nœud mobile qui arrivent au réseau d'origine sont interceptés par l'Agent d'Origine et tunnelés jusqu'à l'Agent Étranger (FA). Une fois que les paquets sont rendus au FA, cet agent dé-tunnelle les paquets et les envoie au nœud mobile.

Un autre point à considérer consiste en ce que la tendance est d'utiliser l'APN (Access Point Name) comme mécanisme de sélection de service au lieu d'introduire un nouveau contexte PDP (Paquet Data Protocol), sans modifier le plan de contrôle des systèmes GPRS et UMTS, mais plutôt en transportant tous les messages de Mobilité IP dans le plan d'usager du système UMTS/GPRS. Dans notre cas, nous ferons promotion d'adressage direct IP qui n'a pas de dépendance géographique et permettra l'allocation dynamique d'adresses CoA sous MIPv6.

L'IGSN procurera les services de Mobilité-IP basés sur l'adressage direct IP qui est identifié par le ME (Mobile-Equipment). Bien qu'il soit considéré que tous les nœuds mobiles supporteront la Mobilité-IP à un point donné dans le futur, l'adressage direct IP peut être utilisé pour distinguer entre une requête de Mobilité-IPv4 et une Mobilité-IPv6.

2.4.5 Établissement d'une session Intra-SGSN

Dans le but de clarifier le tout, la procédure complète d'activation de session (PDP Context Activation suivi d'un enregistrement Mobile IP évolué) sera maintenant présentée. En général, après avoir reçu une requête de contexte d'activation PDP du ME, le IGSN envoie un *Activate_PDP_context_accept* au terminal mobile et déclenche un FA pour qu'il envoie une notification à la station mobile qui requiert l'activation d'une session, de la même manière que lorsque le FA se trouve au GGSN. L'établissement d'une connexion PPP et l'exécution d'une procédure d'attachement UMTS/GPRS ont été

ici omis par souci de simplification. Les flèches dénotent les messages entre les nœuds, alors que les losanges représentent la fonctionnalité du nœud. Le tout est tiré de la spécification 3GPP TR23.923. La 2.6 montre l'établissement d'une session Intra-SGSN.

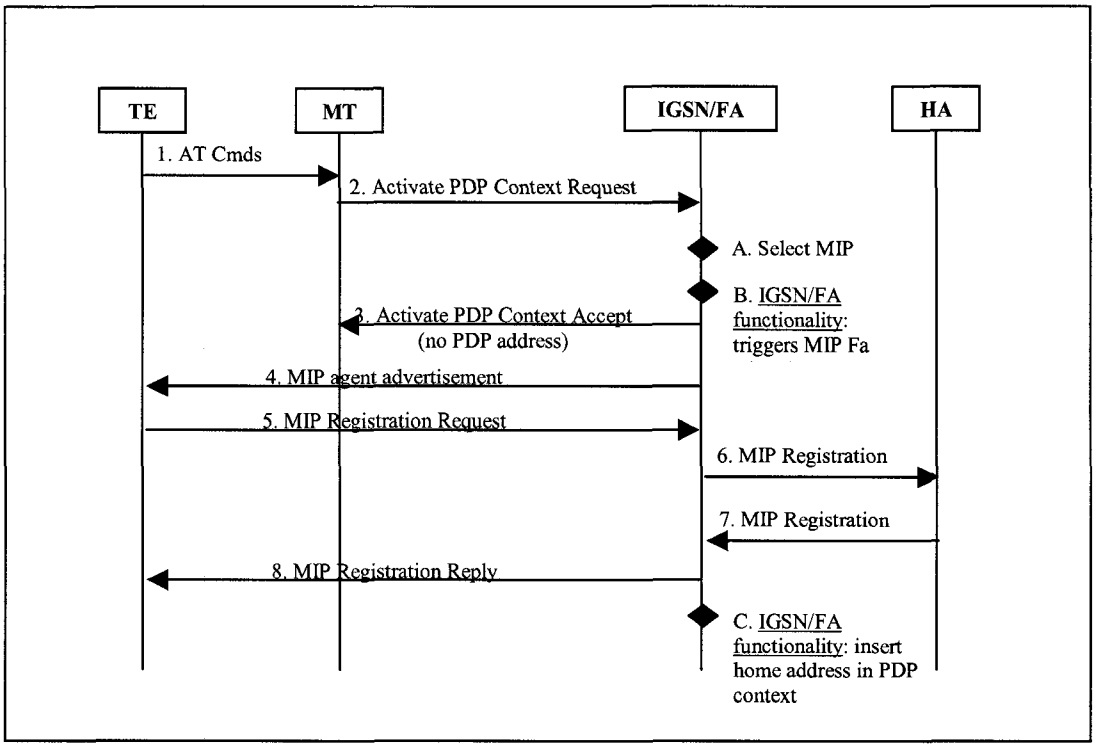

Figure 2.6 Activation de session proposée

Ce scénario est présenté en considérant que la ressource rare est la largeur de bande en provenance des accès sans fil. Les technologies de RSVP et de Diff-Serv (Differentiated Services) demeurent les mécanismes de QoS de choix considérés à même la spécification TR23.923 [8].

Considérons donc par exemple Int-Serv (Integrated Services) qui utilise RSVP comme protocole de signalisation. Des tunnels dans les deux directions (de HA à FA et de FA à HA) peuvent suivre des chemins déjà établis avec QoS, tout en utilisant des ressources appropriées de gestion de file d'attente et des mécanismes d'ordonnancement, de même qu'un routage basé sur des politiques et sur la classification. Alternativement, les réservations sur les chemins peuvent être établies en utilisant des extensions de

tunnels de type RSVP, mais dans ce cas, une encapsulation UDP est requise pour le transport des paquets sur tunnels RSVP. Lorsque Mobilité-IP est utilisée en conjonction avec un environnement Int-Serv, deux choses restent à considérer:

- la Mobilité-IP utilise l'encapsulation IP dans IP pour tunneler les paquets entre les agents de mobilité [9], et ces tunnels rendent les messages de bout-en-bout de RSVP invisible aux routeurs intermédiaires ;
- dans le cas de relève de Mobilité-IP, de nouvelles réservations le long du nouveau chemin de tunnel doivent être établies.

Le point principal que l'on veut mettre en évidence ici est d'avoir une session RSVP séparée entre les points-en-bout du tunnel. Le point d'entrée du tunnel sert de transmetteur pour le tunnel de la session RSVP, alors que le point de sortie du tunnel sert de récepteur. Le tunnel de la session RSVP peut exister indépendamment des messages RSVP de bout-en-bout.

Premièrement, plusieurs nœuds mobiles, utilisant le service de mobilité de ces agents mobiles, peuvent partager un tunnel RSVP et minimiser les états ajoutés dans le réseau. Deuxièmement, un nouveau tunnel RSVP peut être établi séparément pour chaque nœud ou flot. Ces deux alternatives représentent les deux extrêmes du spectre. Nous tendons dans notre cas à promouvoir une solution entre les deux, pour satisfaire les besoins de la cause. Donc, lorsqu'un nœud mobile se déplace vers un réseau visité, les réservations pour le nouveau tunnel doivent être établies. Dans le but de minimiser l'interruption de service durant la relève, le nouveau tunnel entre les agents mobiles peut être pré-configuré jusqu'à un certain degré.

Notre préférence, cependant, est d'utiliser Diff-Serv sur MPLS de manière à supporter l'établissement de session et la gestion de mobilité macro comme évolution de GTP. Diff-Serv peut aussi être utilisé dans le but de contrôle d'admission, en regroupant dans des classes agrégées, en effectuant un contrôle de politique, en conditionnant le trafic et finalement en marquant pour le rejet possible de paquets lors de congestions. D'un autre côté, MPLS peut devenir le mécanisme sous-jacent pour l'établissement de tunnels. Diff-Serv a un potentiel de haute évolutivité. Mais avant de sauter hâtivement

aux mécanismes proposés de GTP évolué, penchons-nous quelque peu sur l'idée d'intégration VPN pour les besoins de sécurité.

2.4.6 Concepts VPN (Virtual Private Network)

Par définition, une adresse VPN-ID est construite en concaténant un champ de longueur fixe, appelé un " Route-distinguisher", à une simple Adresse-IP. Un " Route-distinguisher" est structuré de telle manière à permettre à chaque fournisseur de service VPN de créer des " Route-distinguishers" de leur propre chef, sans le risque que le même " Route-distinguisher" soit assigné par d'autres fournisseurs.

Un "Route-distinguisher" consiste en trois champs, résultant en une entité de 8 octets, c'est-à-dire : un champ appelé *Type* (2 octets), un autre appelé *Autonomous System Number* (2 octets), et finalement l'*Assigned Number* (4 octets). Le champ Type fournit de l'information concernant le Type de VPN. Le champ *Autonomous System Number* contient le numéro du système autonome du fournisseur de service VPN. Finalement, chaque fournisseur de service VPN contrôle lui-même sa propre désignation du champ *Assigned Number*. Dans le cas le plus commun, un fournisseur de service y assigne ainsi une valeur numérique séquentielle à un VPN donné. De cette façon, aucun autre VPN ne partage le même "Route-distinguisher". Les adresses IP sont supposées être uniques à l'intérieur d'un VPN et, de ce fait, il s'ensuit donc que les VPN-Ids sont globalement uniques.

D'un point de vue BGP [10], traiter des routes avec des adresses VPN-ID n'est pas différent que de traiter des routes avec simples adresses IP, puisque la capacité multi-protocolaire de BGP le rend capable de traiter des routes de familles d'adresse multiples. La structure des adresses VPN-IP transformées en adresses VPN-ID, ainsi que la structure de la composante "Route-Distinguisher" du VPN-ID, est totalement opaque à BGP. Quand BGP compare deux préfixes d'adresse de VPN-ID, il en ignore tout simplement la structure. Nous pouvons donc en conclure qu'aucun nouveau mécanisme n'est introduit au protocole BGP lorsque le RFC-2547 est implémenté. Il est important de

souligner que la conversion de VPN-ID à une adresse VPN-IP prend place aux routeurs PE (Provider Edge).

Les adresses VPN-ID sont transportées seulement par les protocoles d'acheminement (routage), et non dans les entêtes IP. Ces identificateurs VPN ne sont pas utilisés directement pour l'envoi de paquets, ce qui est fait par la technologie MPLS. L'information d'accès est exprimée en termes d'adresses VPN-ID. Le routeur PE peut être imaginé comme un LER (Label Edge Router) MPLS. À cet effet, la Figure 2.7 illustre une configuration possible de réseaux à plusieurs VPNs.

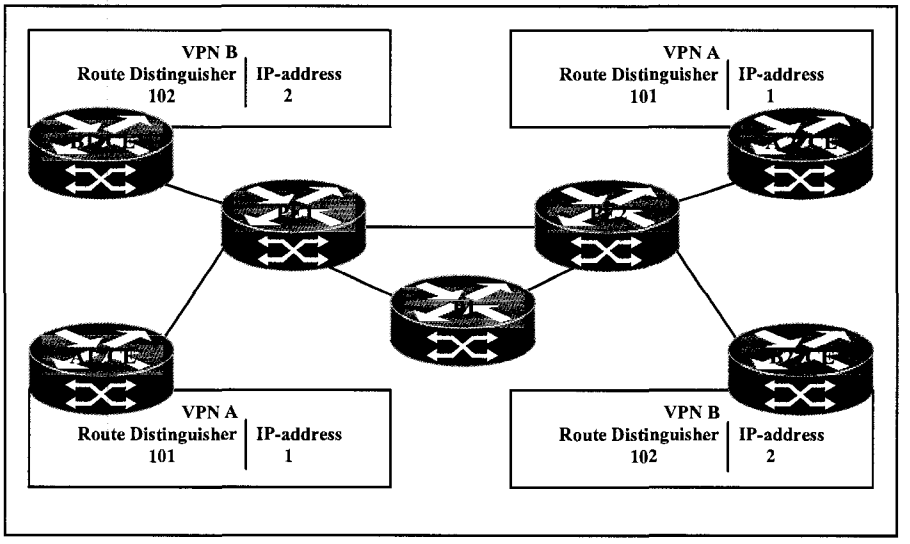

Figure 2.7 Représentation topologique de deux réseaux virtuels privés

Nous observons tout d'abord que deux VPNs y sont présents, soient VPN A et VPN B. Chaque terminaison VPN est représentée par un CE (Customer Equipment). Ces terminaisons se branchent à la dorsale IP à partir de PE (Provider Edge). L'aiguilleur P (Provider) agit comme nœud de transition.

Quand un routeur PE reçoit un paquet d'un usager [10] [11], il utilise l'interface d'arrivée du paquet pour identifier le VPN auquel celui-ci appartient et identifie la table d'envoi FIB (Forwarding Information Base) qui lui est associée. Ce FIB provient de la table tri-dimensionnelle VRF indexée dont chaque index correspond à un VPN ou FIB donné. Ceci fait, le routeur PE exécute alors une visualisation normale à partir de l'entête

IP, en utilisant l'adresse de destination IP lue dans le paquet en question. Ceci résulte en ce que le routeur PE ajoute l'information de l'étiquette au paquet et envoie ce paquet. L'évolutivité du réseau partagé VPN est contrôlée en utilisant une pile à deux étiquettes, celle à l'intérieur pour le routage BGP, et celle à l'extérieur pour le LSP du domaine ISP.

2.4.7 Concepts MPLS utilisés avec GTP évolué

Dans le monde de MPLS, les chemins LSP (Label Switched Paths) pour flots à fine granularité et pour les flots agrégés sont établis, maintenus et enlevés, en utilisant un des protocoles de signalisation, soit LDP (Label Distribution Protocol) ou RSVP-TE (ReSerVation Protocol – Traffic Engineering). Le protocole de distribution utilisé sera donc un moyen par lequel les LSRs (Label Switched Routers) établiront des LSPs (Label Switched Paths) au travers d'un réseau donné en faisant une mise en correspondance entre l'information de routage de la couche réseau et les chemins commutés de la couche liaison (couche 2). Ceci est accompli en utilisant des étiquettes de commutation. Ces étiquettes sont utilisées pour créer un paradigme simple d'envoi de paquets, appelé aussi commutation d'étiquettes.

Tel que mentionné auparavant, un des buts principaux de la commutation étiquetée de multi-protocoles ou encore MPLS, c'est d'assurer une augmentation de vitesse d'envoi, et une séparation de contrôle de la partie d'envoi pour offrir une robustesse accrue qui ne crée pas une interdépendance de fiabilité entre la partie traitement de paquets et celle d'envoi. Cette dernière est reflétée par la condition des liens entre les nœuds. Ainsi, si un processeur de routeur tombe en panne, la partie d'envoi peut continuer de fonctionner et vice-versa. Cette séparation entre contrôle et envoi demeure importante dans la pré-définition de nos chemins tunnelés parce que cette pré-définition ne peut être affectée que minimalement par une panne quelconque.

Le mécanisme d'envoi utilise une table d'envoi maintenue par le routeur lui-même et une étiquette attachée au paquet. Le mécanisme de contrôle, lui, s'occupe de la construction et de l'entretien de cette table d'envoi. En effet, l'envoi de paquets basés sur IP seulement consomme relativement beaucoup de temps pour le traitement du plus long

préfixe. Il est important de mentionner ici que tout ce qui existe déjà en termes de protocoles d'acheminement IP demeure et est ré-utilisé par MPLS pour le traitement d'envoi. Quelques étapes (simplifiées comparativement à IP) prennent place lors de l'envoi de paquets basés MPLS. Il va sans dire que MPLS emploie une approche de connexion virtuelle pour le transfert de paquets au travers du domaine MPLS. Ces étapes sont :

1. Lors de la mise en marche d'un réseau MPLS, un protocole de distribution d'étiquettes MPLS est activé, qui dérive des tables de routage pré-définies par les protocoles de routage IP. Ce protocole de distribution est soit CR-LDP ou RSVP-TE. Ce dernier demeure le choix de la majorité et sera celui considéré pour GTP évolué.
2. Un paquet IP entrant dans un domaine MPLS par l'intermédiaire d'un nœud d'entrée LER (Ingress ou Label Edge Router) est mis en correspondance avec une classe particulière appelée FEC (Forward Equivalence Class).
3. Une fois dans le LER, le FEC du paquet IP est associé à une opération NHLFE (Next Hop Label Forward Entry) qui peut être soit "PUSH", "SWAP" ou "POP". Dans le cas du nœud Ingress-LER, cette opération en est une de "PUSH" étiquette sur le paquet IP. Cette étiquette est une entité de 4 octets, composée d'une portion étiquette de 20 bits, d'une portion CoS (Class of Service), d'un indicateur Stack et finalement d'un champ TTL. L'association FEC à opération d'étiquette est permise par l'utilisation d'une table tridimensionnelle appelée FTN (FEC-to-NHLFE). Donc, d'un FEC, on obtient une opération NHLFE correspondante. En plus d'une correspondance FEC-to-NHLFE, nous considérerons un double emboîtement pour le support futur de tunnelage VPN dans le but d'améliorer la sécurité au niveau du réseau cœur UMTS de la solution GTP évolué.
4. Par la suite, les paquets MPLS sont dirigés vers les nœuds de transition appelés LSR (Label Switch Router) pour une commutation rapide basée sur des étiquettes de chemins pré-définies (LSP, Label Switched Path) lors de la distribution d'étiquettes. Chaque chemin est associé à un FEC donné (flot de paquets établi sur des caractéristiques communes telles que Adresse d'Origine, Adresse de Destination, Port

Logique d'Application, etc.), pour un traitement de chemin similaire. Chaque nœud de transition effectue donc une opération "SWAP" sur les paquets MPLS en utilisant une table bi-dimensionnelle ILM (Input Label Match).

5. Finalement, lorsque les paquets MPLS atteignent le nœud de périphérie de sortie, aussi appelé "Egress-LER" ou PE dans le cas de VPN pour GTP évolué, les paquets MPLS sont dépouillés de leur étiquette MPLS pour retrouver leur allure originelle de paquet IP.

L'avantage de cette commutation MPLS est donc un envoi rapide qui ne dépend plus d'un traitement d'une Entête IP de 20 octets mais plutôt d'un traitement rapide de commutation déjà défini selon un chemin donné (LSP) basé sur une entête MPLS de 4 octets. L'aspect orienté connexion donnera à la solution GTP évolué un potentiel additionnel de privatisation et d'agrégations de flots. En effet, il y a possibilité d'effectuer une encapsulation emboîtée à l'intérieur de plusieurs niveaux d'étiquettes MPLS, dans le but d'établir des tunnels de tunnels de chemins LSP. Ceci est très utile dans l'opération de réseaux privés virtuels ou encore dans le but d'effectuer de l'ingénierie de trafic, ou dans le cas de bris, d'utiliser le détour rapide de nœuds ou de liens en panne.

CHAPITRE III
AMÉLIORATIONS PROPOSÉES AU PROTOCOLE GTP

Ce chapitre décrit les améliorations proposées au protocole GTP dans le but de lui associer un chemin évolutif approprié. Pour ce, nous nous concentrerons sur les Versions 5 et suivantes d'UMTS, une technologie d'adressage et de mobilité (micro et macro) basée sur IPv6, un mécanisme de tunnelage orienté VPN pour des raisons de sécurité, et finalement un moyen de transport caractérisé par MPLS. Dans un premier temps, nous exposerons les fondements de ces améliorations. Dans un deuxième temps, nous présenterons en détails notre proposition d'améliorations de GTP. Nous complèterons le chapitre par une analyse du diagramme de séquence des messages selon GTP évolué.

3.1 Fondements des améliorations proposées

Notre proposition repose sur un certain nombre de mécanismes évolutifs combinant RSVP et MIPv6 [7], pour fins d'obtention d'un contrôle de mobilité IP plus efficace. Elle se veut une valeur ajoutée au protocole GTP en se référant à la spécification 3GPP de Version R99, de façon à mieux comprendre le besoin de migration vers les Versions 5 et suivantes. En effet, la Version R99 fait la promotion d'une couche 2 de type ATM, alors que les Versions 5 et suivantes tendent vers le tout-IP.

