Enabling Roaming in Heterogeneous Multi-Operator Wireless Networks

Oscar Salazar Gaitan

École Doctorale
d'Informatique,
Télécommunications
et Électronique de Paris

Thèse

présentée pour obtenir le grade de Docteur de l'École Nationale Supérieure des Télécommunications

Spécialité : **Informatique et Réseaux**

OSCAR SALAZAR GAITÁN

Mobilité des services dans les réseaux hétérogènes dans une contexte multi-fournisseur

Soutenue le 24 septembre 2007 devant le jury composé de

Abraham O. Fapojuwo	Président
Omar Cherkaoui	Rapporteur
Philippe Godlewski	Examinateur
Philippe Bertin	Examinateur
Philippe Martins	Directeur de thése

Enabling Roaming in Heterogeneous Multi-Operator Wireless Networks

Allowing inter-network cooperation in wireless environments beyond 3G

by

OSCAR SALAZAR GAITÁN

Submitted to the department of Computer Sciences & Networks in partial fulfillment of the requirements for the degree of

DOCTOR of PHILOSOPHY in TELECOMMUNICATIONS

at the

École Nationale Supérieure des Télécommunications

September 24, 2007

To my wife, family, and friends

Acknowledgements

This thesis would not be possible without the support of so many people. People that provided me the inspiration, the support, the necessary feedback and advices. People whom I spent countless hours sharing four years of my life.

I would like to thank you all without mentioning any names. I am certain that you know who you are.

Oscar

Résumé en français

1.1 Introduction

Dans un futur proche, les technologies d'accès sans fil telles que le WiFi, le WiMAX et l'UMTS coexisteront. Néanmoins, cette coexistence quotidienne ne signifie pas qu'ils seront en mesure d'être pleinement inter-opérationnels. Nous pensons alors que le roaming concernera différents réseaux d'accès gérés par différentes entités : des opérateurs de réseaux cellulaires, des fournisseurs d'accès Internet (*FAI*) et certaines organisations ou individus. Notre travail de recherche s'articule ainsi autour d'un scénario de roaming hétérogne, concernant les réseaux d'accès sans fil gérés par différents opérateurs. Notre objectif principal est de fournir une architecture de roaming qui permet l'interopérabilité des réseaux hétérogènes dans le cadre d'un environnement multi-opérateurs, et ce sans grands changements dans les architectures sans fil actuelles. L'un des objectifs de la nouvelle génération des réseaux sans fil NGWNs (*Next GenerationWireless Networks*) est la possibilité de transférer de manière transparente les services entre les réseaux sans fil de différentes technologies d'accès. L'objectif est de tirer partie de la popularité et des hauts débits des UWNs (*Unlicensed Wireless Networks*) afin d'améliorer les services mobiles. Bien qu'il existe déjà des solutions qui tentent de répondre à la problématique de la mobilité de service, la plupart d'entre elles sont basées sur l'hypothèse que les opérateurs mobiles sont également propriétaires des réseaux WiFi / WiMAX. La réalité est que, en dépit du fait que les opérateurs mobiles investissent énormément dans le déploiement de leurs propres UWNs, de nombreux UWNs indépendants peuvent être utilisés comme des extensions de ces réseaux mobiles. La principale problématique est alors le manque d'accords de roaming entre les réseaux mobiles et les opérateurs indépendants (*UWNs*), et l'absence d'une architecture solide permettant cette intégration. Nous considérons que les UWNs indépendants sont des réseaux sans fil qui n'ont pas établi des accords contractuels de roaming avec les opérateurs mobiles. Actuellement, les compagnies de télécommunication offrent des services tels que le *triple play* Ce dernier est le terme commercial qui désigne la fourniture de deux services à large bande passante (*l'accès Internet haut débit et la télévision*) et un service à faible bande passante (*la téléphonie*), via une seule connexion à large bande passante. Dans le *quadruple play*, la communication sans fil est présentée comme un autre moyen qui permet de fournir du contenu multimédia tel que la vidéo, l'accès Internet et le service de téléphonie vocale. La combinaison du triple et du quadruple play donnent une tendance claire vers la FMC (*Fixed Mobile Convergence*). L'objectif est de fournir de tels services avec un numéro de téléphone

unique (*dual mode handset*), quelque soit le réseau d'accès sans fil. Cependant, même si l'appareil mobile est techniquement en mesure de bénéficier des avantages de la FMC, il se heurte à la difficulté de trouver un UWN qui autorise la fourniture du service de roaming. Nous énonçons dans ce qui suit un scénario typique de roaming entre opérateurs hétérogènes :

Alice a un téléphone portable avec deux interfaces réseaux (UMTS et WiFi). Chaque fois qu'elle rentre à la maison, son téléphone est en mesure de passer du réseau UMTS au réseau WiFi, et delui fournir des services de voix et de données (définies dans le profil de service d'Alice). Cela est rendu possible car l'opérateur mobile d'Alice est également son fournisseur d'accès Internet. Néanmoins, lorsque Alice est de retour à son bureau, son téléphone portable n'est pas en mesure de passer sur le réseau WiFi de sa société. La raison est que le réseau sans fil de son entreprise ne dispose pas d'un contrat de roaming avec son opérateur mobile.