La Mobilité-IP de base (IPv4) est décrite dans le RFC-2002 et est intitulée "Support de Mobilité-IP". Ce RFC décrit comment acheminer les paquets à un nœud mobile qui n'est pas dans son réseau d'origine. Le transport de paquets à partir du réseau visité et jusqu'au nœud mobile est obtenu avec différents mécanismes de tunnelage décrits dans les RFC(s) 2003, 2004 et 2344 respectivement. La raison pour laquelle la Mobilité-IPv4 est décrite plutôt que celle de Mobilité-IPv6 est que cette

dernière se doit d'évoluer davantage vers une maturité convenable avant d'être utilisée de façon efficace dans un monde mobile radio.

La spécification 3GPP, étiquetée TR23.923 [8], présente les améliorations de protocole qui permettent l'acheminement transparent de datagrammes IP aux nœuds mobiles, en utilisant l'Internet. Chaque nœud mobile est toujours identifié par sa propre adresse d'origine, sans égard au point courant d'attachement à l'Internet. Lorsque ce nœud mobile est situé loin de son point d'origine, il est aussi associé à une *adresse-au-soin-de* (CoA), qui fournit de l'information concernant son point courant d'attachement à l'Internet. Le protocole de mobilité permet l'enregistrement de *l'adresse-au-soin-de* avec l'agent d'origine. L'agent d'origine envoie les datagrammes destinés au nœud mobile en passant par un tunnel jusqu'à l'*adresse-au-soin-de*. Une fois arrivé au bout du tunnel, chaque datagramme est ensuite délivré au nœud mobile.

Cette spécification 3GPP décrit aussi une méthode par laquelle un datagramme IP peut être encapsulé (transporté comme contenu de donné ou 'payload'), à l'intérieur d'un datagramme IP. L'encapsulation est suggérée comme moyen pour altérer l'acheminement normal IP pour les datagrammes, en les faisant passer par une destination intermédiaire qui, autrement, ne serait pas sélectionnée par le champ de l'adresse IP de destination dans l'entête IP d'origine. L'encapsulation peut servir à une variété de buts, tel que livraison d'un datagramme à un nœud mobile utilisant Mobilité-IP.

Cette spécification propose des extensions de compatibilité inverse de Mobilité-IP de manière à supporter des tunnels de retour. Mais TR23.923 n'essaye pas de résoudre les problèmes posés par les coupes-feu localisés entre l'agent d'origine et l'a*dresse-au-soin-de* du nœud mobile.

De plus, les serveurs AAA identifient typiquement les clients en utilisant le NAI (Network Access Identifier) de Mobilité-IP. Dans ce but, les spécifications 3GPP proposent que le NAI soit utilisé avec Mobilité-IP quand les nœuds mobiles émettent un 'Registration_Request'. Nous nous en tenons ici à ré-évaluer la situation en considérant plutôt l'utilisation directe d'adresses IP.

Mais avant tout, l'impact de cette spécification agit sur trois volets distincts, soient ceux de:

- l'enregistrement d'un mobile (Figure 2.5) ;
- l'établissement d'une session (Figure 2.6) ;
- la relève inter-GGSN (Figure 3.2).

L'intérêt sans cesse grandissant de Mobilité-IP comme solution potentielle de macro-mobilité pour les réseaux cellulaires nous amène à de nouvelles solutions et extensions de protocoles existants. Il existe aussi un besoin de clarifier les requis concernant Mobilité-IP, et ce, à partir d'une perspective de réseaux cellulaires. Ceci nous permettrait d'harmoniser l'évolution de Mobilité-IP avec les solutions existantes de mobilité dans les réseaux cellulaires, et ce, dans le but de migrer vers GTP évolué.

L'activation d'une session pour un ME (Mobile Equipment) demandant un service de Mobilité-IP et équipé avec un client de Mobilité-IP est quelque peu différent du modèle courant GPRS, dans le fait que le tout est traité localement dans l'IGSN, sans avoir recours au GGSN. À partir de la perspective d'un ME, l'activation de session et l'enregistrement initial de Mobilité-IP sont complètement identiques au cas où le FA est placé au GGSN. L'hypothèse est que chaque IGSN est équipé d'un FA ou MAP. La méthode d'établissement de session selon TR23.923 est décrite à la section 2.4 du chapitre 2.

3.2 Évolution proposée de GTP

La solution VPN considérée ici s'associe grandement au BGP/MPLS VPN, qui est une implementation du modèle pair. La raison pour laquelle ce modèle est appelé "pair" est que, d'un point de vue du routage, le réseau du fournisseur de service agit comme pair aux réseaux des usagers. Le BGP/MPLS VPN réfère quelquefois à la technologie de VPN de couche 3 basé sur MPLS. La norme associée à ce type de VPN est le RFC2547bis [10] et [11] mentionné auparavant. Il existe également une autre proposition connue sous le nom de technologie de VPN de couche 2 basé sur MPLS. Cette dernière est une adaptation à des technologies de couche 2 qui sont compatibles

avec des méthodes déjà existantes telles que celles de Relais de Trames, d'ATM, etc., mais ne sera pas utilisée ici, puisque c'est la couche 3 qui nous intéresse pour des raisons de mobilité.

Pour permettre à plusieurs VPNs de discriminer entre eux (par exemple, VPN-A et VPN-B), il demeure crucial de créer un nouveau type d'adresse, que nous appelons l'adresse VPN-IP ou VPN-ID, et nous nous assurons que ces adresses sont uniques, puisque l'environnement VPN seul possède des adresses IP qui sont privées. Ceci est dû au fait que les préfixes d'adresses peuvent être ré-utilisés.

Notre solution GTP évoluée jouira donc d'une capacité d'agrégation à deux niveaux avec traduction d'APN en un VPN_ID. D'où l'utilisation directe d'adresse IP et une compatibilité accrue BGP/MPLS VPN. Mais avant d'aborder une description succincte de ce mécanisme, considérons une perspective d'ensemble de cette proposition.

Nous sommes maintenant au point culminant de la proposition finale. Tel que mentionné auparavant, le protocole GTP encapsule en tunnels les paquets multi-protocolaires au travers du réseau cœur de GPRS, entre les nœuds GSN (c'est-à-dire SGSN et GGSN). Dans le plan de signalisation, GTP-C spécifie un tunnel de contrôle et un protocole de gestion permettant au SGSN de fournir des services GPRS pour un UE. Il est à noter ici qu'un UE du monde GPRS est équivalent à un MN du monde MIPv6. Cette signalisation crée, modifie et enlève les tunnels utilisés dans le but d'instaurer la macro-mobilité entre SGSNs ou entre GGSNs.

Le protocole UDP est utilisé comme protocole de transport pour transférer les messages de signalisation entre les nœuds GSN. Dans le plan de transmission, GTP-U utilise un mécanisme de tunnelage pour transporter les paquets de données de l'usager. Nous recommandons à cet effet d'utiliser plutôt un tunnelage basé sur MPLS pour des raisons de temps de réponse lors de relève. Les concepts de la technologie MPLS utilisés sont décrits à la section 2.4 du chapitre 2.

Nous pouvons donc remarquer que la technologie MPLS nous procure toute une panoplie d'outils et d'applications qui font partie des mécanismes disponibles à même

MPLS, et que nous traiterons dans la sous-section suivante. Nous identifierons par le fait même les applications que nous utiliserons plus précisément dans le cadre du mécanisme de GTP évolué.

Applications de MPLS utilisées par GTP-évolué

Il serait bon de mentionner brièvement quelques applications [12] ou utilisations possibles de MPLS comme technologie de commutation. Ces applications sont :

1. L'envoi et la commutation rapide de paquets IP ;
2. L'utilisation d'un mécanisme de QoS associé appelé Différenciation de Services ou encore Diff-Serv ;
3. L'utilisation d'encapsulation emboîtée (Label – Stacking) dans le but de permettre le détour rapide de nœuds en panne. Dans notre cas spécifique de GTP évolué, cette encapsulation emboîtée permettra la privatisation de flots avec l'établissement de groupes d'utilisation de réseaux virtuels (VPN) ;
4. La mise sur pieds de routes explicites basées sur de nouvelles contraintes de délai, de priorité, de largeur de bande, etc., et ce, en utilisant le protocole de routage CSPF (Constrained-OSPF) en complémentarité avec l'ancrage non-strict (Loose-Pinning) ou l'ancrage strict (Strict-Pinning) ;
5. La remise en service de sentiers rompus peut être implémentée en collaboration avec le TE-CMS pour le calcul d'optimisation hors-réseau..

Description d'un mécanisme à deux niveaux d'agrégation

Pendant qu'un nœud mobile est attaché à un lien étranger (dans un réseau visité) et loin du point d'origine, ce MN (Mobile Node) est aussi adressable par une ou plusieurs *adresses-au-soin-de* ou CoA, en plus de son adresse d'origine. Une CoA, comme mentionné auparavant, est une adresse IP associée à un nœud mobile pendant qu'il est en visite et est rattaché à un lien étranger. Le préfixe du sous-réseau de l'adresse CoA de ce MN est un préfixe sous-réseau (ou un des préfixes du sous-réseau) sur le lien étranger visité par le MN; si le MN est connecté à ce lien étranger pendant

qu'il utilise l'adresse CoA, les paquets adressés à cette adresse CoA seront acheminés au MN dans son point de localisation courante, loin de son point d'origine.

La notification et la procédure de découverte de l'agent demeurent inchangées, selon la norme [2]. Il devient donc possible d'établir des tunnels et des tunnels emboîtés dans des tunnels, de façon à implémenter différents niveaux de granularité à l'intérieur d'un réseau d'accès, tels que ceux de l'UTRAN et du réseau cœur d'UMTS.

Nous proposons donc de définir de l'agrégation de classes de flots telle que présenté à la Figure 3.1 (selon [13]) et, dans ce but, de construire une agrégation à deux niveaux avant de s'interconnecter au réseau dorsal IP, de la façon suivante :

- **Premier niveau d'agrégation** au premier routeur d'Accès ou regroupement des stations de base, aussi appelées Nodes B ;
- **Deuxième niveau d'agrégation** au IGSN ou regroupement des nœuds SGSN / GGSN.

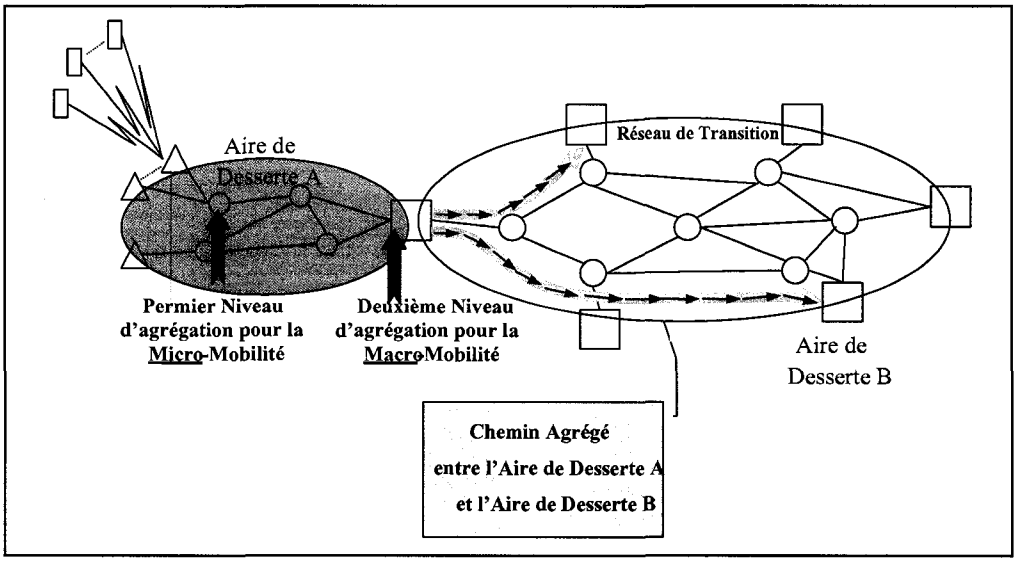

Figure 3.1 Deux agrégations à deux niveaux, LSPs pré-définis

Le premier niveau d'agrégation prend soin des micro-flots (uFlots) et de la micro-mobilité (équivalent au routage BGP-VPN), alors que le deuxième niveau d'agrégation servira à la macro-mobilité et combinera à cet effet le trafic de n x Access Routers (semblable au domaine ISP, de la Section 3.3) avant même d'atteindre le

réseau IP dorsal / multi-réseaux de transition, en se dirigeant vers une destination comme par exemple celle du "Routing Area B" (voir Figure 3.1). Cet emboîtement d'étiquettes à deux niveaux suivra celui décrit à la section 2.4.6 et réutilisera ainsi la pile à deux étiquettes proposée, avec une étiquette interne pour le routage BGP, et une externe pour le LSP du domaine ISP. Cette proposition sera toutefois le sujet d'une recherche future. De plus, le protocole GTP-C lui-même peut être sensiblement remplacé par de la signalisation RSVP-TE utilisée pour établir les LSPs de la technologie de MPLS. Ce protocole RSVP-TE peut à la rigueur évoluer pour permettre l'adaptation à la mobilité, ce qui ne sera pas décrit dans le cadre de ce mémoire.

Par exemple, le diagramme de séquence des messages présenté auparavant à la Figure 2.6 demeure toujours applicable. Cependant, pour ce qui est de l'"Inter-IGSN Routing-Area-Update", l'établissement de LSP peut être défini comme suit (Voir Figure 3.2).

La mobilité Inter-IGSN est prise en charge par Mobilité-IP, dans le cas de terminaux équipés avec un client de Mobilité-IP. Quand un Equipement-Mobile (ou encore ME) se déplace d'un ancien IGSN/FA vers la desserte d'un nouveau IGSN/FA, ce ME devra effectuer un enregistrement de type Mobilité-IP. Durant cet intervalle de temps nécessaire pour établir un tunnel LSP du HA jusqu'au nouveau IGSN/FA dans le cas de MIPv4, ou encore jusqu'au IGSN/MAP dans le cas de MIPv6, les paquets seront quand même envoyés à l'ancien IGSN/FA ou IGSN/MAP, en utilisant l'ancien tunnel de MIP.

En tenant compte de cette proposition-ci, MPLS-RSVP définit une procédure de transfert de paquets de l'ancien SGSN jusqu'au nouveau SGSN lorsqu'une mobilité inter-SGSN prend place et qu'un nouveau tunnel LSP au nouveau IGSN est en train d'être établi, tout en se basant sur des flots de chemin pré-défini et optimisé (en utilisant une entité hors-réseau ou externe telle qu'un TE-CMS). Similairement, un transfert de paquets de l'ancien IGSN au nouveau IGSN, en utilisant un LSP de MPLS peut être envié pour supporter le besoin de l'étape 5 du diagramme de séquence des messages de la Figure 3.2.

Pour que la Mobilité-IP puisse supporter convenablement la mobilité, il doit être possible de déclencher des enregistrements de Mobilité-IP, c'est-à-dire définir un moyen de détection de mouvement macro entre deux RAs (Routing Areas). Les messages de type "Routing-Area-Update" ont déjà été mentionnés à cet effet comme élément de déclenchement approprié pour la détection de mouvements macro. Nous pouvons donc affirmer que les "Inter IGSN RA-Updates" peuvent être utilisés pour déclencher un enregistrement de type Mobilité-IP.

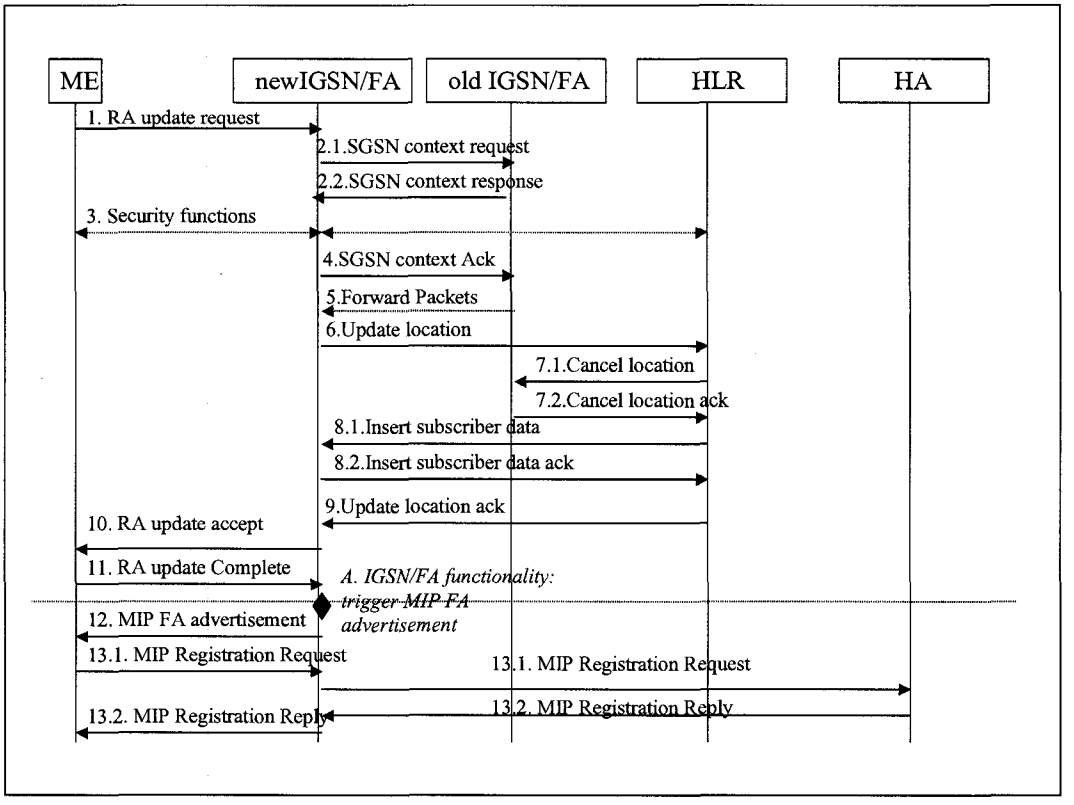

Figure 3.2 Mises à jour RA, utilisant des LSPs de MPLS

Finalement, une description à haut niveau de la procédure d'une mise-à-jour RA est présentée dans ce qui suit :

Dans ce diagramme de séquence de messages, nous pouvons voir que le "IGSN-context-request" déclenche l'utilisation d'un LSP agrégé pré-défini entre RA "A" and RA "B". Le type de LSP considéré en est un d'agrégation de niveau 2 (selon la

hiérarchie à deux niveaux définie auparavant), sans tenir compte si des changements de micro-flots prennent place au niveau 1 qui est encapsulé. En d'autres mots, la micro-mobilité devient opaque à la macro-mobilité à l'intérieur de cette proposition de solution, puisque l'agrégation de niveau 2 s'effectue à l'IGSN/MAP qui a déjà pris en charge le besoin de micro-mobilité. Il est important de souligner que le déclenchement d'un LSP implique de la signalisation jusqu'au nœud IGSN ciblé, d'un LSP pré-défini (en utilisant une pré-définition fournie par un TE-CMS hors réseau). Ceci est prévu de manière à ce que lorsque l'on se retrouve à l'Étape 5 mentionnée plus haut, l'accès à deux niveaux d'agrégation soit prêt à envoyer des paquets, avec un délai moindre d'établissement de tunnel. Cette pré-définition peut donc être accomplie avec un TE-CMS (Traffic Engineering – Configuration Management System) qui se base sur des algorithmes et/ou heuristiques pour l'optimisation de réseaux. Ce TE-CMS est bâti pour déterminer la topologie fonctionnelle à être optimisée en passant par plusieurs itérations de requêtes topologiques pour la gestion des ressources (c'est-à-dire largeur de bande, capacité de traitement, espace mémoire, etc.). Le tout s'avère utile pour la gestion de flots en réajustant l'information de réseau, à partir d'une boucle de réponse entre le TE-CMS et ce réseau en question.

Ce qui rend aussi cette solution nouvelle, est que ce même niveau 2 d'agrégation peut être considéré comme privé par un groupe donné d'usagers, d'où l'idée que des VPNs sont formés et définis explicitement. Les VPNs basés sur BGP, avec leur principe de communautés, peuvent être placés comme recouvrement des systèmes proposés, et ce, entre deux RAs spécifiques. Dans les réseaux cœur de UMTS, l'APN indique la référence au IGSN à être utilisée. D'une façon pratique, l'APN est utilisé pour décider à quelle adresse IP le tunnel GTP résultant (ou dans ce cas-ci le LSP résultant) s'y rattache. De plus, l'APN peut, à l'intérieur du IGSN, identifier le réseau IP dorsal externe. Donc, l'APN peut finalement être utilisé pour déterminer à quelle adresse IP un tunnel subséquent (i.e. PPP, L2TP or IPsec) originant de l'IGSN considéré, est attaché.