Si la société d'Alice veut faire partie de la réseau mobile d'Alice, elle devrait mettre en place un accord contractuel de roaming avec l'opérateur mobile, chose certainement pas facile à réaliser. Pour supporter un roaming transparent et hétérogène, les UWNs doivent être en mesure de fournir l'accès transparent aux utilisateurs mobiles qui veulent y accéder s'ils veulent offrir de meilleurs services mobiles. Cette opération suscite cependant plusieurs questions: Pourquoi un opérateur indépendant UWN doit donner accès à un utilisateur inconnu? Pourquoi les UWNs doivent partager leurs ressources pour satisfaire les utilisateurs mobiles? Quel est le profit de l'opérateur UWN? Même si le UWN est prêt à partager des ressources, comment pouvons-nous déterminer s'il peut répondre aux exigences en termes de QoS? Comment l'utilisateur mobile peut recevoir des appels téléphoniques à travers le UWN? Comment un UWN pourrait identifier l'utilisateur mobile et vérifier quel type de services cet utilisateur est autorisé à effectuer? Et finalement, si les UWNs ont prêts à devenir des extensions des réseaux mobiles, comment peuvent-ils régler ces accords de roaming? Notre principale motivation concerne les deux dernières questions et vise la fourniture d'une stratégie efficace sous la forme d'une plateforme de roaming qui prend en charge la mobilité des services entre les réseaux hétérogènes dans un contexte multi-fournisseurs.

1.2 Problématiques étudiéess

La mobilité de services dans un contexte hétérogène inclue l'intégration des opérateurs de réseaux avec différentes interfaces d'accès, diverses stratégies de gestion et des modèles économique. Notre principe est de mettre en place des accords de roaming entre les fournisseurs en chargeant le réseau mobile de la prise de décision relative au roaming. À cet égard, nous avons identifié trois principales problématiques qui revêtent une importance cruciale : l'authentification des utilisateurs mobiles, l'autorisation du profil de service et la facturation.

1.2.1 Authentication d'Utilisateurs

Le provisionnement de l'accès au réseau sans fil dans un UWN n'est pas une tâche évidente. Pour atteindre cet objectif, les opérateurs UWNs doivent valider l'identité de l'utilisateur mobile et vérifier le profil du service. Traditionnellement, le processus d'authentification est effectué en interrogeant les bases de données dans le réseau domestique (*réseau mobile*). Pour accroître la sécurité de la transaction, certaines propositions suggèrent l'utilisation du réseau UWN des mécanismes d'authentification tels que EAP-SIM ou EAP-AKA. Néanmoins, de tels mécanismes ne peuvent être appliqués que si les deux opérateurs mobiles ont accès à ce type d'authentification et ont déjà établi un rapport de confiance lié par des accords de roaming (*roaming agreements*).

1.2.2 Authorization du Profil d'Utilisateur

Quand un utilisateur mobile arrive dans un VN (*Visiting Network*), et essaie d'instancier ou de poursuivre un service déjà instancié, le VN détermine, selon le profil utilisateur, les droits de l'utilisateur d'accéder à ces services. Dans un environnement multi-opérateurs, cette action représente un nouveau défi, car il n'est pas recommandé de partager le profil utilisateur du service avec des VNs malicieux ou fournir toute information relative à cet utilisateur mobile s'il n'y a pas une relation de confiance. Le profil utilisateur contient des informations confidentielles sur le client, ainsi il ne peut pas être partagé avec les UWNs indépendants. Un autre inconvénient est que les réseaux mobiles ne répondront à aucune requête liée au profil des utilisateurs mobiles et provenant d'un UWN malicieux.

1.2.3 Facturation

Il n'y a aucun doute sur le fait que le roaming hétérogène apportera, aux opérateurs mobiles, des avantages technologiques. D'autre part, les opérateurs des réseaux UWN offriront des services à valeur ajoutée tels que les appels voix sur IP (*VoIP*) vers les téléphones mobiles et les réseaux PSTN. Les gains aussi bien économiques que technologiques motivent les opérateurs UWNs à ouvrir leurs réseaux et à coopérer avec les opérateurs mobiles autour de la fourniture de mobilité de services verticale. À partir de cette prémisse, nous considérons que, même si la question de la facturation n'est pas traitée dans notre recherche, notre plate-forme de roaming devrait fournir un framework pour l'appui de cette fonctionnalité afin de poursuivre le développement d'une plate-forme de facturation hétérogène.

1.3 Organisation de la Thèse

La structure de cette thèse est organisée comme suit:

- Le chapitre 3 présente une introduction générale au NGWNs et présente les enjeux et les défis du roaming dans de tels environnements. Nous montrons également pourquoi les solutions actuelles dans ce domaine ne remplissent pas les conditions requises pour un roaming sans soudure dans un contexte

hétérogène et nous soulignons les différences avec nos travaux de recherche. Ainsi, nous nous concentrons sur la prochaine génération de services mobiles. Nous analysons la question de la mobilité de service dans les réseaux sans fil hétérogènes et nous présentons les technologies émergentes dans le domaine la mobilité invisible des services. Dans ce chapitre, nous donnons également la logique derrière notre proposition.

- Le chapitre 4 est consacré à notre première contribution. Nous y présentons notre architecture de roaming basée sur SIP. Nous définissons des éléments architecturaux et nous décrivons toutes leurs fonctionnalités. Une description détaillée du protocole de signalisation est également fournie. Enfin, nous présentons une évaluation de notre architecture à travers une simulation et une étude de faisabilité.

- Dans le chapitre 5, nous abordons des questions importantes de notre architecture : le compromis en terme de signalisation de notre architecture et le renforcement des mécanismes d'authentification et d'autorisation. Une importante contribution est également présentée dans ce chapitre : les certificat d'attributs embarqués dans le protocole SIP.

- Enfin, le chapitre 6 conclut cette thèse et donne un aperçu des perspectives et des travaux futurs dans ce domaine de recherche.

La bibliographie complète et la terminologie indice se trouvent à la fin de ce document. Les détails de mise en place sur les concepts et les bancs d'essai se trouvent dans les annexes de cette thèse.

1.4 Contributions

Au cours de nos travaux de recherche, nous avons étudié les différents aspects de l'intégration dans les réseaux hétérogènes. Dans cette thèse, nous présentons une architecture de roaming. Notre proposition repose sur une approche basée sur un courtier (*broker*) qui contribue à réduire le nombre de relations de confiance entre les opérateurs réseau et les UWNs, tout cela basé sur le principe de transitivité. Ainsi, le réseau domestique (*Home Network*), par le biais d'une entité appelée *Roaming Broker*, établit une relation de roaming avec les UWNs indépendants. Dans notre architecture, les accords de roaming ont pour objectif l'établissement de la confiance mutuelle entre les opérateurs de réseaux hétérogènes en assurant le respect des Service Level Agreements et en fournissant des mécanismes de contrôle d'accès efficaces (*authentification et autorisation*) pour le roaming des utilisateurs mobiles dans les UWNs. Nous avons décidé d'utiliser SIP comme protocole de signalisation en raison de sa simplicité et du fait qu'il joue déjà un rôle important dans le IMS (*IP Multimedia Subsystems*) dans les réseaux 3G.

En identifiant, étudiant, et analysant les enjeux et les défis du roaming entre plusieurs opérateurs hétérogènes et en fournissant une plateforme de roaming efficace basée sur SIP, nous avons réalisé les contributions suivantes :

- Nous avons analysé l'impact du roaming entre plusieurs opérateurs hétérogènes sur la FMC. Nous avons identifié les problèmes et les défis de ce type de roaming.
- Nous avons réalisé une analyse approfondie des relations de confiance et des rôles des différents éléments de notre architecture. Nous avons introduit une infrastructure de confiance qui permet de lier des relations de confiance entre les opérateurs mobiles et les UWNs par l'intermédiaire d'un élément nouveau appelé le Roaming Broker.
- Nous avons étudié l'impact de l'AAA (*Authentication, Authorization, and Accounting*) sur les mécanismes du processus de handover vertical. À travers cette analyse, nous avons identifié les problèmes de l'authentification et de l'autorisation comme un potentiel handicap pour un roaming transparent.
- Nous avons amélioré le processus d'authentification et d'autorisation basé sur SIP grâce à l'utilisation des mécanismes distribués de PKI et de PMI.
- Enfin, nous avons proposé une technique qui permet une validation locale (*dans le VN*) du profil du service de l'utilisateur, sans compromettre la sécurité de l'utilisateur mobile. Notre méthode repose sur des certificats d'attribut XML embarqués dans SIP.