L'APN est l'entité d'identification dans UMTS qui informe l'Agent-d'Interconnection (Gateway-Agent), où le tunnel en question doit se rattacher. Les deux parties de l'APN sont pleinement qualifiées comme noms de domaine selon les conventions d'appellation réservées au DNS (Dynamic Naming System). Les VPN-ID peuvent ainsi émaner de l'APN ou plus spécifiquement selon notre proposition de solution, de la partie "Network Identifier". Dans ce cas, le réseau IP dorsal externe auquel nous nous référons, devient tel qu'attendu, spécifique à un VPN donné. Les détails d'implémentations de la mise-en-correspondance d'APN à VPN-Ids ne font pas partie de la recherche donnant lieu à ce mémoire. Mais la solution GTP évolué doit s'assurer qu'elle offre l'infrastructure requise pour la supporter pour éviter la dépendance géographique engendrée par le mécanisme d'APN.

Nous confirmons également que la spécification existante de RSPV-TE est assez mature pour établir des tunnels de type LSP pour la macro Mobilité-IP. Les adresses respectives d'origine, du FA ou MAP dans le but d'établissement de tunnels de type LSP, seront fournies durant les procédures de découverte d'agents, de MAP et d'enregistrement, inclues comme faisant partie de la panoplie des technologies identifiées MIPv4 ou MIPv6.

Finalement, non seulement avons-nous répondu au besoin d'évolution du protocole GTP vers un protocole plus adapté qui est celui d'établissement de LSP à partir de la technologie de MPLS, mais nous avons aussi couvert l'implémentation potentielle de VPN-ID à partir d'un concept d'agrégation à deux niveaux. Cet aspect fera l'objet de recherche future pour des besoins de raffinement et de validation de protocoles.

3.3 Le diagramme de séquence des messages

Penchons-nous maintenant sur les détails d'implémentations du mécanisme considéré de pré-définition de chemins. Pour ce, analysons le diagramme de séquence des messages de la Figure 3.3 simplifiée, avec emphase mise sur l'établissement de

chemins LSP pré-définis, ce qui nous aidera à diminuer le temps de réponse d'une relève impliquant un RAU (Routing Area Update).

La Figure 3.3 est une représentation simplifiée d'un RAU en utilisant des sentiers LSP de MPLS. Les étapes 1 et 2 représentent donc les messages de déclenchement d'une relève inter SGSN. Par la suite, la série de messages 3 et 4 viennent supporter l'accusé de réception du nouveau contexte et l'aspect de sécurité authentifiée au serveur AAA. Sans s'attarder sur ces éléments courants, considérons maintenant la partie ombragée verte qui représente l'essence même de l'évolution apportée au protocole GTP par cette proposition. Les primitives principales se présentent comme suit :

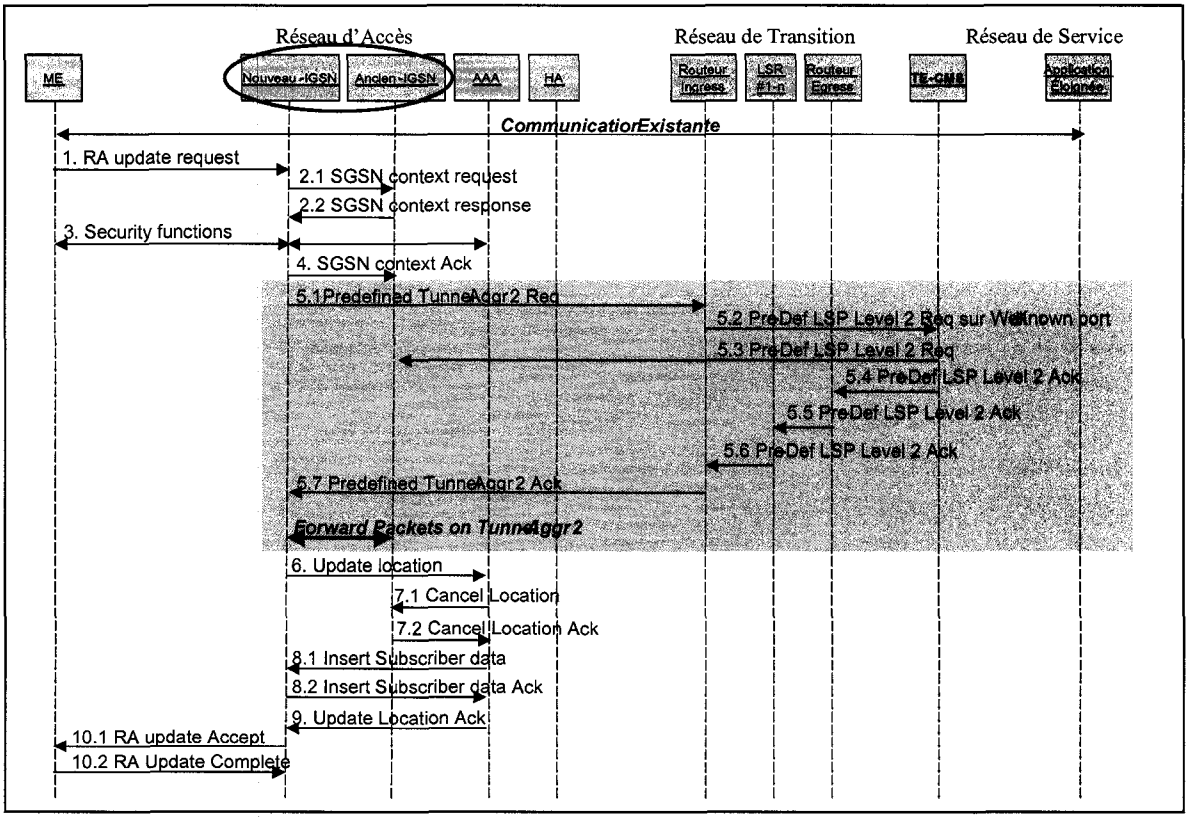

Figure 3.3: Mises à jour RA avec utilisation de LSP

- Message 5.1 **Predef Tunnel_Aggr 2 Req** initié par le nouveau IGSN pour rediriger les flots de communications, lors de la relève ;
- Messages 5.2, 5.3 **Predef LSP Level 2 Req** représentant la requête d'établissement de chemins, sur le réseau cœur constitué de nœuds MPLS, dont les nœuds Ingress, LSR et Egress. De plus, une entité TE-CMS (Traffic Engineering-Configuration Management System) est également utilisée pour le calcul hors-réseau, de chemins LSP pré-définis ;
- Messages 5.4, 5.5, 5.6 et 5.7 **Predef LSP Level 2 Ack** représentant l'accusé de réception de l'établissement de chemins, sur le réseau cœur dans la direction inverse jusqu'au nouveau IGSN.

Les étapes 6, 7.1 et 7.2 sont les éléments réguliers d'annulation du chemin antérieur entre le réseau cœur et le GGSN. Les étapes 8.1 et 8.2 décrivent les messages usuels de mise à jour des données d'usagers concernant sa nouvelle localisation qui est confirmée par l'étape 9 avec un **Update Location Ack**. Finalement, les messages 10.1 et 10.2 terminent la complétion de cette mise à jour de chemin ou RAU.

Nous devons comprendre que les étapes 5.1 à 5.7 peuvent être soumises à une exécution concurrente antérieure plus rapide, suite aux contraintes de fautes que nous vérifierons lors de la validation de protocole, sur *SPIN*. Il est aussi entendu que cette pré-définition de sentier LSP se fait hors-réseau par une méthode algorithmique d'optimisation qui devra, par approximation écourtée, répondre à l'intérieur d'une période d'harmonisation relativement rapide avec une marge de précision d'environ 5% de la largeur de bande effective.

Cet algorithme peut être décrit de façon simple, par les étapes suivantes :

1. Lire la topologie actuelle du réseau cœur et calculer la largeur de bande effective de chaque lien ;
2. Utiliser une heuristique approximative (e.g. recherche taboue) pour le dimensionnement du réseau cœur, compte tenu de la nouvelle assignation de chemins ;

3. Traduire cette assignation de chemins en commandes de configurations LSP (e.g. CLI) ;
4. Y a-t-il requête de nouveaux sentiers pré-définis en provenance du CN?
5. Si Oui, déployez la nouvelle assignation de chemins pré-définis et retournez à l'étape 1 ;
6. Si Non, retournez à l'étape 1, après un temps d'harmonisation **∇t** pré-établi.

La bande passante effective [14] considérée ici, est celle d'un processus arrivant **A**, et peut être définie comme suit :

$$\alpha(s,\tau) = \sup_{t\geq 0}\left\{\frac{1}{s\tau}\log E\left(e^{s(A(t+\tau)-A(t))}\right)\right\}, s,\tau \in (0,\infty) \qquad (1)$$

Le paramètre τ est appelé le paramètre de temps et indique la longueur d'un intervalle de temps. Le paramètre **s** est appelé le paramètre d'espace (space) et contient de l'information à propos de la distribution des arrivées de paquets. Généralement, la largeur de bande effective α d'un flot de trafic varie entre la moyenne et le régime de crête du trafic. α procure donc un lien entre les caractéristiques du trafic d'un flot et les ressources requises en termes de largeur de bande et de grosseur de mémoire tampon, dans le but de supporter un niveau de service garanti pour ce flot. Il en résulte donc :

- Lorsque s → 0, la largeur de bande effective se rapproche du régime moyen de trafic ;
- s → ∞, cette largeur de bande effective est principalement reflétée par le régime de crête du trafic.

L'utilisation principale de ce concept nous aide à comprendre que tant et aussi longtemps que la largeur de bande effective d'un groupe de flots demeure en deçà de la capacité **Cap** d'un lien donné, la probabilité de perte de paquets due à une surcharge de mémoire tampon diminue exponentiellement en fonction de la profondeur de cette mémoire tampon.

CHAPITRE IV
RÉSULTATS EXPÉRIMENTAUX ET ANALYTIQUES

Dans le chapitre précédent, nous en sommes venus à formuler, un algorithme permettant de fournir un cadre opérationnel pour l'utilisation de chemins LSP, ceci dans le but d'accélérer la relève macro de type inter-SGSN caractérisée par l'utilisation de RAU. Dans ce chapitre, nous présenterons et analyserons les résultats expérimentaux et analytiques découlant de nos propositions. Dans cette optique, nous présenterons premièrement l'environnement nécessaire à l'évolution de cette expérimentation. Dans un deuxième temps et pour en venir à une telle preuve de concept, nous aborderons le plan d'expérimentation concernant l'évolution de GTP (utilisé dans le cas des trafics mission critique tels que ceux de la voix et de la vidéo-conférence). Ceci fait, nous pourrons par la suite décrire les simulations (*OPNET Modeler*^{TM} et *SPIN*) effectuées dans le cadre de cette recherche et analyser les résultats obtenus. Finalement, nous serons en mesure de suggérer une démarche mathématique simplifiée pour le dimensionnement de tels trafics (voix et vidéo-conférence), dans le but de s'assurer qu'une telle implémentation est réaliste avec les aiguilleurs (routeurs) couramment disponibles.

4.1 Environnement de simulation

En ce qui concerne les outils de recherche utilisés pour l'évaluation de notre méthode d'amélioration de relève RAU, nous avons installé deux applications de simulation. La première consiste en un simulateur basé sur l'exécution d'événements temporels et appelé *Modeler*^{TM} *Version 10.5*. Ce logiciel provient de la compagnie Opnet Technologies Inc et inclut toute une librairie de protocoles et technologies de télécommunications normalisés tels que RSVP, MIPv6, OSPF, LDP, ATM, MPLS, UMTS, etc. Une description plus succincte d'*OPNET Modeler*^{TM} peut être trouvée à l'adresse Internet http://www.opnet.com . La deuxième application s'appelle *SPIN*

Version 4.0.7, un outil utilisé pour valider la consistance de systèmes concurrentiels (tels que les protocoles de communications), et pour vérifier certaines propriétés de Model Checking (dans ce cas-ci le blocage d'état et la sûreté). Il consiste à l'entrée, en la spécification d'un modèle du système global considéré, basé sur le langage Promela et procure en sortie une description du graphe d'état reflétant l'exécution des processus concurrentiels. *SPIN* (Simple Promela INterpreter) est un outil logiciel gratuit qui demeure disponible à l'adresse Internet http://spinroot.com . *OPNET Modeler*^{TM} et *SPIN* demeurent deux applications grandement utilisées à l'intérieur des institutions universitaires et commerciales, surtout dans le domaine de la recherche.

Les simulations sur *OPNET Modeler*^{TM} et *SPIN* ont été exécutées à partir d'un matériel utilisant une plate-forme caractérisée par les spécifications suivantes :

- Système Opérationnel Windows-2000 Professionnel de Microsoft
- Version 5.0.2195 du SO avec mise à jour SP 4 et constructeur 2195
- Modèle de l'ordinateur → HP Compaq nc8000
- Type de système logiciel → ordinateur basé sur X86
- Processeur → Pentium 4 M
- BIOS – Version 1.1
- 523,632 kilo octets de mémoire physique RAM
- 40 giga octets de disque dur
- compilateur pour langage C et C^{++} → Visual C^{++} V6.0
- Vitesse d'horloge → 1.5 GHz

Ces deux outils peuvent opérer simultanément sur la même plate-forme informatique. Le tout fonctionne dans un mode d'opération solitaire sans besoin de ressources réseaux, sauf pour l'obtention d'une licence *OPNET Modeler*^{TM} résidant sur un serveur de licence. Le temps de simulation dépend de la complexité des modèles définis, d'où l'intérêt important de se limiter à un certain niveau d'abstraction dans la description des modèles considérés. Une simplification adéquate de ces modèles fut

considérée à l'étape d'implémentation de la preuve de concept, et a été appliquée dans la définition des modèles (*OPNET Modeler*^{TM} et *SPIN*) en question.

Nous en venons maintenant à la description du plan d'expérience qui a été établi dans le cadre de cette expérimentation.

4.2 Plan d'expérience

Il convient maintenant de décrire les étapes suivies lors de l'expérimentation pour la preuve de concept de l'utilisation de chemins LSP, dans le but d'accélérer la relève macro de type inter-SGSN caractérisée par l'utilisation de déclenchement RAU. À cet effet, l'expérimentation mise en place se déroulera en quatre phases précises de manière à obtenir une perspective globale de l'amélioration du temps de relève RAU. Comme mentionné auparavant, un cadre général a été défini au chapitre 3, par la description de l'algorithme global de définition de chemins pré-définis. Ce même algorithme a été repris ici et présenté encore une fois sous forme schématique pour mieux comprendre l'interaction entre les fonctions utilisées; il est présenté à la Figure 4.1.

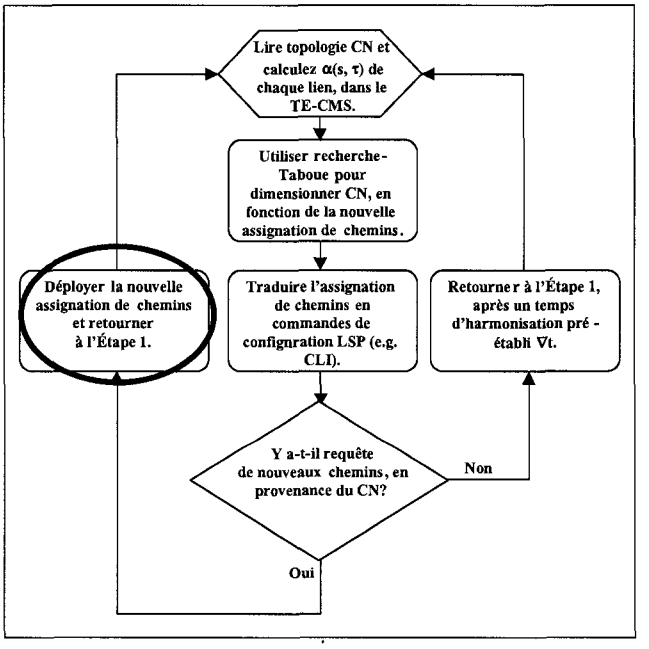

Figure 4.1 : Algorithme global de définition de chemins pré-définis

Il est important de souligner que la méthode détaillée d'optimisation hors réseau de la Figure 4.1 ne fait toutefois pas partie de cette recherche, mais sera considérée pour investigation future, comme cadre pour l'implantation d'un tel mécanisme de GTP évolué dans un contexte de Mobilité-IP et d'infrastructure VPN. Dans cette figure, seule la partie "Déployer la nouvelle assignation de chemins et retourner à l'étape 1" marquée en rouge, s'avère profitable pour le moment, puisque nous ne considérons ici que l'amélioration du temps de relève inter-SGSN. Ce cadre global y est défini pour mieux situer les limites de notre implémentation. À cet effet, quatre étapes principales d'analyse sont envisagées et se présentent comme suit :

1. Une première phase consiste en la caractérisation à partir du temps de relève RAU entre deux RA (Routing Areas). Cette mesure a été prise sur un réseau GPRS semblable au réseau UMTS 3GPP, en utilisant un équipement de test appelé *G_b-Analyzer*, de la compagnie *Netcare International.* Les résultats nous ont été transmis directement par cette compagnie.
2. Par la suite, une deuxième phase permet de proposer une solution d'amélioration du temps de relève en ayant recours à des chemins à étiquettes commutées, de manière à obtenir un temps de relève de moins d'une demie seconde. Cette implémentation sera mise sur pieds à partir de simulations utilisant l'outil *OPNET Modeler™*.
3. De plus, une troisième phase permet de valider le nouveau protocole de migration, nous dirigeant vers GTP évolué. Cette validation de protocole se fera au moyen de l'outil de vérification *SPIN.* Promela s'avère être un langage descriptif de systèmes concurrents. Tel est le cas de l'établissement parallèle de chemins LSP avant la relève RAU.
4. Finalement, une démarche analytique sera proposée pour le dimensionnement des trafics de voix et de vidéo-conférence, dans le but de s'assurer qu'une telle implémentation est réaliste avec les aiguilleurs (routeurs) couramment disponibles. Cette démarche utilise le même concept

de bande passante efficace qui fait partie de l'algorithme d'allocation de bande passante pour le choix des chemins appropriés.

Une fois ces quatre étapes effectuées, nous pourrons en tirer les conclusions recherchées de mécanismes possibles pour l'amélioration du temps de relève inter-SGSN. Il est important de noter ici que nous nous sommes limités à des modèles simples consistant, dans le cas de la simulation *OPNET Modeler*^{TM}, à un seul mobile se déplaçant d'un Nœud B_1 à un Nœud B_2, et dans le cas de la validation de protocole *SPIN*, à trois mobiles. En effet, l'intérêt dans la simulation *OPNET Modeler*^{TM} demeure l'évaluation du temps de relève inter-SGSN en assumant, pour les types de trafics considérés, qu'il n'existe pas de contraintes de charges détériorant la relève inter-SGSN. Pour ce qui est de la validation *SPIN*, une augmentation itérative du nombre de mobiles démontrera que trois mobiles sont suffisants pour déterminer si un LSP peut être établi sans congestion de messages concurrents (protocole d'établissement).

4.3 Expérimentation et analyse des résultats

Tout comme indiqué dans le plan d'expérience, nous débuterons premièrement par une mesure du temps de réponse de RAU dans le but d'établir une base de référence sur laquelle nous pouvons comparer.

4.3.1 Mesure du temps de réponse RAU

Considérons donc maintenant un réseau typique urbain GPRS/UMTS avec une pluralité de stations Nobe-B rattachées à quelques RNC, qui eux-mêmes sont rattachés à quelques SGSNs. Nous faisons déjà face à deux niveaux d'aggégation sans toutefois considérer le prochain niveau hiérarchique qui sera celui du GGSN.

Chaque SGSN nous procure donc la possibilité de pouvoir se déplacer à l'intérieur d'un RA. En faisant ainsi, nous entrons dans un mode de mobilité macro. Ce mode de relève s'appelle ainsi puisqu'il s'étend sur deux RAs (voir Figure 4.2).