1.5 Conclusion et Perspectives

Notre architecture offre le roaming grâce à l'ajout d'un nouvel élément architectural appelé le *Roaming Broker*. Ainsi, grâce à cet élément, les relations de confiance sont construites entre le réseau mobile et les UWNs. La confiance entre les réseaux sans fil hétérogènes est établie de la manière suivante : la confiance entre la RB et le réseau mobile s'appuie sur les accords de roaming contractuels alors que la confiance entre le RB et le UWNs se fait avec des SLAs placés au niveau des points d'accès WiFi. L'échange de messages de signalisation est réalisé à travers une version améliorée de SIP. Par conséquent, SIP est modifié suivant les recommandations de l'architecture afin de transmettre des informations critiques au pocessus de roaming. En particulier, deux des messages SIP modifiés sont le message *SIP REGISTER* et le message *SIP INVITE*. Après notre modification, les deux messages transmettent de l'information concernant l'utilisateur mobile, le UWN, et les SLAs. Ainsi, le HN peut déterminer si l'utilisateur peut ou pas errer dans le VN. L'évaluation de notre architecture a été réalisée par simulation sur ordinateur. Une étude de faisabilité a également été effectuée pour démontrer l'applicabilité réelle de notre proposition. Dans cette perspective, les résultats simulation sur ordinateur ont montré que le déploiement de notre architecture de roaming basée sur SIP augmente légèrement le délai de traitement des messages SIP. D'autre part, en raison des informations liées au roaming et transmises via les messages SIP (*augmentant la taille des messages SIP*), le surcoût de la signalisation marque aussi une légère augmentation. Néanmoins, comme indiqué par les résultats de la simulation, ces changements ne représentent pas

un impact considérable sur SIP. L'étude de faisabilité a également démontré que, basée sur les caractéristiques techniques actuelles des points d'accès sans fil i.e. *la capacité de traitement des calcul et de mémoire de stockage, l'implémentation d'une telle plate-forme est possible*. Une supposition importante de notre architecture est que le réseau mobile est toujours le *decision maker*. Par conséquent, il doit être interrogé en permanence grâce à des messages de signalisation. Ainsi, nous avons décidé de diminuer l'effet de la triangulation tout en assurant la sécurité des réseaux robustes et sans compromettre le profil d'utilisateur. Pour atteindre cet objectif, notre architecture intègre la solution PKI & PMI. La mise en Ïuvre de ces mécanismes de sécurité repose également sur notre protocole de roaming basé sur SIP. Dans cette perspective, nous nous appuyons sur une infrastructure de confiance approuvée par des éléments PKI & PMI répartis dans l'architecture de roaming et de l'utilisation de certificats XML embarqués dans le protocole SIP, afin de prouver les droits de l'utilisateur mobile au moment du roaming ou lorsque l'utilisateur lance les services à partir d'un VN. Pour l'évaluation de l'impact des ces nouveaux éléments de sécurité sur notre architecture, nous nous sommes fondés sur la simulation par ordinateur. Le performance de notre architecture avec nos contributions a montré une amélioration globale en terme de délai de signalisation, en particulier dans l'enregistrement et le lancement de session. Le fait que le processus d'authentification et d'autorisation, sont désormais effectués, réduit l'échange de messages de signalisation et le retard introduit par l'Internet.

1.5.1 Perspectives

Nous considérons que le roaming dans les réseaux hétérogènes n'est pas une tâche évidente; ainsi, afin de fournir un environnement efficace de roaming, nous devons compter sur : des techniques robustes de vertical handover, et des protocoles de gestion de la mobilité efficaces et une prise de décision optimale pour le roaming. Nos travaux futurs consistent à évaluer la qualité de service tels que les paramètres de délai, la gigue (*jitter*) ou de la perte des paquets, afin d'analyser et d'étendre les capacités du vertical handover dans notre architecture. L'extension à d'autres environnements de réseau sans fil, tels que les systèmes de distribution sans fil ou les réseaux maillés sans fil est également recommandée. En outre, bien que nos travaux de thèse ne mettent pas l'accent sur les questions liées au processus de roaming hétérogène, nous considérons que ce sujet doit être traité afin d'améliorer la transparence et l'efficacité de ce processus.

Pendant nos travaux de recherche et durant la phase de conception de notre architecture de roaming, nous avons identifié certaines questions qui ont attiré notre attention et qui, nous en sommes convaincus, pourraient être des sujets de recherche intéressants.