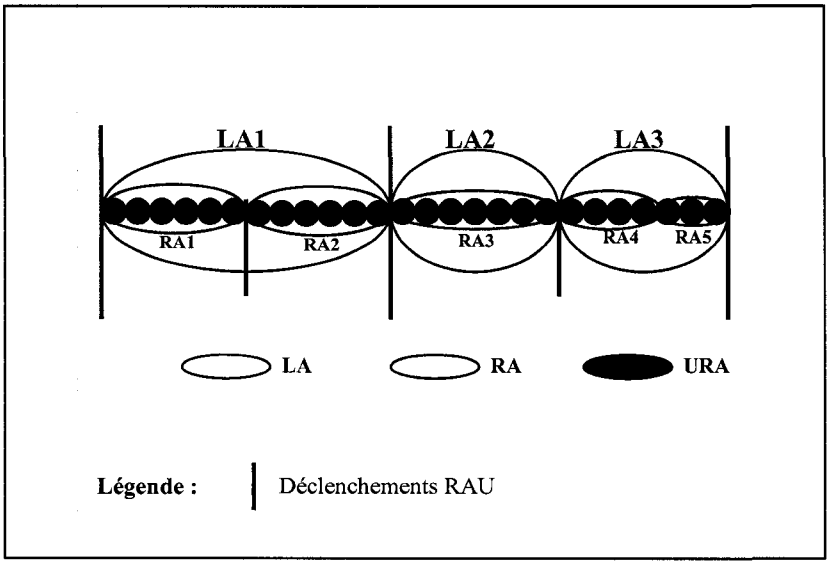

Figure 4.2 Représentation graphique des aires de mobilité

Un RA est une sous-division du LA (Location Area) et est géré par un SGSN. Il va sans dire qu'un déplacement entre deux RAs implique donc une relève entre deux SGSNs et est géré par le GGSN d'agrégation. De plus, un URA (UTRAN Registration Area) représente une sous-partie de RA, mais nous n'en tiendrons pas compte dans notre recherche puisqu'il existe un pluralité de URA à l'intérieur d'un seul SGSN.

L'instrument de mesure identifié comme étant l'*analyseur* G_b (pour interface G_b) a été branché à l'interface G_b entre le BSC/RNC et le SGSN d'un système typique GPRS, pour déterminer le temps de réponse de relève RAU. L'*analyseur* G_b consiste en une application qui fonctionne sur plate-forme Microsoft Windows. Le tout résulte en des temps moyens de RAU représentant deux scénarios, suite à un déplacement Inter-SGSN, d'un RA_A à un RA_B, et un autre type de mesure concernant une série de messages RAU envoyés à cause d'une mise à jour régulière. Les Figures 4.3 (temps de relève RAU, en ordonnée et le numéro d'identification de mobile, en abscisse) et 4.4 présentent les résultats obtenus. Les temps réponse moyens encourus sont donc de 2 à 10 secondes pour les RAUs dû à une relève et d'environ 0.7 seconde pour les RAUs dus à des mises à jour régulières. Une autre source confirmant ces résultats, est [15].

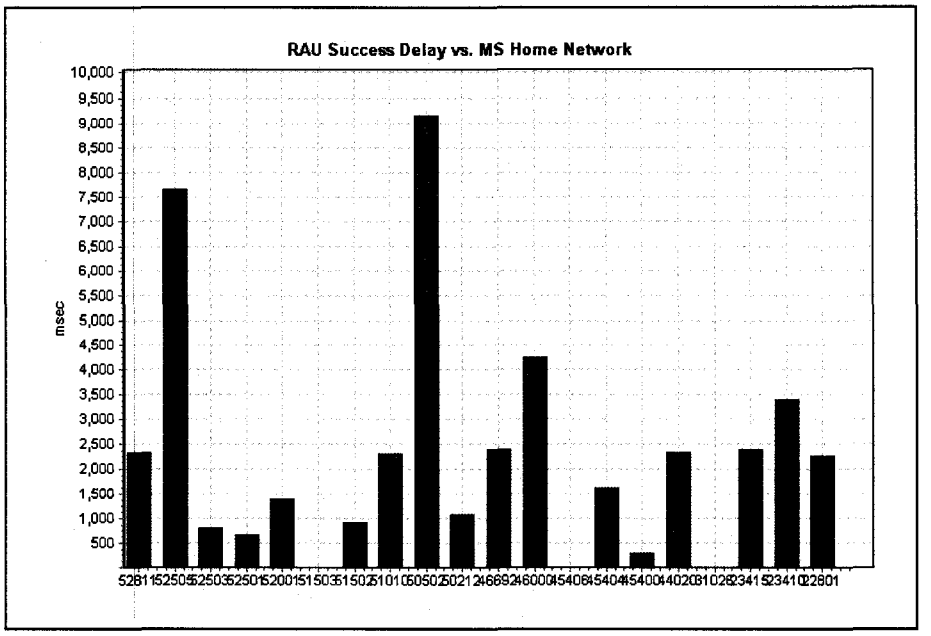

Figure 4.3 Délai de relève à partir du réseau d'origine

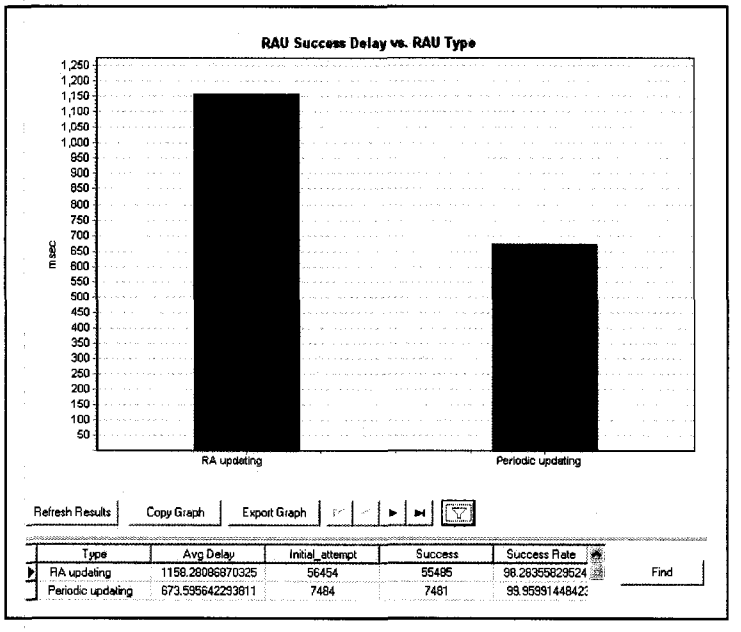

Figure 4.4 RAU dû à un déplacement et RAU dû à une mise à jour régulière

4.3.2 Implémentation du nouveau protocole sur *OPNET Modeler*^{TM}

Tel que mentionné auparavant, le protocole de GTP amélioré demeure un des constituants de l'algorithme défini auparavant et illustré à la Figure 4.1. De cet

algorithme, dérivons un mécanisme d'établissement de chemins LSP pour établir une preuve de concept d'amélioration du temps de réponse RAU. La présente section nous permet d'y décrire un tel modèle de réseau avec chemins pré-définis et non pré-établis. Le modèle considéré a été mis sur pieds à partir de l'outil *OPNET Modeler*^{TM} V10.5, avec une infrastructure Radio UMTS et un CN basé sur MPLS.

Modèle

Considérons maintenant un réseau UMTS tel que celui représenté à la Figure 4.5.

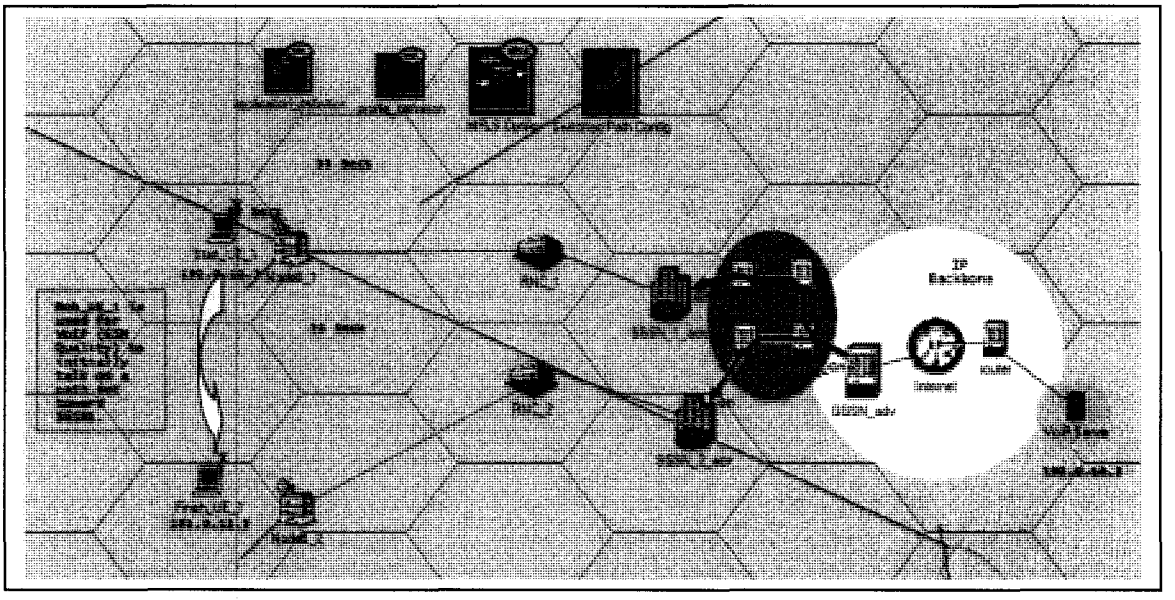

Figure 4.5 Modélisation d'un réseau UMTS avec deux chemins pré-définis

Nous pouvons ainsi remarquer que les chemins Bleu et Rouge sont représentés par des chemins pré-définis respectivement LSP1 et LSP2. Le LSP1 devient le chemin primaire avant la relève RAU, alors que le LSP2 représente le chemin alternatif après la relève. Nous observons donc à la Figure 4.5 que le mobile UE_1 se déplace de sa position *Start* (départ) jusqu'à sa position *Finish* (arrêt). Ce déplacement permet de mettre en œuvre un événement de relève RAU Inter-SGSN. La topologie du réseau d'accès en est une en arbre dont le tronc principal origine du GGSN ou encore du nœud en périphérie de la dorsale. Cette topologie est celle qui d'un point de vue

coût/performance demeure la plus utilisée par les opérateurs de réseaux cellulaires. Nous nous limiterons ainsi à utiliser un tel arrangement topologique. Tel que mentionné au Chapitre 3, des chemins prédéfinis LSP1 et LSP2 sont établis avant même que la requête de tunnelage soit initiée entre l'ancien SGSN et le nouveau SGSN. Ceci nous procure ainsi une façon plus rapide de répondre à une relève RAU. Les primitives 5.1 à 5.7 de la Figure 3.3 peuvent être traitées en parallèle avec les messages subséquents d'une requête RAU. Pour bien comprendre le modèle utilisé, nous allons tout d'abord décortiquer la séquence d'événements qui prend place lors d'une relève RAU Inter-SGSN. Nous réutiliserons à cet effet, les événements des primitives relatées à la Figure 3.3. Ces messages sont comme suit :

- Les étapes 1 et 2 représentent le déclenchement d'une relève Inter-SGSN ;
- Par la suite, la série de messages 3 et 4 viennent supporter l'accusé de réception du nouveau contexte et l'aspect de sécurité authentifié au serveur AAA ;
- Message 5.1[1] **Predef Tunnel_Aggr 2 Req** initié par le nouveau SGSN pour rediriger les flots de communications, lors de la relève ;
- Messages 5.2[1], 5.3[1] **Predef LSP Level 2 Req** représentant la requête d'établissement de chemin, sur le réseau cœur constitué de nœuds MPLS, dont les nœuds Ingress, LSR et Egress. De plus, une entité TE-CMS (Traffic Engineering-Configuration Management System) est également utilisée pour le calcul hors-réseaux, de chemins LSP pré-définis en utilisant par exemple une fonction de coût [16];
- Messages 5.4[1], 5.5[1], 5.6[1] et 5.7[1] **Predef LSP Level 2 Ack Predef LSP Level 2 Req** représentant l'accusé de réception de l'établissement de chemin, sur le réseau cœur dans la direction inverse jusqu'au nouveau SGSN ;

[1] Sujet à un repositionnement en fonction des résultats de validation *SPIN* obtenus.

- Les étapes 6, 7.1 et 7.2 sont les éléments réguliers d'annulation du chemin antérieur entre le réseau cœur et le GGSN ;
- Les étapes 8.1 et 8.2 décrivent les messages usuels de mise à jour des données d'usagers concernant sa nouvelle localisation qui est confirmée par l'étape 9 avec un **Update Location Ack** ;
- Finalement, les messages 10.1 et 10.2 terminent la complétion du RAU.

Le fait de pré-définir les LSPs 1 et 2 peut en venir à dire d'effectuer les étapes 5.1 jusqu'à 5.7 en parallèle à partir des étapes 1-4 qui représentent le déclenchement d'une relève RAU Inter-SGSN. Une fois ces LSP obtenus et réservés, nous pouvons remplacer les messages 5.1 – 5.7 par un message global d'Envoi-des-Paquets (Forward Packets). Ce processus concurrent résulte en un temps d'activation plus rapide de tunnel qui, dans ce cas-ci, sont de type LSP (par exemple LSP1 et LSP2). Cette première ébauche de la Figure 4.6 inclut tous les nœuds présents dans un réseau courant UMTS, prêts pour une relève RAU Inter-SGSN. En simplifiant davantage le modèle de ce réseau, nous obtenons la deuxième ébauche de la Figure 4.7.

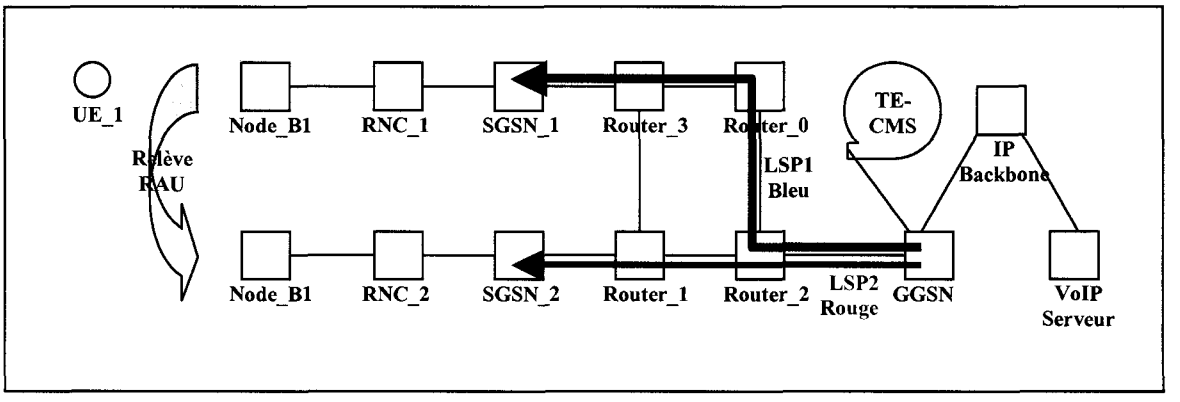

Figure 4.6 Réseau UMTS avec deux chemins pré-définis, première ébauche

Cette deuxième ébauche nous dirige ainsi vers un accès sans fil avec multi-chemins LSP1 et LSP2 respectivement, nous connectant en permanence au serveur de VoIP. En simplifiant encore davantage le tout, nous pouvons nous satisfaire du modèle de la Figure 4.8.

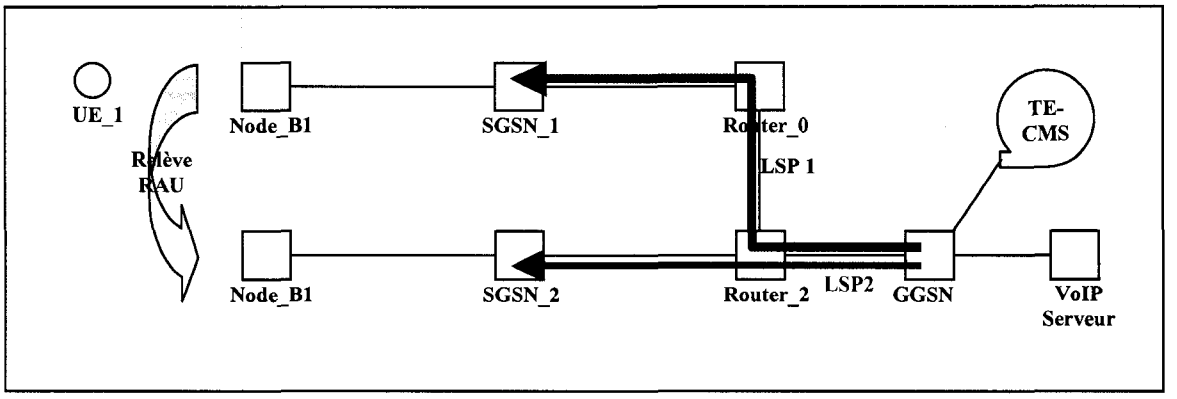

Figure 4.7 Modélisation du même réseau, deuxième ébauche

Nous remarquons donc que la Relève RAU ne représente ni plus ni moins qu'une commutation rapide entre les SGSN_1 et SGSN_2 si les LSPs 1 et 2 sont pré-définis et établis en parallèle à partir du déclenchement de requête RAU. Le TE-CMS demeure compatible avec l'algorithme présenté à la Section 4.2 et cherche à optimiser le réseau CN considéré.

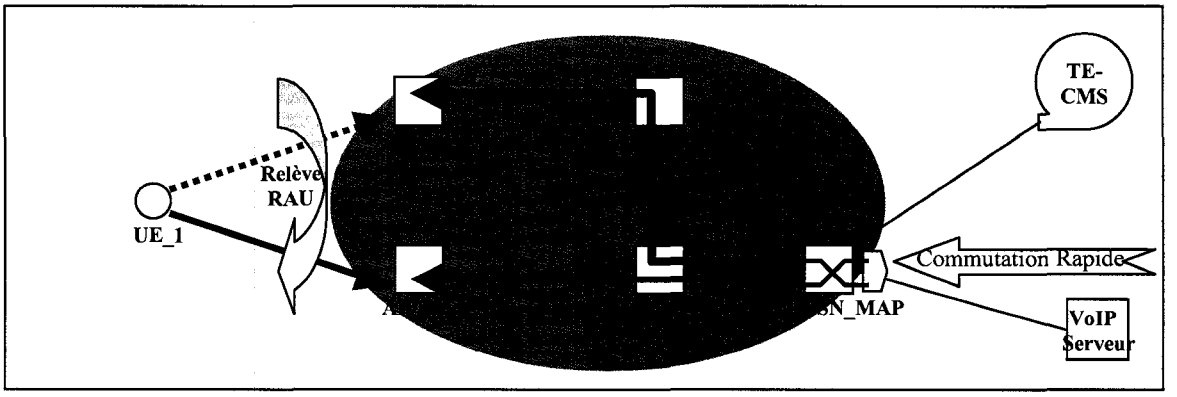

Figure 4.8 Modélisation d'un réseau UMTS, troisième ébauche

De plus, le GGSN (remplacé par un IGSN) sert d'élément actif à la commutation rapide entre les LSP 1 et 2. Le but de cette transition rapide est d'atteindre un temps de commutation de moins d'une seconde. Mais, la tendance des réseaux publiques est d'améliorer le protocole GTP pour changer de tunnels à l'intérieur d'une

demie seconde. Notre but est que nous puissions commuter rapidement à moins de 50 mSec (spécification Sonet qui est < 80 mSec de la relève GSM), ce qui serait dix fois plus rapide que la tendance actuelle proposée sous GTP, et supporterait ainsi une relève acceptable pour les trafics sensibles au délai. La voix sur IP représente un tel trafic.

Il est toutefois attendu de détecter un certain niveau de dégradation ayant la même durée que ce temps de commutation rapide, mais le trafic de voix survit assez facilement à une telle anomalie de 50 mSec par l'utilisation d'algorithme de correction, lors de pertes de paquets. Pour ce qui est de la vidéo-conférence, une méthode de masquage de trames détériorées par des trames prédites peut également aider considérablement à améliorer la perception de l'usager (exemple MPEG-2).

Dans notre cas, nous nous concentrerons sur un trafic de voix, pour plus précisément fournir une infrastructure appropriée, avec une fondation évolutive vers d'autres types de trafic.