- L'intégration des mesures de QoS basées sur le protocole RTCP dans le VN: Le RB utilise l'information sur la QoS pour gérer les SLAs dans les VNs. Grâce à cette information, il crée et met à jour la liste des réputations des VNs. De plus, en se fondant sur des statistiques en temps réel la qualité de service, le RB fournit les informations nécessaires à la HN afin de déterminer

s'il y a lieu ou non de déclencher le processus de roaming. La faisabilité de cette technique a été confirmée lors de la mise en Ïuvre des expérimentations sur ordinateur. Nous avons développé un agent logiciel qui surveille de manière continue des statistiques sur la QoS provenant des appels VoIP en utilisant des paquets RTCP dans le réseau. Malheureusement, nous n'avons pas pu évaluer les performances d'un tel mécanisme. L'objectif des mesures de la QoS basée sur RTCP est de fournir des renseignements aux points d'accès sans fil, et plus précisément 802.11, afin d'être conscient des statistiques de la QoS dans le VN.

- CAC (*Call Admission Control*) basé sur la qualité de la voix:
  Traditionnellement, les CAC basés sur 802.11 tiennent compte des limitations physiques de l'accès réseau sans fil afin de décider d'accepter ou de rejeter un appel dans le réseau. Les mécanismes CAC actuels limitent les appels au nombre de stations dans le réseau. L'hypothèse de base est que toutes les stations utilisent les mêmes applications de VoIP, les mêmes vocoders et les mêmes protocoles. Dans la littérature, nous pouvons trouver certaines analyses qui évaluent la capacité maximale d'un réseau sans fil 802.11 basée sur mes mesures subjectives de la qualité de la voix, plus précisément, en utilisant ITU E-Model. Nous considérons que ces paramètres peuvent contribuer à l'élaboration d'un CAC basé sur la qualité de la voix. Un tel mécanisme pourrait accepter ou rejeter les appels VoIP basés sur la manière dont l'appel peut être perçu par l'utilisateur mobile à travers l'état actuel du réseau sans fil plutôt que sur le nombre de stations actuellement dans le réseau.

- Amélioration de la prise de décision pour le roaming:
  L'adoption et le déploiement des technologies sans fil sans licence contribue à la rapide augmentation de la densité des UWNs dans certaines villes. Ë cet égard, en tentant d'errer dans un VN, il sera fréquent de trouver plus d'un UWN à disposition prêt à offrir des services de roaming. Ici, le réseau mobile (le décisionnaire du roaming) doit choisir le réseau optimal pour l'utilisateur mobile. La combinaison de la théorie de la décision, les approches MADM (*Multiple Attribute Decision Making*) et les techniques d'optimisation peuvent offrir d'importantes perspectives en vue d'améliorer ces processus.

Nos recherches ont porté principalement sur les aspects techniques du roaming hétérogène. Nous sommes cependant conscients que les réseaux hétérogènes ne diffèrent pas seulement dans la technologie mais aussi dans les modèles économiques et stratégies de gestion. Dans cette perspective, nous avons identifié trois domaines de recherche qui pourraient améliorer le roaming entre plusieurs opérateurs hétérogènes : des mécanismes efficaces de facturation, de nouveaux modèles économiques et des mesures incitatives pour le provisionnement du roaming.

- Des nouveaux mécanismes efficaces de facturation:
  La question de la tarification est primordiale pour offrir des services de roaming, les opérateurs mobiles ne déploieront aucune architecture de roaming s'il n'y a aucun moyen de facturer les utilisateurs mobiles. Ainsi, nous

estimons que de nouveaux mécanismes de facturation visant à fournir une plateforme distribuée efficace de tarification motiveraient les opérateurs réseaux ou fournisseurs de services à participer au roaming hétérogène.

- Des nouveaux modèles économiques:
  Des nouveaux modèles économiques peuvent apporter des solutions sur la manière de commercialiser ce type de roaming et peut-être définir une future application. Rappelons qu'au sein de notre architecture, trois principaux éléments participent à ce sujet : le réseau mobile, le roaming broker et le réseau de visite. Par conséquent, il est nécessaire de parvenir à créer de nouveaux modèles qui fournissent une claire stratégie économique pour permettre à tous les participants de notre architecture d'être gagnants économiquement.

- Incitations au provisionnement du roaming:
  Les UWNs indépendants sont des éléments importants dans notre architecture dans la mesure où ils sont prêts à partager leurs ressources pour offrir des services de roaming à des utilisateurs mobiles.

Nous sommes certains qu'en s'attaquant à ces problématiques techniques et économiques, les performances de notre plate-forme de roaming s'amélioreront considérablement et pourrait être envisagée pour un futur déploiement sous des scénarios réels.

Abstract

In a future, existing wireless access technologies such as WiFi (*Wireless Fidelity*), WiMAX (*Worldwide Interoperability for Microwave Access*) and UMTS (*Universal Mobile Telecommunication System*) will coexist on daily basis. Nevertheless, this daily coexistence does not imply that they will be able to fully interoperate.

We considered that in a future, the roaming landscape will be characterized by different access networks managed by different entities, such as cellular network operators, ISPs (*Internet Service Providers*), and independent organizations or individuals. Thus, our research work focuses on a specific roaming scenario which is formed by heterogeneous wireless access networks managed by different operators. Our main objective is to provide a seamless (*for the mobile user*) roaming architecture to enable network interoperability under a *heterogeneous multi-operator wireless environment*, all this without major changes in current wireless architectures.

The SIP-based roaming architecture aims at the creation of a suitable environment for the support of seamless service mobility in heterogeneous multi-operator wireless networks, this without major changes in current wireless network architectures. To achieve this, we rely on an element called the Roaming Broker. The goal of this entity is to establish *mutual trust* between the cellular and the Unlicensed Wireless Networks. As trust can be built by different means, for our research we consider that mutual trust between the cellular network and the Roaming Broker is endorsed by contractual agreements. On the other hand, trust between the Roaming Broker and the independent wireless networks is endorsed by SLA (*Service Level Agreements*) and efficient access control mechanisms. Furthermore, we assume that roaming between heterogeneous wireless networks, follows economic rather than technical reasons. Consequently, heterogeneous roaming must be allowed only if the Visiting Networks respects the accorded Service Level Agreements.

The SIP-based Roaming Architecture enables roaming in heterogeneous multi-operator wireless environments. In this architecture, the HN is always the roaming decision maker as it is determines, based on the mobile user's and VN's credentials, whether or not the mobile user can roam in to the VNs. That is mobile user authentication and authorization are always performed by the HN and from the HN. Furthermore, as VNs are not allowed to communicate directly with the HN, all these roaming signaling messages must pass through the RB. Thus, as confirmed by the simulation results, this process clearly contributes with an increment in the overall delay of the network registration and session initiation process. In this regard, we aim at reducing the roaming signaling exchange caused by

the authentication and authorization mechanisms in the network registration and session initiation process while providing robust network security. To achieve this, we rely on the use of Public Key Infrastructure and Privilege Management Infrastructure) techniques deployed on top of our architecture. Moreover, for authorization, we propose the use of PMI and XML Attribute Certificates to define the rights and roles of the key elements of our architecture.

The results obtained through computer simulation indicated that the use of this approach reduces significantly the network registration and session initiation delay, hence outperforming the traditional Roaming-SIP method. We also confirmed that the wireless delay introduced by the VN increases considerably when increasing the traffic congestion level in the Visiting Networks. Thus, from the simulations results we can state that by reducing the signaling message exchange, and maintaining acceptable congestion levels in the wireless network, hence reducing the wireless transmission delay, we can improved the overall delay in both network registration and session initiaition process.

Contents

List of Figures

List of Tables

Acronyms

AAA	Authentication, Authorization, and Accounting
ACS	Ambient Control Space
AS	Application Server
ATM	Asynchronous Transfer Mode
CDMA	Code Division Multiple Access
CSMA/CA	Carrier Sense Multiple Access/Collision Avoidance
DES	Data Encryption Standard
DHCP	Dynamic Host Configuration Protocol
DIFS	Distributed Inter-frame Space
DSL	Digital Subscriber Line
DSSS	Direct Sequence Spread Spectrum
DCF	Distributed Coordinating Function
EAP-SIM	Extensible Authentication Protocol-Subscriber Identity Module
EAP-AKA	Extensible Authentication Protocol-Authentication and Key Agreement
EIR	Equipment Identity Register
ETSI	European Telecommunications Standards Institute
FHSS	Frequency Hoping Spread Spectrum
FMC	Fixed Mobile Convergence
GANC	Generic Access Network Controller
GERAN	Generic Radio Access Network
GGSN	Gateway GPRS Support Node
GLR	Gateway Location Register
GSM	Global System for Mobile Communications
HLR	Home Location Register
HN	Home Network
HSDPA	High Speed Downlink Packet Access
HSS	Home Subscriber Server
IMS	IP Multimedia Subsystem
IMT-2000	International Mobile Telecommunications 2000
ISDN	Integrated Services Digital Network
ISM band	Industrial, Scientific and Medical band
I-CSCF	Interrogating Call Session Control Function
IP	Internet Protocol
ISP	Internet Service Provider
LTE	Long Term Evolution
MSC	Mobile Switching Center