Simulation

Tel que mentionné, notre recherche se base plus spécifiquement sur l'outil de simulation *OPNET Modeler*^{TM}. L'accès est celui radio d'UMTS W-CDMA R99. L'encodeur supporte Voix sur IP avec qualité basée sur GSM. Plus précisément, le trafic VoIP-GSM produira la caractéristique de flot présentée à la Figure 4.9 (le nombre de paquets par seconde, en ordonnée et le temps simulé, en abscisse). Nous pouvons remarquer que ce trafic effectue une pause à environ 48 secondes, ce qui est tout simplement une composante voulue dans le flot total de voix, en espérant déroger de la monotonie. Le flux total est de 100 paquets/sec ce qui correspond bien à un trafic de type G.729, par exemple, avec génération inter paquets tous les 10 msec. Dans notre cas test, nous induirons une relève RAU forcée au temps 70 secondes. La raison de ce choix est de considérer un point du régime permanent lors de la génération du trafic de voix qui soit considéré comme état stable, sans trop de variation, ce qui nous permet de mieux évaluer les délais encourus lors de la relève.

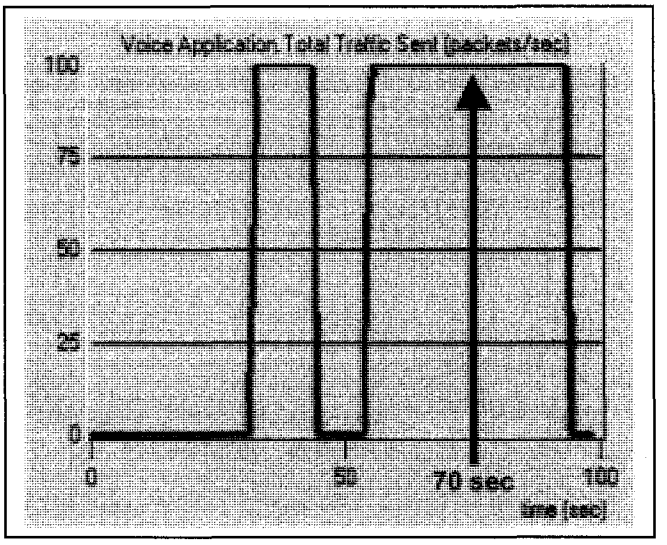

Figure 4.9 Trafic de voix VoIP-GSM

Le Réseau Cœur dépend des nœuds SGSN et GGSN en entrée et en sortie, qui supportent MPLS (LER). Les routeurs intermédiaires (routeurs 0, 1, 2 et 3), ne servent que de points de transit dans le réseau cœur et agissent comme LSR. Ils supportent donc aussi MPLS. Nous utilisons une configuration où deux nœuds séparés, soient SGSN1 et SGSN2, servent de points d'entrée dans le réseau cœur sur deux LSP complètement différents. Il est important de comprendre ici que chaque LSP demeure uni-directionnel. Nous nous contenterons de ne considérer que la direction de GGSN vers SGSN, puisque ce qui nous intéresse, se limite au temps de commutation d'un LSP à un autre lorsqu'ils sont pré-définis et établis concurentiellement durant le déclenchement initié lors d'une requête RAU Inter-SGSN.

Résultats

Avant de considérer les résultats obtenus, nous présenterons en tout premier lieu la performance optimale recherchée du protocole GTP, dans les installations publiques. Il est possible d'atteindre un temps de commutation d'environ 0.5 seconde (voir Figure 4.10) pour une relève RAU Inter-SGSN.

	Avec Relève, Paquets Commutés (estimé)
Changement de Cellule	< 0.4 seconde
RAU de type Intra SGSN	< 0.4 seconde
RAU de type Inter SGSN	< 0.5 seconde

Ref.: Ericsson CN-Evolution Phase 2, 2004

Figure 4.10 Performance de commutation avec GTP

Tel que mentionné auparavant, nous ciblons une performance de 50 mSec. Si nous considérons par contre la Figure 4.11 (le débit binaire par LSP, en ordonnée et le temps simulé, en abscisse) utilisant deux LSP statiques, soient LSP1 et LSP2, nous obtenons un temps de commutation de relève RAU Inter-SGSN de l'ordre de 1 sec.

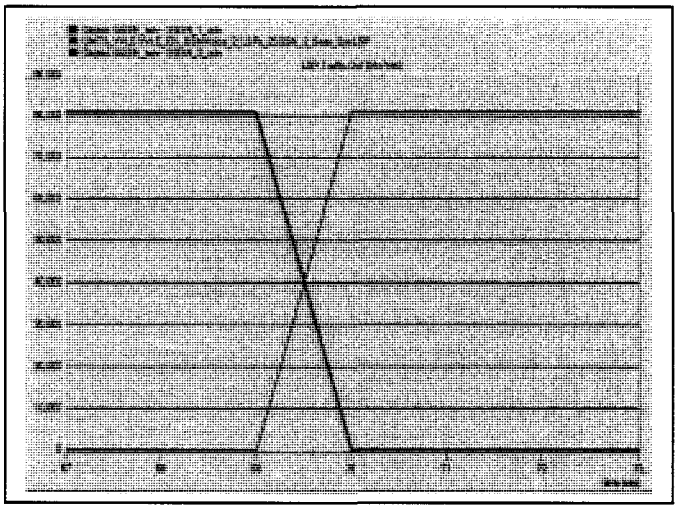

Figure 4.11 Performance de commutation avec GTP évolué

Mais, ce dernier délai représente le temps d'attente de commutation au GGSN entre les LSP1 et LSP2. Ce délai de commutation n'est pas vraiment le temps de commutation perçu au mobile UE_1 puisque, dans notre cas, deux LSP avaient été pré-

définis dans un passé récent, et qui avaient été établis concurentiellement au déclenchement d'une relève lors d'une requête RAU. Il devient clair qu'à un point donné, le mobile UE_1 opère à partir de deux chemins LSP sans interruption et que la relève radio prend place avec un recouvrement RF qui limite l'incidence d'un délai. Le résultat complémentaire obtenu à cet effet est celui de la Figure 4.12 qui démontre l'avantage d'un établissement de chemins concurentiellement à la requête de relève RAU.

Si l'on considère le sous-diagramme du haut à gauche (le nombre d'octets par seconde du trafic agrégé, en ordonnée et le temps simulé, en abscisse), nous voyons le trafic d'application de voix (G.729 par exemple) utilisé avec une pause aux environs de 48 secondes. Pour ce qui en est du diagramme du haut à droite (débit binaire relatif perçu au GGSN, en ordonnée et le temps simulé, en abscisse), celui-ci démontre qu'il y a commutation de chemins au GGSN, avec le délai obtenu (~ 1 sec) équivalent à celui de la Figure 4.11. Il devient intéressant d'observer le comportement du mobile UE_1 à partir du résultat du bas et à gauche (débit relatif perçu au mobile, en ordonnée et le temps simulé, en abscisse). La résultante est une interruption perçue nulle du point de vue UE_1 puisque, durant la relève, le mobile UE_1 reçoit au moyen de la propagation hertzienne, les deux chemins en même temps (à cause de recouvrement radio des cellules), ce qui permet au GGSN d'effectuer une commutation de chemins sans toutefois qu'une interruption quelconque prenne place (en utilisant un mécanisme approprié de combinaison de chemins). Finalement, le diagramme du bas et à droite (le délai d'insertion d'un LSP en seconde, en ordonnée et le temps simulé, en abscisse) nous montre qu'il y a même eu dans ce cas-ci, amélioration du délai de chemin lorsque nous passons du chemin LSP-1 au chemin LSP-2, ceci dépend toutefois de l'acheminement final considéré.

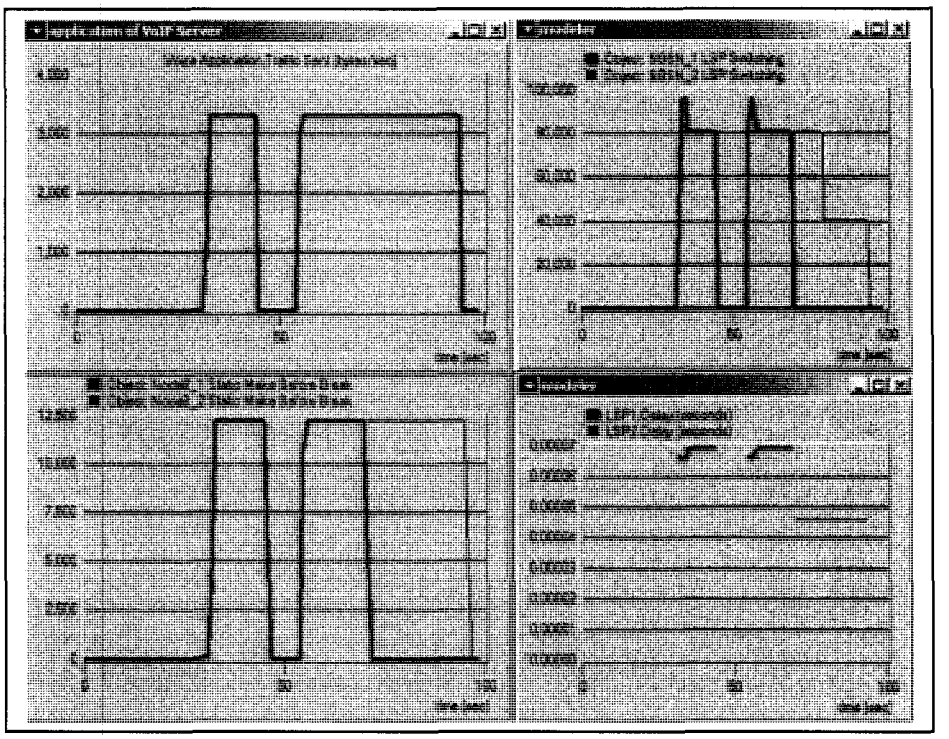

Figure 4.12 Perception de cette commutation au UE_1

Certains pourraient avancer que les chemins LSP dynamiques sont de meilleure augure pour des raisons d'évolutivité. Il est donc possible d'accepter un certain temps d'établissement de LSP, si nous avons recours à des LSP dynamiques. Ces derniers requièrent selon [17] moins de 25 msec de temps d'établissement. La modélisation de LSP dynamiques sera décrite plus loin dans cette sous-section.

Il demeure alors sans équivoque que nous avons dépassé nos attentes d'un facteur d'amélioration d'au moins 2 (< 25 msec si l'on utilise des LSP dynamiques dans le pire cas) comparativement au délai de 50 msec visé pour la commutation rapide, et d'un facteur d'amélioration de 20 comparativement au 500 msec visé par l'industrie à partir d'un tunnelage GTP évolué. Le meilleur résultat est donc celui obtenu par la solution basée sur des LSP statiques et dont le temps de commutation perçu (0 seconde) est représenté à la Figure 4.13 (débit relatif au mobile, en ordonnée et le temps simulé, en abscisse).

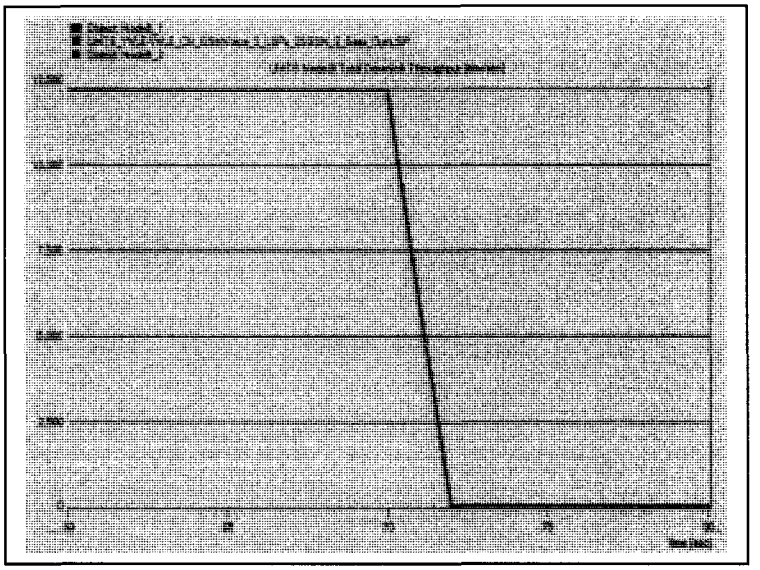

Figure 4.13 Temps de commutation perçu avec LSP statiques

Analyse des résultats

De ces résultats, nous pouvons conclure que deux alternatives s'offrent à nous, concernant la conception d'un protocole de GTP évolué (voir Figure 4.14).

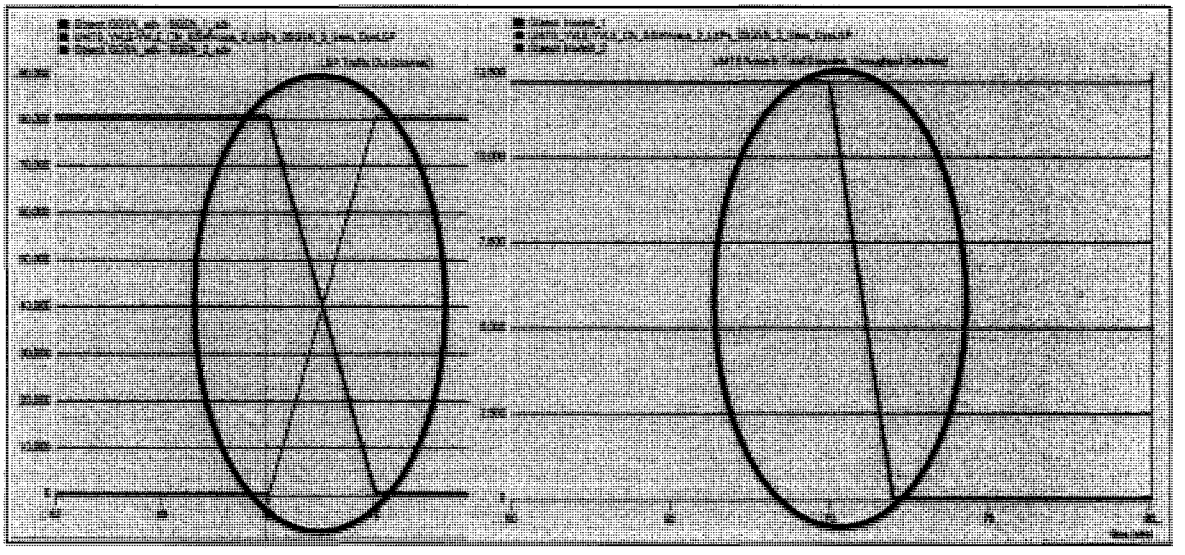

Figure 4.14 Compilation de résultats des modes SBBM et SMBB

Une première façon est de considérer une approche dite "*Static Break-before-Make*" ou encore SBBM qui résulte en un temps de commutation d'environ 1 seconde

(Figure 4.14, gauche avec débit relatif au GGSN, en ordonnée et le temps simulé, en abscisse).

Ce n'est que lorsque le chemin LSP1 est enlevé que le chemin LSP2 est activé avec un temps réponse prohibitif. Une deuxième façon est d'utiliser une approche "*Static Make-before-Break*" ou encore SMBB qui opère de telle manière que les chemins LSP1 et LSP2 sont tous deux établis avant même qu'une relève prenne place (Figure 4.14, droite avec débit relatif au GGSN, en ordonnée et le temps simulé, en abscisse). Il en résulte que, lorsqu'il y a relève d'un mobile UE d'une cellule à une autre cellule Inter-SGSN, le temps de réaction du réseau demeure pratiquement nul sous le mode d'opération SMBB. Il en est de même pour la section radio qui utilise déjà cette méthode pour obtenir une relève-douce (Soft-HandOff) et qui, selon la Figure 4.15, jouit d'au moins deux cellules de support avant qu'une relève ne prenne place.

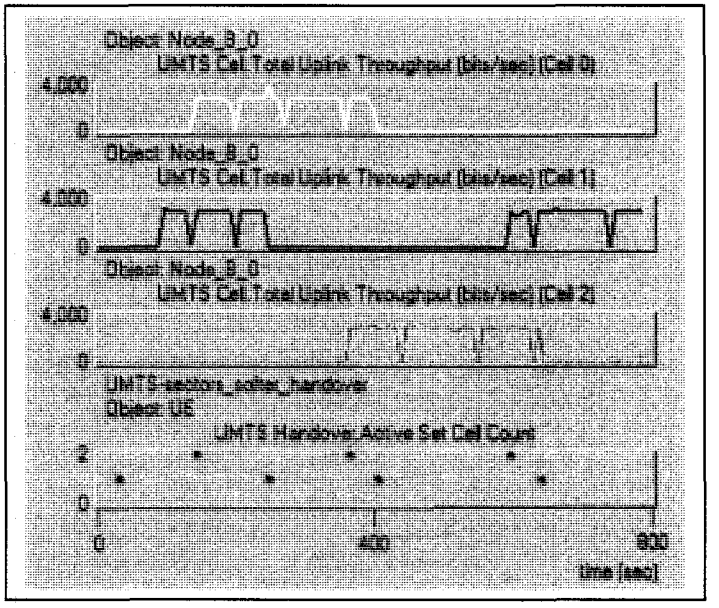

Figure 4.15 Mode SMBB déjà utilisé au niveau radio

Dans le cas précis de la Figure 4.15 (débit binaire en bits/sec relatif au Nœud-B et nombre de cellules actives, en ordonnée et le temps simulé, en abscisse), un mobile

UE passe d'une cellule sectorisée Brune, à une Jaune, à une Orange et de retour à une cellule sectorisée Brune.

Il est possible également d'opter pour une solution mitoyenne telle que celle du "*Dynamic Break Before Make*" ou encore DBBM basée sur au moins deux LSP dynamiques. Dans ce dernier cas, la commutation est plus rapide que celle du cas statique puisque la notification de changement de LSP est intégrée dans les rafraîchissements réguliers du protocole de signalisation de ressources appelé RSVP. Mais encore une fois, ce n'est que lorsque le chemin LSP1 est enlevé que le chemin LSP2 est activé, mais ici, avec un temps de réponse inférieur à 25 msec, ce qui cadre bien à l'intérieur du 50 msec recherché.

De plus, il demeure essentiel que, dans le cas du SMBB, qu'un module combinateur illustré à la Figure 4.8 soit utilisé à l'IGSN. Ce module est mieux représenté à la Figure 4.16. Il se compose d'une partie de séparation ou division et d'une partie de combinaison, d'où le nom de *Combinateur bi-directionnel.* Ce module combine dans une direction et divise dans l'autre direction. Il est donc bidirectionnel. Par-dessus tout, cet élément supporte la commutation de LSP, mais doit combiner avant de commuter. Cela explique comment il est possible d'obtenir un mode d'opération SMBB avec un temps de commutation quasiment nul. L'opération d'un *Combinateur bi-directionnel* peut être implémentée, dans un sous-système de commutation localisé au IGSN.

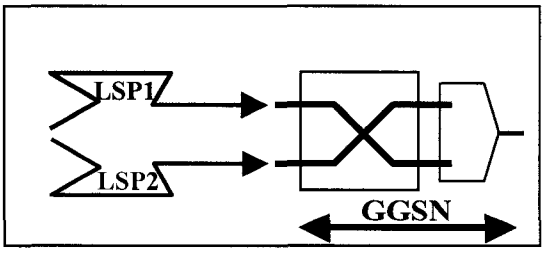

Figure 4.16 Utilisation du combinateur bi-directionnel pour le mode SMBB

En ce qui concerne le DBBM où des LSP dynamiques sont utilisés, nous avons effectué le même exercice que celui représenté à la Figure 4.5, mais cette fois-ci, en

utilisant des LSP dynamiques. Les configurations principales, reflétant le scénario des LSP dynamiques, consiste à initier l'application de voix à 150 secondes de manière à s'assurer que les LSP ont eu le temps d'être initialisés au départ (à environ **100 secondes**) lorsque le réseau complet a été mis en marche. Ces contraintes opérationnelles du modèle utilisé forcent une relève dynamique au temps de **220 secondes** pour permettre au trafic de voix de demeurer dans un régime stable durant une période de temps appréciable. Le résultat obtenu relatif au temps d'établissement des LSP dynamiques est représenté à la Figure 4.17 à gauche (débit relatif au GGSN, en ordonnée et le temps simulé, en abscisse) et à droite (le temps d'établissement de LSP, en ordonnée et le temps simulé, en abscisse).