NGWN Next Generation Wireless Network
OFDM Orthogonal Frequency Division Multiplexing
P-CSCF Proxy Call Session Control Function
PDA Personal Digital Assistant
PDG Packet Data Gateway
PDP Packet Data Protocol
PKI Public Key Infrastructure
PMI Privilege Management Infrastructure
PMI Privilege Management Infrastructure
PSTN Public Switched Telephone Network
PCF Point Coordinating Function
QoS Quality of Service
RAN Radio Access Network
RB Roaming Broker
RNC Radio Network Controller
S-CSCF Serving Call Session Control Function
SCTP Stream Control Transmission Protocol
SDP Session Description Protocol
SEGW Security Gateway
SGSN Serving GPRS Support Node
SIFS Short Inter-frame Space
SIP Session Initiation Protocol
SLA Service Level Agreement
TDMA Time Division Multiple Access
UE User Equipment
UMA Unlicensed Mobile Access
UMTS Universal Mobile Telecommunications System
UNC UMA Network Controller
UTRAN UMTS Terrestrial Radio Access Network
UWN Unlicensed Wireless Network
VHE Virtual Home Environment
VLR Visitor Location Register
VN Visiting Network
VoIP Voice over IP
WiFi Wireless Fidelity

WiMAX Worldwide Interoperability for Microwave Access
WISP Wireless ISP
WLAN Wireless Local Area Network
WMAN Wireless Metropolitan Area Network
WPA Wireless Protected Access
WEP Wired Equivalent Privacy
XML Extensible Markup Language
3G Third Generation
3GPP The 3G Partnership Project
4G Fourth Generation

Chapter 2

Introduction

In a future, existing wireless technologies such as WiFi (*Wireless Fidelity*), WiMAX (*Worldwide Interoperability for Microwave Access*) and UMTS (*Universal Mobile Telecommunication System*) coexist on daily basis. Nevertheless, this daily coexistence does not mean that they will be able to fully interoperate.

Driven by IP-centric architectures and protocols, the so-called all-IP network architectures have been proposed as underlying platform for NGWN (*Next Generation Wireless Networks*). As consequence, we will see the raising of a new generation of mobile services hence also posing new issues and challenges regarding seamless service mobility in such heterogeneous wireless environments.

We considered that In a future, the roaming landscape will be characterized by different access networks managed by different entities, such as cellular network operators, ISPs (*Internet Service Providers*), and independent organizations or individuals. Our research work focuses on a specific roaming scenario which is formed by heterogeneous wireless access networks managed by different operators. Our objective is to provide a seamless (*for the mobile user*) roaming architecture to enable network interoperability under a *heterogeneous multi-operator wireless environment*, all this without major changes in current wireless architectures.

The outline of this chapter is as follows: section 2.1 presents the background and motivation of our research. Section 2.2 provides the problem statement addressed in this thesis. Furthermore, section 2.3 explains the main contributions of our research to this domain. Finally, section 2.4 provides the thesis overview.

2.1 Background and Motivation

One of the goals of NGWNs refers to the ability of services to be seamlessly transferred between wireless networks of different access technologies. The objective is to exploit the popularity and high data rates of UWNs (*Unlicensed Wireless Networks*) to enhance cellular services. Although, there are certain approaches that attempt to address this kind of service mobility, most of them rely on the assumption that cellular operators also own the WiFi/WiMAX networks. The reality is that in spite of the effort that cellular operators put in the deployment of their own UWNs, there is an enormous amount of independent UWNs that can be used as

extensions of cellular networks. The main issue here is the lack of roaming agreements between cellular networks and independent UWN operators, and a solid architecture to enable such integration. In our research, independent UWNs are wireless network that does not have established contractual roaming agreements with cellular operators.

Currently, telecommunication companies offer services such as the *triple play*, this service is a marketing term for the provisioning of two broadband services (*high-speed Internet access and television*) and one narrowband service (*telephone*), over a single broadband connection. Due to the explosive growth of UWNs, telcos have proposed the *quadruple play* as the wireless evolution of the triple play. Here, wireless communication is introduced as another medium to deliver multimedia content such as video, Internet access and voice telephone service. The combination of triple and quadruple play yields a clear trend in the form of FMC (*Fixed and Mobile telephony Convergence*). The aim is to provide such services with a single phone (*dual mode handset*), that is able of switching between different wireless access networks. Thus, the mobile user can support both the cellular access (*wide area*) and the local/metropolitan area technology. Nevertheless, even the mobile device is technically capable of enjoying the benefits of the FMC, there is till the constraint of finding a UWN that is enabled to provide our roaming services. A typical scenario of heterogeneous multi-operator of roaming is the following:

> *Alice has a dual interface mobile phone (UMTS-WiFi) and every time she gets back home, her phone is able to switch over her WiFi network to provide her voice and data services (defined in Alice's service profile). This is possible because Alice's cellular operator also owns the Alice's ISP. Nevertheless, when Alice's gets back into her office, her cell phone is not able to switch over the company's WiFi network and get her cellular services. The reason is that her company's wireless network does not have established any roaming relation with her cellular operator.*

If Alice's company wants to support FMC and become extension of Alice's cellular network, it should establish a contractual roaming agreement with the cellular operator and this is certainly not easy to accomplish.