Nous pouvons rapidement apprécier que l'établissement des LSP dynamiques aux temps de **100** et de **175 secondes** s'avère très acceptable. Le graphique de gauche nous montre le flot de trafic de voix passant du LSP1 au LSP2 au temps 220 secondes, alors que leur temps de pré-établissement était respectivement de **.00049 seconde** pour le LSP1 et de **.00036 seconde** pour le LSP2. On peut donc considérer un temps d'établissement de LSP dynamique typique inférieur à 1 msec.

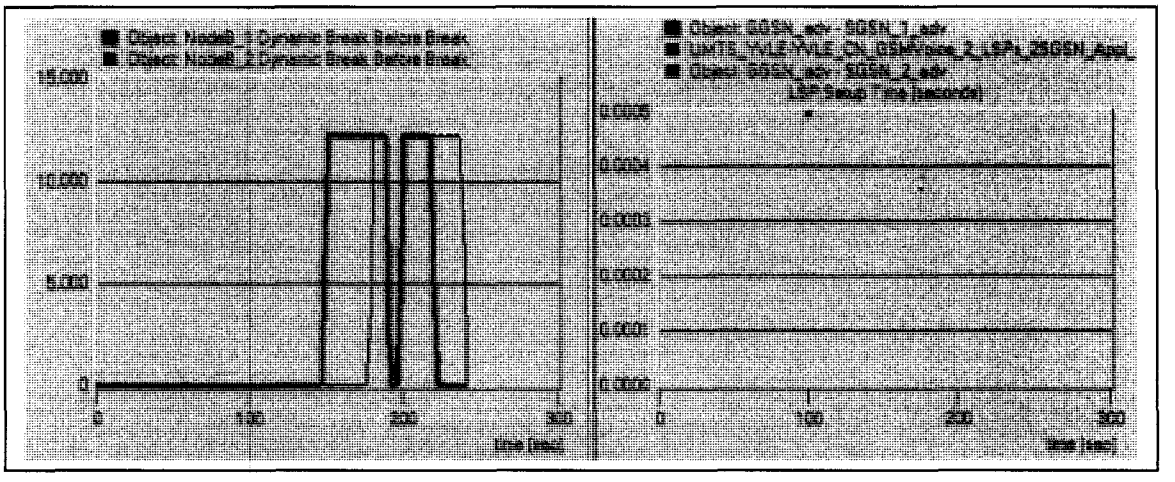

Figure 4.17 Temps d'établissement des LSP dynamiques

Si nous récapitulons ces résultats, nous obtenons donc dans le domaine réseau CN, les trois alternatives suivantes:

- En mode SBBM, un temps de relève réseau d'environ 1 seconde, sans combinateur ;
- En mode DBBM, un temps de relève réseau inférieur à 25 mSec, sans combinateur ;
- En mode SMBB, un temps de relève réseau quasiment nul, avec combinateur (qui représente la valeur ajoutée de cette recherche).

Finalement, il serait bon de considérer l'aspect de sécurité de ces chemins LSP1 et LSP2 lorsqu'ils sont établis, puisque nous essayons toujours de conserver le juste équilibre entre la QdS, la mobilité (dont GTP évolué fait partie), et la sécurité. Concernant cette dernière, si nous utilisons IPsec comme mécanisme de sécurité, il n'y a aucun moyen de savoir combien de temps un paquet d'information prendra pour atteindre sa destination, à moins qu'un chemin déterministe soit défini, d'où l'utilisation d'une technologie orientée connexion, telle celle de MPLS. En effet, IPsec répond au besoin d'authentification du paquet de l'envoyeur jusqu'au receveur, mais ne touche aucunement à l'aspect QdS relié à la Sécurité.

Pour cette raison, nous avons ici considéré l'utilisation de VPN de type couche 3 (L3) avec signalisation BGP pour installation d'agrégations. Il devient important d'identifier quels seront les points optimaux d'agrégations au niveau réseau cœur nous permettant ainsi de simplifier l'acheminement de paquets dans le réseau dorsal IP. En effet, la question demeure "Est-ce que la sécurité sur micro-flots sera maintenue partout dans le réseau cœur?" Il devient évident que, lorsque les flots de trafic atteindront le réseau dorsal IP, ils auront déjà été agrégés en classes de services à quelques points que ce soit, puisque que le réseau dorsal IP ne supporte pas des flots individuels pour des raisons d'évolutivité. La question ultime devient donc : "Où devrions-nous considérer agréger les flots de trafic en classes de services, à l'intérieur du réseau cœur? Puisque IPsec aura de la difficulté à procurer le niveau de QdS/Sécurité voulue pour les classes

agrégées, quel sera alors le mécanisme d'agrégation acceptable pour satisfaire ce duo de performance QdS-Sécurité?"

Nous proposons donc de définir, tel que mentionné auparavant dans la Section 3.2, une agrégation de classes de flots telle que celle présentée à la Figure 4.18.

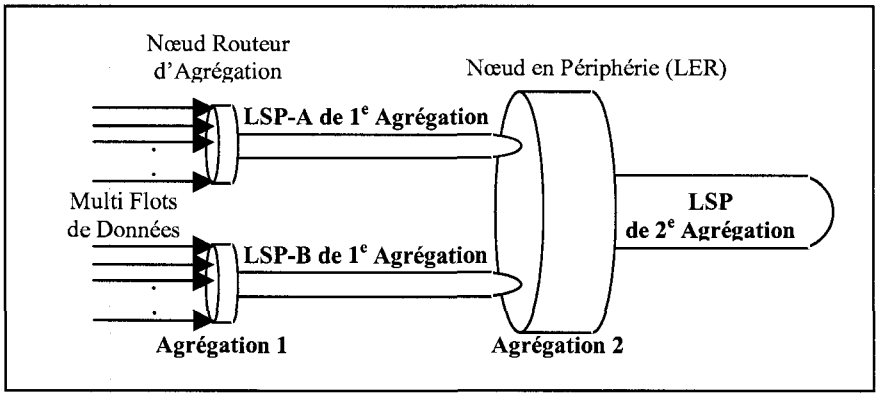

Figure 4.18 Deux niveaux d'agrégation des micro-flots avant la dorsale IP

Notre but demeure ainsi de construire une agrégation à deux niveaux avant de s'interconnecter au réseau dorsal IP, ce qui peut être décrit de la façon suivante :

- **Premier niveau d'agrégation** au premier routeur d'Accès ou regroupement des stations de Base, aussi appelées Noeuds B ;
- **Deuxième niveau d'agrégation** au IGSN ou regroupement des nœuds SGSN / GGSN.

Le premier niveau d'agrégation prend soin des micro-flots (uFlots) et de la micro-mobilité, alors que le deuxième niveau d'agrégation servira à la macro-mobilité et combinera à cet effet le trafic de *n* routeurs d'accès avant d'atteindre le réseau IP dorsal / multi-réseaux de transition, tout en se dirigeant vers une destination comme par exemple celle du "Aire de Desserte B" (voir Figure 3.1). Le protocole GTP-C lui-même peut être sensiblement remplacé par de la signalisation RSVP-TE utilisée pour établir les LSP de la technologie de MPLS. Ce protocole RSVP-TE peut à la rigueur évoluer pour permettre l'optimisation et l'adaptation à la mobilité.

Il est bon de noter ici que cette section n'abordera que la notion de base d'intégration VPN aux modes d'opération de types DBBM ou SMBB, alors que les sujets d'optimisation d'agrégats et d'adaptation à la mobilité demeureront un sujet futur de recherche.

Une des raisons principales pour l'adoption d'une architecture incluant VPN pour le transport sur dorsale IP est l'idée d'un duo de performance QdS-Sécurité ou encore triplet QdS-Sécurité-Mobilité. En effet, nous retrouvons également la problématique de rendre le mécanisme d'ancrage appelé APN (Access Point Name) plus flexible et moins dépendant de sa localisation géographique. Nous allons tout d'abord considérer ces deux points principaux qui sont celui du support du triplet QdS-Sécurité-Mobilité et celui de l'amélioration de la flexibilité du mécanisme d'ancrage APN. Référons-nous donc à la Figure 4.19, pour capturer l'essence même du contexte dans lequel nous opérons.

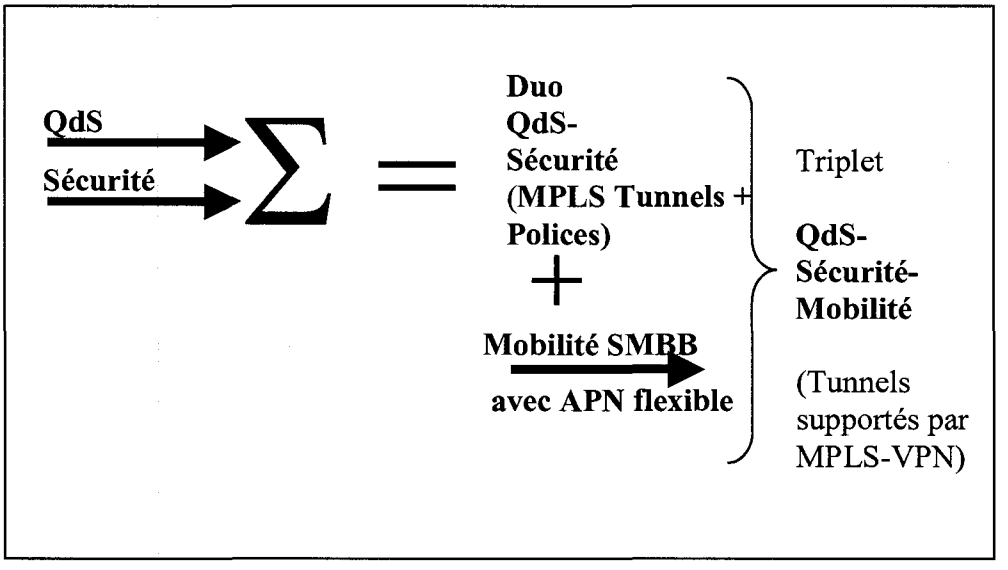

Figure 4.19 BGP-MPLS-VPN pour le support du triplet

La Figure 4.19 illustre notre préoccupation de conserver un juste équilibre entre la QdS et la sécurité, en établissant des tunnels établis par MPLS, et qui peuvent même être optimisés pour l'ingénierie de trafic. Une fois la sécurité d'agrégation de flots

obtenue (ce qui est différent d'IPsec, puisque les entités gouvernementales doivent être capables d'effectuer de l'intercepte légal), nous devons, en plus, y ajouter une touche de mobilité, idéalement SMBB ou encore DBBM avec un temps réponse de commutation négligeable. La mobilité qui est permise en mode SMBB se trouve couramment limitée à un ancrage APN qui permet à un mobile UE de savoir à quel réseau cœur il appartient. Donc, dans un réseau GPRS/UTMS, l'APN est une donnée de référence qui pointe au GGSN ou encore au routeur en bordure à être utilisé. Pratiquement, l'APN permet de décider à quelle adresse IP le tunnel GTP évolué est ancré. À la limite, l'APN peut, à partir du GGSN, identifier le réseau dorsal IP externe auquel le réseau cœur est attaché.

Tel que décrit auparavant, l'APN se compose de deux parties, soit l'Identificateur-Réseau externe et l'Identificateur-Opérateur du réseau d'accès. Ce dernier demeure donc une référence du réseau d'accès UMTS d'origine dans le cas de relèves de type RAU. Nous pouvons maintenant nous référer au Tableau 4.1 pour visualiser quelques exemples de référence APN orientés nom et spécifiques à des pays donnés, pour des réseaux existants donnés.

Tableau 4.1 Exemples d'APN

Canada	Fido	internet.fido.ca
Denmark	Orange	web.orange.dk
Norway	Netcom	internet.netcom.no
Sweden	Telia	online.telia.se
USA	Cingular	isp.cingular

Nous pouvons facilement discerner que, pour Telia par exemple, nous avons un APN spécifié comme étant online.telia.se, où *online* représente le service auquel le mobile UE est souscrit, *telia* représente le réseau d'Accès UMTS considéré, et finalement *se* représente le pays d'origine (dans ce cas-ci la Suède).

Il devient alors évident que l'APN est basé sur un schéma d'adressage orienté nom qui a une dépendance géographique, alors qu'un adressage direct IP nécessiterait moins d'intervention d'une fonction de traduction. En effet, l'adresse IP considérée serait locale à l'endroit où le mobile UE se trouve sans dépendance géographique. L'adressage IP direct nous permet donc de référer dynamiquement notre mobile UE (Route Optimization de IPv6) au réseau d'accès local où il se trouve (d'origine ou visité) sans jamais avoir à retourner au réseau d'origine, ce qui est différent du cas d'un mode d'adressage orienté nom.

Nous proposons donc de remplacer l'APN par un Route-Distinguisher (entité de 8 octets) de VPN qui nous référerait à un groupe d'utilisation privé. Non seulement avons-nous maintenant une agrégation tunnelée supportant la QdS et la Sécurité, mais nous y ajoutons une amélioration de mobilité indépendante géographiquement. De plus, l'adressage IPv6 nous procure les avantages qui sont l'acheminement dynamique et la localisation mobile qui est mise-à-jour dans une banque de données locale, en périphérie du réseau d'accès considéré.

La solution utilisant des tunnels BGP/MPLS VPN répond bien encore une fois aux besoins de QdS-Sécurité et de Mobilité. Une question demeure toujours: "Comment pouvons-nous intégrer l'algorithme d'optimisation des ressources avec le tunnelage par VPN et un mode d'opération SMBB ou DBBM pour la relève RAU?" Pour répondre à cette question, référons-nous à la Figure 4.20.

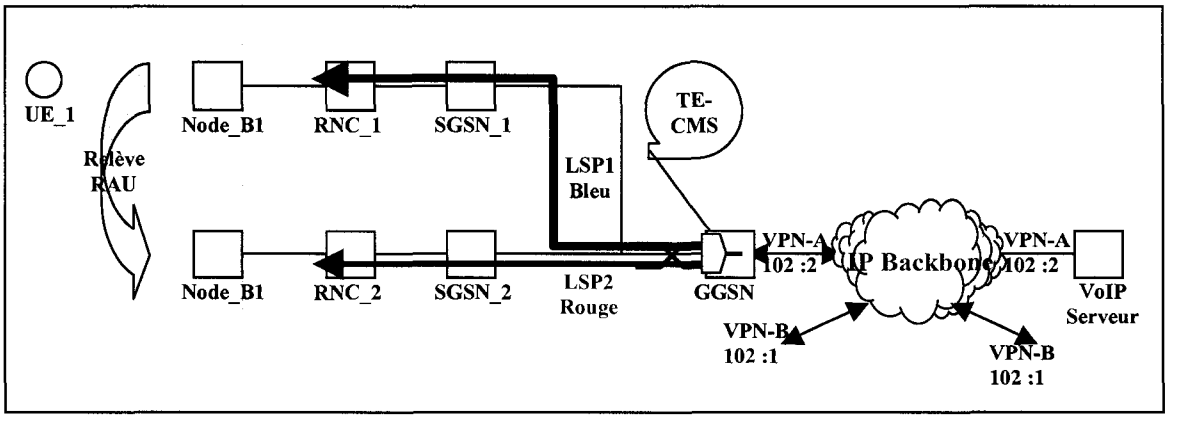

Figure 4.20 Vue d'une architecture supportant QdS-Sécurité-Mobilité

Dès le départ, il est possible de comprendre que le principe d'utiliser un mécanisme VPN au niveau de la dorsale-IP nous permet non seulement de différencier entre une pluralité de réseaux virtuels séparés, mais aussi de résoudre notre problème initial de référence APN qui est limité par une dépendance géographique. Cette résolution de problème se fait en remplaçant l'APN par un "Route-Distinguisher" plutôt que d'utiliser un pointeur de référence physique géographique. Ainsi, en utilisant par exemple le "Route-Distinguisher" **102:2** aux nœuds GGSN et Serveur VoIP, nous pouvons facilement localiser ces deux nœuds à des endroits différents, comme par exemple à Montréal, au Canada et à Stockholm, en Suède respectivement, sans confusion apportée concernant le réseau de bout-en-bout considéré.

Les modes d'opération SMBB ou DBBM ne changent pas et peuvent toujours être supportés par une infrastructure LSP-MPLS. Nous venons donc à la fois de répondre aux besoins de QdS-Sécurité combinés avec celui d'APN flexible.

4.3.3 Vérification des résultats en utilisant *SPIN*

Tel que mentionné auparavant, une troisième phase permettra de valider le nouveau protocole de migration, nous amenant ainsi vers un GTP évolué (sentiers LSP prédéfinis). Cette validation de protocole se fera en utilisant l'outil de vérification *SPIN* (Simple Promela INterpreter) [18]. Promela s'avère être un langage descriptif de systèmes opérant avec des processus concurrentiels, tel est le cas des protocoles de communications. En effet, ces derniers définissent la sémantique, la syntaxe et la séquence d'événements des messages de communication qui s'y rattachent.

Il est important de mentionner ici qu'une représentation modélisée sur *SPIN* d'un protocole quelconque implique nécessairement un certain niveau d'abstractions. L'aspect qui nous intéresse le plus est celui de l'implication de l'établissement de chemins LSP prédéfinis, en fonction des primitives déjà existantes lors d'une relève RAU. À cet effet, nous avons produit un modèle se

composant de trois mobiles UE_1, UE_2, UE_3 dans le but de s'assurer qu'il n'y a pas d'interaction de caractéristiques (feature interaction), de façon à concevoir un protocole robuste. Les nœuds considérés sont identifiés à la Figure 4.21.

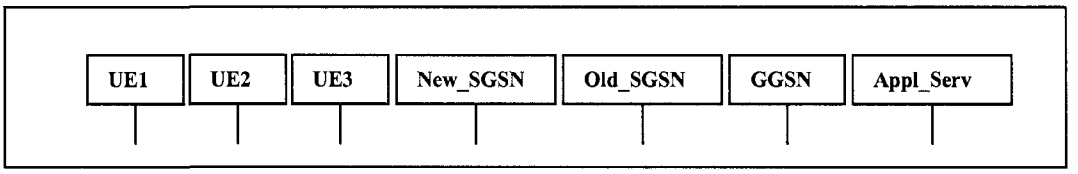

Figure 4.21 Nœuds considérés dans le modèle *SPIN* de relève RAU

Nous y voyons entre autres un nouveau SGSN (ou AR2), un ancien SGSN (ou AR1), un nœud en bordure de la dorsale IP, appelé GGSN, et finalement un serveur d'applications qui aura pour effet de consommer les paquets générés par les trois mobiles en question. Le tout nous permettra d'effectuer une relève RAU entre l'ancien SGSN et le nouveau SGSN, à partir d'un message RAU initié par un mobile quelconque. *SPIN* prend en charge d'y ajouter un effet stochastique qui ne nous permet pas de savoir d'avance quel mobile demandera une relève (Voir Figure 4.22 à cet effet).

On peut remarquer, à partir de la Figure 4.22, qu'une relève RAU prend bien place à partir d'un mobile UE1 (un mobile pour raison de simplification graphique). Le point d'intérêt sur lequel nous nous concentrons est de voir comment l'établissement de LSP peut être affecté par l'opération globale du système UMTS modélisé. Ceci nous permettra de voir d'emblée les limitations (s'il y en a) de la séquence d'événements du protocole considéré, et d'amener des corrections adéquates.

Le langage Promela, utilisé pour définir le système déjà mentionné, se retrouve en Annexe I, pour référence.

De ce modèle, nous obtenons un premier groupe de résultats qui nous démontrent le comportement du protocole proposé. Ces résultats sont représentés à

la Figure 4.23, qui décrit bien l'arrivée de la relève Iu_ps1!RAU initiée ici par le mobile UE1.

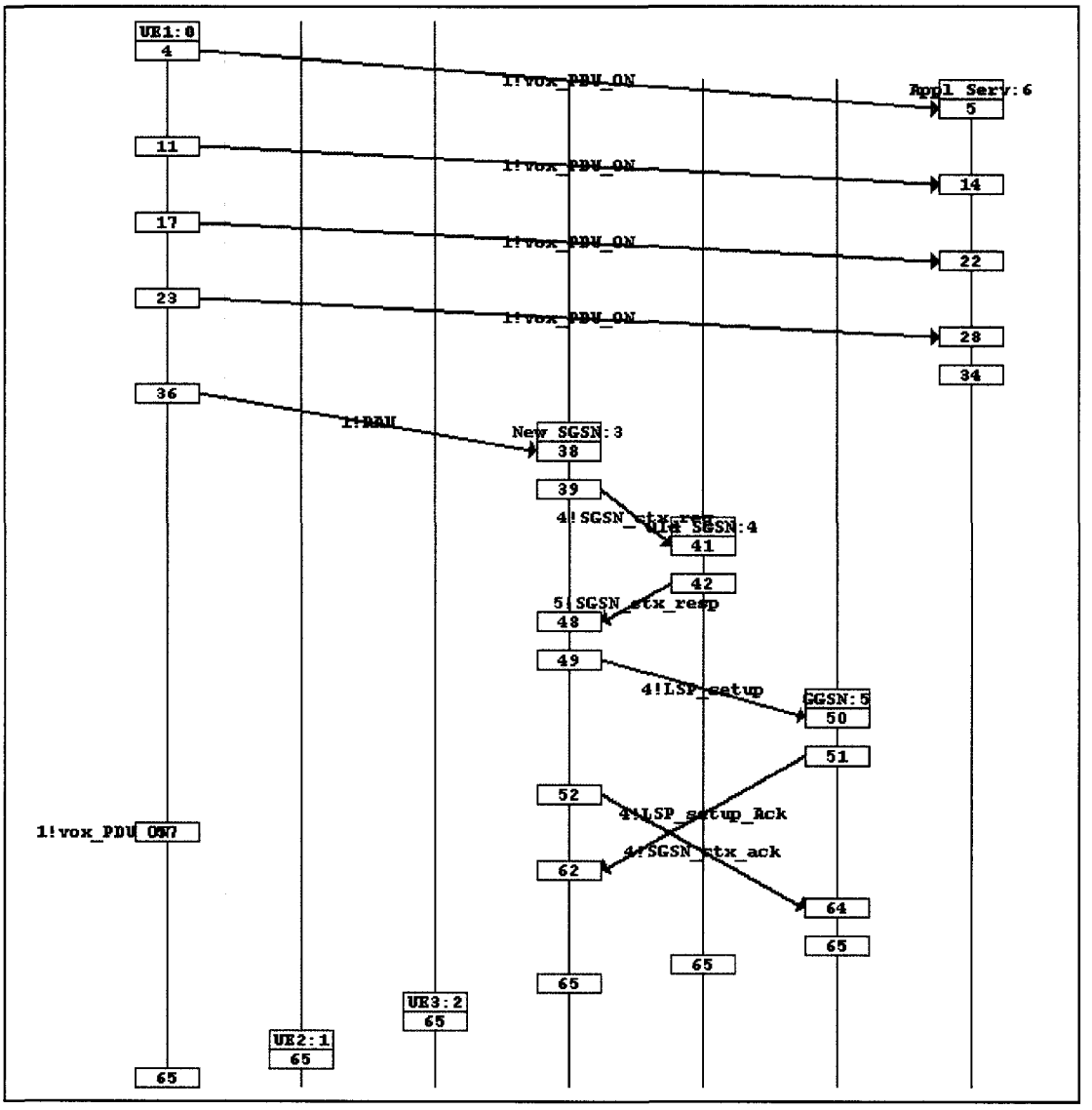

Figure 4.22 Modèle *SPIN* avec demande de relève RAU initiée par un seul UE

Tout se déroule bien de la façon suivante:

- Transmission de paquets de voix — Iu_ps3!vox_PDU_ON
- Réception d'un paquet de voix — Iu_ps3?vox_PDU_ON
- Transmission de paquets de voix — Iu_ps2!vox_PDU_ON

- ..... autres vox_PDU_ON envoyés et reçus
- Réception d'un paquet de voix Iu_ps2?vox_PDU_ON
- Mobile UE1 initie une Relève RAU **Iu_ps1!RAU**
- Le nouveau SGSN reçoit ce RAU Iu_ps1?RAU
- Mobile UE2 initie une Relève RAU Iu_ps2!RAU
- Transmission de paquets de voix Iu_ps3!vox_PDU_ON
- Réception d'un paquet de voix Iu_ps3?vox_PDU_ON

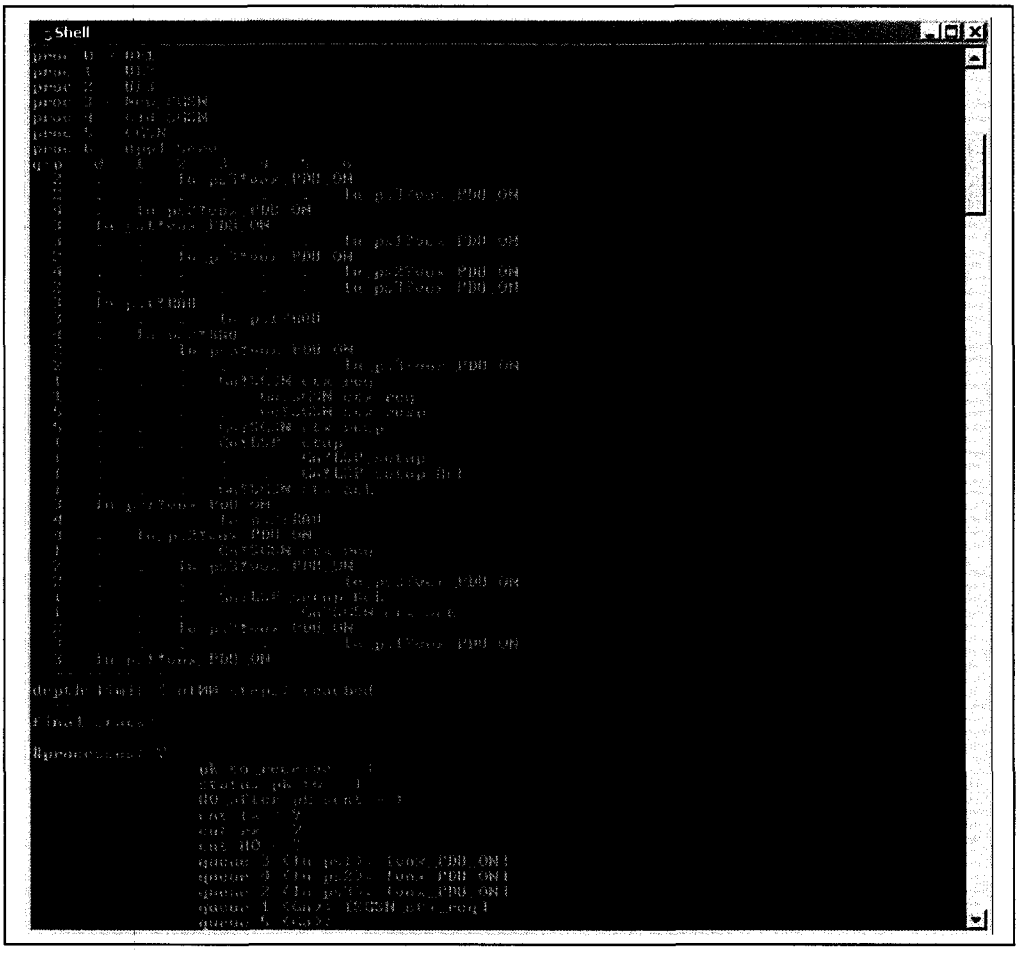

Figure 4.23 Résultats spin407 -c -u100 sur Handover_GTP_R99_V18_wFAULT_3_UEs.pml

- Le nouveau SGSN envoie Gn!SGSN_ctx_req

- L'ancien SGSN reçoit Gn?SGSN_ctx_req
- L'ancien SGSN envoie Go!SGSN_ctx_resp
- Le nouveau SGSN reçoit Go?SGSN_ctx_resp
- Le n-SGSN demande un LSP Gn!LSP_setup
- Le GGSN reçoit le Gn?LSP_setup
- Le GGSN envoie un Gn!LSP_setup_Ack
- Le n- SGSN envoie un ctx_ACK Gn!SGSN_ctx_ack
- Le n- SGSN reçoit un LSP_ACK Gn?LSP_setup_Ack[2]
- ..... avec requêtes RAU enchevêtrées!

Tout se passe bien selon la Figure 4.23 jusqu'à maintenant, avec une progression donnée de la séquence de messages du protocole recommandé, mais cette progression n'est pas nécessairement garantie en tout temps. Maintenant, nous effectuons une demande d'exécution pour vérifier si le graphe d'exécution rencontre un ou des états erronés. Selon la Figure 4.24, nous obtenons un état final erroné à l'étape 35.

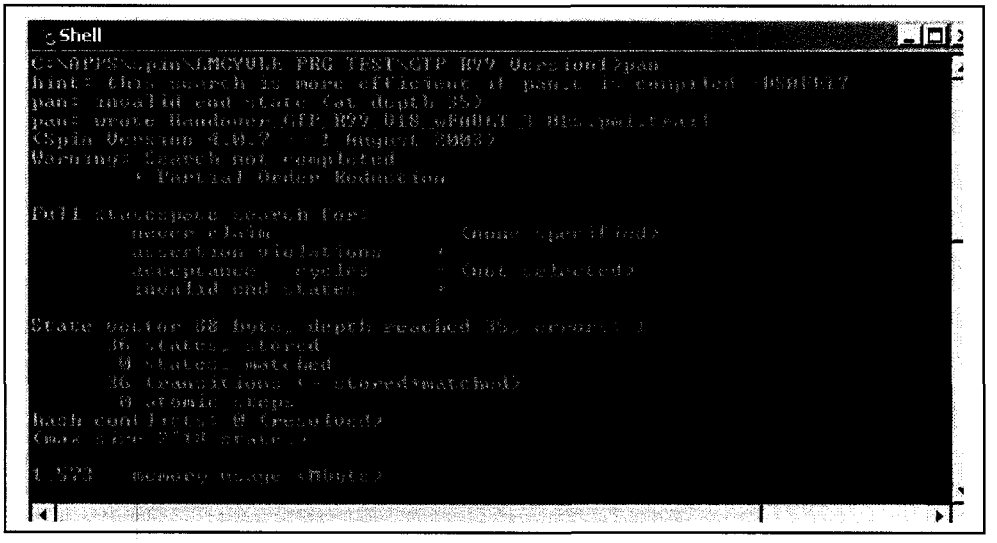

Figure 4.24 Résultats pan sur Handover_..._3_UEs.pml

[2] LSP_setup_Ack est reçu avant que le SGSN_ctx_ack soit reçu, ce qui est bien.

Nous devons alors effectuer une exécution-dirigée (trace) avec un fichier trace *.pml.trail, qui nous permettra d'identifier le problème. En exécutant la commande spin407 -t –c ..., nous obtenons le résultat de la Figure 4.25. Il en résulte que trois paquets de voix sont transmis mais gardés en suspens et non complétés, tout comme le paquet de voix de la Figure 4.22 qui demeure en suspens lors de la relève RAU, pour un système simplifié (pour des raisons de clarté graphique), et utilisant un seul mobile UE_1. Ce que *SPIN* nous identifie, c'est qu'un système de file d'attente doit à tout prix être utilisé pour conserver ces paquets lors de relève RAU, sinon ils n'effectuent pas de transitions vers un nœud destinataire. Cet état non complété est découvert à partir du graphe d'exécution établi par *SPIN*, lors du déroulement de la simulation du modèle de la relève RAU. Cela peut nous amener aussi à établir une horloge (Timer = t_1) d'expiration qui dicte au système que tout paquet non consommé après un certain temps doit être éliminé. Ce premier test est appelé test de blocage d'état.

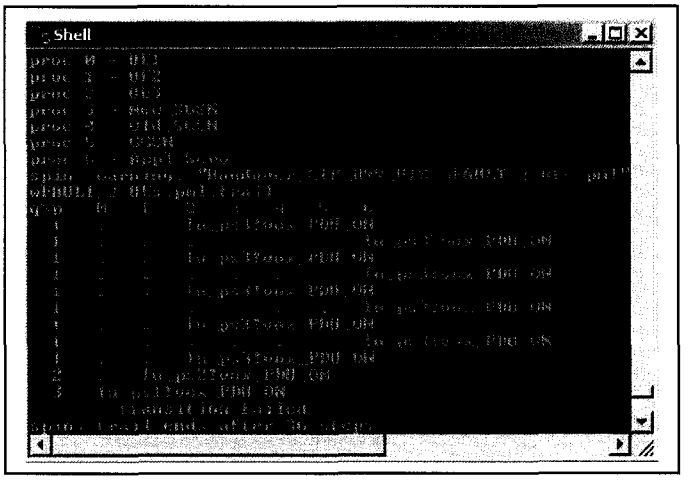

Figure 4.25 Résultats spin407 -t –c

Finalement, nous allons exécuter une revendication interdite (Never Claim) pour déterminer si un message "SGSN_ctx_ack" peut, dans certains cas, arriver avant un message "LSP_setup_Ack". À cet effet, nous ajoutons donc une

revendication improbable au modèle (voir Annexe II pour plus de détails). Cette revendication improbable effectuée sur le modèle UMTS dicte qu'il n'existe **aucune possibilité** (sinon, nous avons un problème) que tous les états suivants arrivent séquentiellement (où la réception d'un SGSN_ctx_ack prends place avant celle d'un LSP_setup_Ack) soit :

- Un vox_PDU_ON est reçu
- Un SGSN_ctx_ack est reçu
- Un vox_PDU_ON est reçu
- Un vox_PDU_ON est reçu
- Un LSP_setup_Ack est reçu

Il est important de comprendre que ces états de la revendication interdite peuvent arriver de façon aléatoire. Pour simuler cette revendication improbable de "SGSN_ctx_ack" arrivant avant "LSP_setup_Ack", nous devons exécuter les commandes suivantes :

- **spin407 -a -PCL -E/E** effectué sur la filière modèle *.pml pour initier une simulation ;
- **cl -DBFS -o pan pan.c** dans le but d'effectuer une recherche d'erreur d'états et l'enregistrer dans une filière source appelée pan.c ;
- **pan** pour visualiser la filière pan.c ;
- **spin407 -t –c** sur la filière modèle *.pml pour guider l'exécution de la simulation selon la trace générée par "pan" .

Les résultats obtenus furent :
"pan: claim violated! (at depth 31)
pan: wrote Handover_GTP_R99_V18_wFAULT_3_UEs_FINAL.pml.trail"

qui est la trace résultante et, à partir de cette trace, nous obtenons les états reflétant la revendication peu probable, mais possible de l'Annexe II, qui s'attend à ce que

plusieurs paquets voix soient envoyés avant la réception des messages "SGSN_ctx_ack" et "LSP_setup_Ack". La question demeure : Est-il possible d'arriver à une condition erronée où le message de Contexte de QdS appelé "SGSN_ctx_ack" soit reçu avant le message "LSP_setup_Ack"? Si nous considérons la trace d'exécution obtenue à la Figure 4.26, nous constatons rapidement que cette condition, bien que peu probable, est possible. C'est ce qui fait la beauté du vérificateur de modèle qui n'est pas limité qu'à des états probables, mais considère plutôt tous les états possibles, tels que les états suivants :

-	Transmission de paquets de voix	Iu_ps3!vox_PDU_ON
-	Réception d'un paquet de voix	Iu_ps3?vox_PDU_ON
-	Transmission de paquets de voix	Iu_ps3!vox_PDU_ON
-	Réception d'un paquet de voix	Iu_ps3?vox_PDU_ON
-	Mobile UE3 initie une Relève RAU	Iu_ps3!RAU
-	Le nouveau SGSN reçoit ce RAU	Iu_ps3?RAU
-	Le nouveau SGSN envoie	Gn!SGSN_ctx_req
-	L'ancien SGSN reçoit	Gn?SGSN_ctx_req
-	L'ancien SGSN envoie	Go!SGSN_ctx_resp
-	Le nouveau SGSN reçoit	Go?SGSN_ctx_resp
-	Le n-SGSN demande un LSP	Gn!LSP_setup
-	Le GGSN reçoit le	Gn?LSP_setup
-	Le n- SGSN envoie un ctx_ACK	Gn!SGSN_ctx_ack
-	Le GGSN envoie un	Gn!LSP_setup_Ack
-	Le GGSN reçoit un ctx_ACK	Gn?SGSN_ctx_ack[3]
-	Le n- SGSN reçoit un LSP_ACK	Gn?LSP_setup_Ack

spin: trail ends after 31 steps

Nous concluons de cette analyse de sûreté, deux points principaux :

[3] SGSN_ctx_ack est reçu avant que le LSP_setup_Ack soit reçu, ce qui est un problème.

- les paquets de voix en suspens doivent être soit mis en file d'attente, soit rejetés après un certain temps (minuterie) ;
- l'établissement d'un LSP doit prendre place avant le message de "context_ack" (test effectué à partir de 3 mobiles UE), ce qui nous force à ramener les étapes 5.1 – 5.7 de la Figure 3.3 à une exécution parallèle avec le déclenchement RAU.

Figure 4.26 Résultats *pan* sur Handover_GTP_R99_V18_wFAULT_3_UEs.pml

4.3.4 Calcul de ressources requises

Cette section, décrit le mode de calcul considéré pour le dimensionnement des trafics de voix et de vidéo-conférence, dans le cadre d'un accès UMTS. Il est important

de noter que chaque LSP sert de chemin d'agrégation pouvant supporter plusieurs flots de trafic, pour une classe donnée (par exemple voix). Ainsi, pour transporter une capacité équivalente de 300 flots de trafic à débit binaire de 8 Kbps chacun (G.729, avec temps d'inter arrivé de paquets de 10 msec), nous devons alors dimensionner ce LSP avec une caractéristique de débit binaire maximum approprié. Nous allons premièrement procéder à la notation analytique de ce dimensionnement pour nous permettre de bien identifier les paramètres utilisés.

Notations

Voici la notation choisie dans le but de décrire mathématiquement le dimensionnement de trafics de voix et de vidéo-conférence :

Erl	Erlangs
Apls	Appels
Min	Minute
Erl / LSP Agr	Erlangs par LSP Agrégés
Bl	Blocage
Can	Nombre de canaux
i	indice
U	Utilisation en bits par seconde (bps)
N	Nombre Total de Flots α_i par LSP
α_i	Largeur de Bande Effective du i^{è} Flot de vidéo-conf.
Cap	Capacité totale d'un LSP

Traffic de Voix

Considérons premièrement le trafic de voix et prenons comme hypothèse que nous opérons dans un grand centre métropolitain consistant en une population d'environ 2.3 millions d'habitants. Avec une pénétration cellulaire mobile de 60% (pire

cas an Amérique du Nord), nous nous retrouvons avec une population de 1.4 millions d'habitants qui possèdent un mobile quelconque pour communication 3G d'UMTS.

Un appel normal durant en moyenne 2 minutes, nous espérons soutenir dans le pire cas 840,000 appels (plus de 60% des mobiles disponibles) simultanés de 2 minutes chacun par pire heure occupée, ou encore 100 agrégations de 8,400 appels de 2 minutes, ce qui résulte en un volume de trafic de :

$$\text{Erl} = \frac{8{,}400 Apls \times 2 Min}{60 Min} = 280 \text{Erl / LSP Agr} \qquad (1)$$

Considérons maintenant une probabilité de blocage acceptable d'appel de 2%. Dans le but de rencontrer ce niveau de blocage, nous nous conformons aux tables d'Erlangs satisfaisant l'équation :

$$\text{Bl} = \frac{(Erl)^{Can}}{Can!} \Big/ \sum_{i=0}^{Can} \frac{(Erl)^{i}}{i!} \qquad (2)$$

Nous devons donc supporter 100 LSPs d'agrégations de 300 flots de 8 Kbps chacun, ou encore 100 LSPs de 2.4 Mbps. Ceci nous amène à considérer plutôt 2.5 Mbps de débit binaire maximum. Ce résultat final de **100 LSPs x 2.5 Mbps** ne représente aucun obstacle pour les nœuds routeurs pouvant commuter facilement 1000 LSPs avec agrégation totale de 20 Gbps (exemple M20 de Juniper). Le dimensionnement requis est possible dans la pratique et permet typiquement des capacités de 1 méga paquet par seconde.

Traffic de vidéo-conférence

Comme première approximation, le concept de largeur-de-bande efficace devient important puisque nous faisons ici l'hypothèse d'un type de trafic vidéo-conférence qui démontre des caractéristiques de dépendance à long terme (auto-similaire). Considérons dans ce cas-ci que 10% de la population de ceux qui utilisent un

mobile UMTS (840,000 usagers) peuvent à la rigueur simultanément utiliser une communication vidéo-conférence durant la pire heure. Ceci représente donc 84,000 flots de 90 Kbps. En effet, selon [19], le codec vidéo de type H.263 basé sur MPEG 2 produit un faible taux binaire moyen d'environ **90 Kbps** tirés de ses caractéristiques illustrées au Tableau 4.2.

Le trafic H.263 donne une taille moyenne de paquets de 1028 octets avec une variation auto-similaire accentuée située entre 921 (-107) et 1074 (+46) octets par paquet [19]. Nous faisons l'hypothèse que l'arrivée des flots vidéos est indépendante, selon [14], les unes des autres. Ceci nous aidera à conserver un certain gain statistique de multiplexage. Une autre hypothèse est que chaque nœud du réseau rejette les paquets qui dépassent une certaine garantie de délai, ce qui est acceptable puisque MPEG2 possède certains mécanisme de masquage de perte de paquets.

Tableau 4.2 Caractéristiques du trafic vidéo de type H.263

Source de trafic	**Vidéo (ex : H.263)**
Taille moyenne des paquets IP (octets)	1028
Taux (paquets par seconde)	23
Paramètre de Hurst	0.87
Auto-similarité du trafic généré	OUI

Des 84,000 flots de 90 Kbps considérés auparavant, nous devons effectuer une analyse à partir de largeur de bande effective, pour calculer le besoin total de ressources réparties sur 100 LSPs (séparés des 100 LSPs de voix). Nous obtenons donc 840 flots de 90 Kbps par LSP. Vérifions que le LSP de capacité *Cap* est capable de procurer le support de N flots α_i (où $N = 840$), en utilisant la formule suivante [14]:

$$U = \sum_{i=1}^{N} \alpha_i \qquad \text{où } U < Cap \tag{3}$$

840 x 90 Kbps = 75,600 Kbps = 75.6 Mbps par LSP (4)

75.6 Mbps par LSP < 100 Mbps → $\text{Capacité}_{\text{Max}}$ par LSP (5)

100 LSPs x 100 Mbps = 10 Gbps inférieur au 20 Gbps du Routeur (6)

Il est intéressant de voir ici que l'agrégation de 840 flots correspond à environ un LSP d'environ ~ 76 Mbps < 100 Mbps. Cette inégalité est vraie si et seulement si (selon [14]) Pr $(B > x) \sim e^{-sx}$ lorsque x → ∞, où B est le régime normal de retard de trafic. En d'autres mots, tant et aussi longtemps que la largeur de bande effective d'un groupe de flots demeure inférieure à la capacité totale *Cap* d'un chemin, la probabilité de perte de paquets résultant d'une congestion de file d'attente diminue exponentiellement en fonction de la profondeur de la mémoire tampon.

De plus, selon les références [14][20], la somme des arrivées d'un agrégat de flots FBM (Fractional Brownian Motion), originant de sources indépendantes auto-similaires reflétant un paramètre Hurst commun, produit un agrégat qui est lui aussi de type auto-similaire.

Finalement, plus les débits binaires de crête sont petits par rapport au débit binaire d'agrégation de LSP, plus les trafics auto-similaires se rapprochent du débit moyen avec un gain de multiplexage statistique acceptable. Ainsi, selon [14], une agrégation de 1000 flots (dans notre cas 840 flots) de trafic auto-similaire sur 100 Mbps se rapproche de son débit binaire moyen lorsque sa valeur de crête inférieure à 400 Kbps comparativement à son agrégat de 100 Mbps. Ces valeurs justifient donc la méthode de calcul utilisée, qui est basée sur des flots vidéo moyen de 90 Kbps chacun, avec valeur de crête possible de ~300 Kbps (31 pps x 1074 octets par paquet x 8 bit/octet).

CHAPITRE V
CONCLUSION

Nous avons donc vu que les réseaux sans fil contemporains se dirigent de plus en plus vers des services Internet Multimédia. Ces derniers supportent davantage des ressources en temps réel et une mobilité accrue. Ce mémoire s'est concentré sur le protocole de tunnelage GTP qui répond peu aux contraintes de délai encouru. GTP utilise IP pour transport, nous forçant vers une double encapsulation IP/GTP/IP. Nous avons donc été amenés à considérer un chemin évolutif de GTP, pour diminuer le temps élevé de relève radio, satisfaisant ainsi le trafic de voix. La méthode consista à utiliser des mesures dans le champ pour identifier le délai courant obtenu. Elle a permis d'améliorer le temps de réponse de GTP en y ajoutant un algorithme de chemins pré-définis. Nous avons eu recours à des chemins à étiquettes commutées pour se comparer à un temps de relève de moins d'une demi-seconde.

5.1 Synthèse des travaux

À l'intérieur de cette recherche nous avons abordé les points principaux suivants, qui sont la mesure du temps de relève RAU typique entre deux RA (Routing Areas), une solution d'amélioration du temps de relève en ayant recours à des chemins à étiquettes commutées (< ½ seconde), la validation de protocole à partir de l'outil de vérification *SPIN* pour déterminer s'il y a occurrence d'états problématiques, une démarche mathématique simplifiée pour le dimensionnement des trafics de voix et de vidéo-conférence, dans le but de s'assurer qu'une telle implémentation est réaliste avec les aiguilleurs (routeurs) couramment disponibles, et en tout dernier lieu une proposition de l'intégration de fonctionnalité pour obtenir un triplet opérationnel de QdS-Sécurité-Mobilité qui soit balancé.

Pour ce qui est de la mesure du temps de relève RAU typique, ce temps s'avère être de l'ordre de 2 à 10 secondes, ce qui est inconcevable pour les trafics sensibles au

délai. Nous avons par la suite décrit une solution d'amélioration du temps de relève en ayant recours à des chemins à étiquettes commutées avec les alternatives suivantes :

- en mode SBBM, un temps de relève réseau d'environ 1 seconde[1] ;
- en mode DBBM, un temps de relève réseau inférieur à .025 seconde[1] [17] ;
- en mode SMBB, un temps de relève réseau quasiment nul[2].

Il est recommandé d'utiliser le mode DBBM pour les relèves qui requièrent un grand nombre de chemins LSP, comme dans le cas d'agrégation de canaux de voix, et ce, dans le but de conserver l'aspect d'évolutivité. Le temps de pré-établissement de LSP dynamiques obtenu était inférieur à 1 msec. Il est toutefois préférable d'utiliser le mode SMBB pour les trafics hautement critiques comme les appels d'urgence 911.

Concernant la validation de protocole à partir de l'outil de vérification *SPIN*, nous en avons conclu que les paquets de voix en suspens doivent être soit mis en file d'attente, soit rejetés après un certain temps (minuterie) et que l'établissement d'un LSP doit prendre place avant le message de "context_ack" (test effectué à partir de 3 mobiles UE), ce qui nous force à ramener les étapes 5.1 – 5.7 de la Figure 3.3 à une exécution parallèle avec le déclenchement RAU.

Par la suite, une démarche mathématique simplifiée pour le dimensionnement de trafic se résume à 100 LSPs x 100 Mbps = 10 Gbps pour la vidéo-conférence (100 LSPs x 2.5 Mbps = 250 Mbps pour la voix) qui demeure bien inférieur au 20 Gbps de capacité totale d'un routeur typique.

En tout dernier lieu, une proposition de l'intégration de fonctionnalité pour obtenir un triplet opérationnel de QdS-Sécurité-Mobilité balancé est possible en utilisant des tunnels BGP/MPLS VPN (pour GTP évolué) qui répond bien encore une fois aux besoins de ce triplet opérationnel. Ceci permet de remplacer potentiellement l'APN par un Route-Distinguisher (entité de 8 octets) de VPN qui nous référerait à un groupe d'utilisation privé, sans référence requise au point d'attache d'origine. De plus, l'adressage IPv6 nous procure des avantages additionnels qui sont l'acheminement

[1] Sans combinateur

[2] Avec combinateur

dynamique, la localisation mobile qui est mise-à-jour dans une banque de données locale, en périphérie (edge) du réseau d'accès considéré et finalement, pas de dépendance géographique pour l'entité APN.

5.2 Limitations principales

Malgré le fait que cette recherche tente de concevoir un cadre global pour satisfaire le triplet opérationnel de QdS-Sécurité-Mobilité, notre démarche expérimentale, surtout pour les phases 2, 3 et 4, découle principalement de simulations sur ordinateurs (*OPNET Modeler*^{TM}, *SPIN*) et d'un développement mathématique de dimensionnement. Ceci s'avère être une première approximation. Il serait opportun de pouvoir reprendre la solution d'évolution de GTP et de l'implémenter sur un banc de test approprié. Ce dernier deviendra disponible dans un proche avenir puisqu'à Ericsson Recherche, nous avons l'intention d'ajouter à notre banc de test IPv6, des routeurs programmables qui nous permettrons de tester de nouveaux algorithmes.

Concernant également la proposition d'intégration de fonctionnalité de réseaux VPN, ce domaine doit non seulement être en ligne avec les pratiques courantes de conceptions de réseaux, mais également doit s'étendre aux accès sans fil. Il serait dès lors opportun de considérer un réseau dorsal qui supporte en même temps la connectivité des réseaux à grande échelle et des réseaux d'accès.

5.3 Indications pour des recherches futures

Ce mémoire couvre un cadre fertile ayant comme perspective la QdS-Sécurité-Mobilité. Il va de soi que ce cadre opérationnel peut facilement se multiplier en plusieurs avenues de recherche à approfondir davantage dans le but d'amener le triplet opérationnel vers un niveau de maturité plus avancé. Ces sujets de recherche future peuvent être décrits comme étant la définition d'un gestionnaire (TE-CMS) hors-ligne destiné à l'optimisation du CN à partir, d'heuristiques adaptées au besoin de réseaux mobiles à grands nombres d'abonnés, avec une variation importante des types de trafic. De plus, nous avons couvert le besoin de mobilité à grande échelle à partir de VPN,

mais ce fut dans les grands termes que cette approche a été abordée. Il serait recommandé de comparer entre eux quelques genres de VPN possibles pour déterminer précisément ce qui s'avérerait plus avantageux. Pour ce qui en est de l'APN, il semble être possible de pouvoir bénéficier des mêmes avantages que ceux offerts par Mobilité-IPv6, mais avons-nous tout considéré en termes d'échanges de messages de protocoles de mobilité? Enfin et par-dessus tout, il devient très important d'aligner davantage le mécanisme proposé de relève RAU améliorée, avec la norme HMIPv6, ce qui requiert une vue fonctionnelle globale entre les protocoles de la couche 3 et ceux de la couche 2.

Ceci dit, satisfaire le support bien équilibré du triplet QdS-Sécurité-Mobilité est une priorité non négociable. C'est ce qui différenciera dans le futur, un produit de télécommunications d'un autre.

Finalement, deux références additionnelles seront considérées dans le futur pour des besoins de comparaison de performance technologique. La première référence [21] identifie une méthode de synchronisation de chemins lors d'une relève intra SGSN basée sur le bi-chemin, alors que la deuxième référence [22] propose l'utilisation de MPLS dans le réseau cœur tout en considérant les bénéfices d'efficacité accrue de transmission, de balance de charge, de service de protection et aussi de support de QdS. Ces deux références ne représentent toutefois pas d'interférence de choix technologique avec les mécanismes proposés dans ce mémoire, puisque la relève macro inter SGSN n'est pas considérée dans [21] et [22].

BIBLIOGRAPHIE

[1] D. Bruce & R. Yakov, "MPLS Technology and Applications", Morgan Kaufmann Publishers, 2000, ISBN 1-55860-656-4.

[2] C. Perkins, "IP Mobility Support", RFC 2002, October 1996.

[3] Draft-ietf-choi-mobileip-ldpext-03.txt, "Extension of LDP for Mobile IP Service through the MPLS Network ", August 2001.

[4] C. Perkins, "Minimal Encapsulation within IP", RFC 2004, October 1996.

[5] R. Bhagavathula , N. Thanthry, W. Lee, and R. Pendse, "Mobility: A VPN Perspective", Circuits and Systems MWSCAS-2002 Symposium, 4-7 Aug. 2002, pp. 89-92.

[6] F. M. Chiussi, D. A. Khotimsky, and S. Krishnan, "Mobility Management in Third-Generation All-IP Networks", IEEE Communications Magazine, September 2002, pp. 124-135.

[7] Y. Lemieux, "RSVP/MIPv6 Integration", Quick Study under DELTA^{2002}, March 2002.

[8] 3GPP TR 23.923 V.3.0.0, "Combined GSM and Mobile IP Mobility Handling in UMTS IP CN", *3GPP*, May 2000.

[9] S. Pierre, "Réseaux et Systèmes Informatiques Mobiles", Presses Internationales, 2003, ISBN 2-553-01038-9.

[10] C. Semeria, "RFC 2547bis: BGP/MPLS VPN Fundamentals", Juniper White Paper, 2001.

[11] C. Semeria, "RFC 2547bis: BGP/MPLS VPN Hierarchical and Recursive Applications", Juniper White Paper, 2001.

[12] Y. Lemieux, A. Lemay, R. Ben Ali, "Planification et Conception de Réseaux MPLS", Travail effectué dans le cadre du cours INF-6470, 2004.

[13] Y. Lemieux, and L. Marchand, "Method and System for Multi-Protocol Label Switching (MPLS) Based Data Flow Aggregation in a Third Generation (3G)

Cellular Telecommunications System", P15266US, applied in USA, July 2002, Patent Application P15266US.

[14] C. Li, A. Burchard, and J. Liebeherr, "A Network Calculus with Effective Bandwidth", Technical Report: University of Virginia, CS-2003-20, November 2003, pp. 1-28.

[15] A. Gurtov, M. Passoja, O. Aalto, and M. Raitola, "Multi-Layer Protocol Tracing in a GPRS Network", IEEE Fall Vehicular Technology Conference 2002, pp. 1-5.

[16] V. Räisänen, "Implementing Service Quality in IP Networks", Nokia Networks OY, Finland, Wiley Edition, 2003, ISBN 0-470-84793-X.

[17] G. Dorvius, Y. Lemieux, S. Pierre, "Pro-Active UTMS QoS-Based Alternate Routing Algorithm", Opnetwork-2003 Conference, Washington DC.

[18] G. J. Holzmann, "The SPIN Model Checker", Addison-Wesley Publishers, 2003, ISBN 0-321-22862-6.

[19] R. Ben Ali, Y. Lemieux, and S. Pierre "UMTS to IP Backbone QoS Mapping Refinement for Multimedia Telephony Services", Opnetwork – 2003 Conference, Washington DC.

[20] A. Szlavik, G. Seres, J. Zatonyi, and J. Biro, "On the Applicability of the On-Off Type Approximation of the Effective Bandwidth Function", ICC 2003 Conference, 11-15 May 2003, pp. 183-187.

[21] S. Qingguo, S. Ruisong, W. Li, "QoS Guaranteeing during UMTS Packet-domain Handover", IEEE Parallel and Distributed Computing, Applications and Technologies -2003 Conference, pp. 387-390.

[22] H. Chueh, K. Wang, "An All-MPLS Approach for UMTS 3G Core Networks", IEEE Fall Vehicular Technology Conference 2003, pp. 2338-2342.

ANNEXE I

Nous avons utilisé le langage Promela pour définir le système déjà mentionné, de la façon suivante :

```
/* ********************************************************** */
byte    pk_to_receive =  3;             /* Generated Packets was 10 optimum */
/* ********************************************************** */
bool            status_pk_fw = 1;
byte            HO_after_pk_sent = 1;
byte            cnt_tx = 0;
byte            cnt_rx = 0;
byte            cnt_HO = 0;
/* ********************************************************** */
/* Message types                */
mtype = { LSP_predef, LSP_setup, LSP_setup_Ack, LSP_teardown };
/* MPLS-LSP Set-Up Messages     */
mtype = { RAU, SGSN_ctx_req, SGSN_ctx_resp, SGSN_ctx_ack};
/* 3GPP CN Session Messages     */
mtype = { data_PDU, vox_PDU_ON };
/* User Data                    */
/* Channel Descriptions         */
chan  Iu_ps1 = [1] of { mtype };
chan  Iu_ps2 = [1] of { mtype };
chan  Iu_ps3 = [1] of { mtype };
chan Gn = [3] of { mtype };
chan Go = [3] of { mtype };
mtype last_received;
/* ********************************************************** */
/*  Has 3 UE on 3 diff channels!!!                  */
active proctype UE1()
{
Generate:
        do
        :: status_pk_fw == 1 -> Iu_ps1!vox_PDU_ON; cnt_tx++;
        :: (cnt_tx  > HO_after_pk_sent && cnt_HO == 0) -> cnt_HO++; Iu_ps1!RAU;
status_pk_fw = 0;       /* Start Hand-Over  w 1x RAU!!! */
        od;
Finish:
```

```
                skip;
}
active proctype UE2()
{
Generate:
        do
        :: status_pk_fw == 1 -> Iu_ps2!vox_PDU_ON; cnt_tx++;
        :: (cnt_tx  > HO_after_pk_sent && cnt_HO == 0) -> cnt_HO++; Iu_ps2!RAU;
status_pk_fw = 0;
        od;
Finish:
                skip;
}
active proctype UE3()
{
Generate:
        do
        :: status_pk_fw == 1 -> Iu_ps3!vox_PDU_ON; cnt_tx++;
        :: (cnt_tx  > HO_after_pk_sent && cnt_HO == 0) -> cnt_HO++; Iu_ps3!RAU;
status_pk_fw = 0;
        od;
Finish:
                skip;
}
active proctype New_SGSN()
{
byte    n_ctx_req = 0;
byte    n_ctx_ack = 0;
HandOver:
        do
        :: Iu_ps1?RAU -> Gn!SGSN_ctx_req; n_ctx_req++;
        :: Iu_ps2?RAU -> Gn!SGSN_ctx_req; n_ctx_req++;
        :: Iu_ps3?RAU -> Gn!SGSN_ctx_req; n_ctx_req++;
        :: Gn?LSP_setup_Ack -> last_received = LSP_setup_Ack; /*IMPORTANT*/
        ::    Go?SGSN_ctx_resp->    ->    Gn!LSP_setup    ->    Gn!SGSN_ctx_ack;
n_ctx_ack++; status_pk_fw = 1;
        od;
Finish_Handover:
        skip;
}

active proctype Old_SGSN()
{
```

```
byte    n_ctx_resp = 0;
Idle:
        do
        :: Gn?SGSN_ctx_req -> Go!SGSN_ctx_resp; n_ctx_resp++; break;
        od
}
active proctype GGSN()
{
HO:
        do
        :: Gn?SGSN_ctx_ack -> last_received = SGSN_ctx_ack;    /*IMPORTANT*/
        :: Gn?LSP_setup -> Gn!LSP_setup_Ack;
        od
}
active proctype Appl_Serv()
{
Sink:
        do
        :: Iu_ps1?vox_PDU_ON -> cnt_rx++ -> last_received = vox_PDU_ON;
        :: Iu_ps2?vox_PDU_ON -> cnt_rx++ -> last_received = vox_PDU_ON;
        :: Iu_ps3?vox_PDU_ON -> cnt_rx++ -> last_received = vox_PDU_ON;

        :: cnt_rx > pk_to_receive  -> goto Finish;
        :: status_pk_fw == 0 -> goto After_HO;
/* Application-Server stops receiving during Hand-Over     */
        od;
After_HO:
        do
        :: Iu_ps2?vox_PDU_ON -> cnt_rx++;
        :: Iu_ps3?vox_PDU_ON -> cnt_rx++;
        :: status_pk_fw == 1 -> goto Sink;              /* No packet received until HO
status status_pk_fw == 1 */
        od;
Finish:
                skip;
}
```

ANNEXE II

La Revendication peu probable est définie de la façon suivante :

```
never {
        do
        :: last_received == vox_PDU_ON -> break
        :: else
        od;

        do
        :: last_received == SGSN_ctx_ack -> break
        :: else
        od;

        do
        :: last_received == vox_PDU_ON -> break
        :: else
        od;

        do
        :: last_received == vox_PDU_ON -> break
        :: else
        od;

        do
        :: last_received == LSP_setup_Ack -> break
        :: else
        od;

        }
```