For the support of seamless heterogeneous roaming, independent UWNs must be able to provide seamless access to visiting mobile users if they want to provide enhanced cellular services. This action however poses several questions: why should an independent UWN operator provide access to an unknown user? why would the UWN share their resources to satisfy mobile users from a different network? what is the profit for the UWN operator? Even if the UWN's is willing to share resources, how can we determined wether or not it can meet the QoS requirements specified my the mobile user? How can the mobile user receives phone calls through the UWN? How a UWN could identify the mobile user and verify what kind of services the user is allowed to perform?and finally, if UWN's are willing to become an extension of cellular networks, how can they settle these roaming agreements?

The main motivation of our research lies on the last two questions and aim at

the provisioning of effective answers in the form of a roaming platform that supports seamless transfer of services within a heterogeneous multi-operator wireless environment.

2.2 Problem Statement

Heterogeneous multi-operator roaming involves the integration of network operators with different wireless interfaces, diverse management strategies and business models. Our premise is to establish inter-provider roaming agreements while granting the roaming decision making to the cellular network. In this respect, we have identified three main issues that are paramount in preventing seamless wireless interoperability: mobile user authentication, service profile authorization, and accounting.

Although accounting relates to billing, and this is the source of revenues of network operators, we considered the billing issue out of the scope of our research.

2.2.1 Mobile User Authentication

The provisioning of wireless network access into the UWN to mobile cellular users is not a trivial task. To achieve this, UWNs operators must validate the mobile user's identity and verify the service profile. Traditionally, the authentication process is performed among network operators by querying the user databases in the home network (*cellular network*). To increase the security of the transaction, some proposals suggest that the UWN network must use authentication mechanisms such as *EAP-SIM or EAP-AKA*. Nevertheless, such mechanisms can only be applied if both network operators support this kind of authentication and have previously established a trust relation bound by roaming agreements.

2.2.2 Service Profile Authorization

When a mobile user roams into a VN (*Visiting Network*) and attempts to initiate or continue an initiated service, the VN determines, base on the user's service profile, the right of the user to perform such services. Under a multi-operator environment, this action introduces a new challenge because it is not recommended to share the user's service profile with untrusted VNs nor to provide any information regarding the mobile user if there is no trust relation. The user profile contains confidential information about the customer hence it cannot be shared with independent UWNs. Another drawback is that cellular networks will not respond any query related to the mobile users profile that comes from an untrusted UWN.

2.2.3 Accounting

There are no doubts that heterogeneous roaming will bring technological benefits to network operators. Cellular operators see their network coverage area

increased, the capacity of cellular network is benefited by the lowering of traffic when mobile users operate from UWNs, high bandwidth applications can be supported from the UWN i.e *peer to peer wireless VoIP communication*. On the other hand, UWN operators gain an add value service within their networks, for example: VoIP calls to mobile phones and PSTN network. Both economical and technological gains are the motivation for independent UWN operators to open their networks and cooperate with cellular operators in providing dynamic and seamless vertical service mobility. From this premise, we consider that even though the issue of accounting is not addressed in our research, our roaming platform should provide a framework for the support of this feature for the further development of a heterogeneous billing platform.

2.3 Thesis Contributions

During our research work, we studied the different aspects of NGW architectures assuming an IP-centric network integration i.e. *a common IP-based interface for data transport over heterogeneous networks*. In this document, we present our vision in this area, and define our roaming architecture with all the key elements. Furthermore, we analyzed and identified the issues and challenges of our research. Finally, we evaluated through computer simulation our proposals in this subject matter.

In this thesis, we present a SIP-based roaming architecture to enable roaming in heterogeneous multi-operator wireless networks. Our proposal relies on a broker-based approach to contribute in reducing the number of trust relationships between network operators and the UWNs based on the *principle of transitivity* (Wilhelm & Butty, 1998). Thus, the HN (*Home Network*), through an entity called the RB (*Roaming Broker*), established a roaming relation with independent UWNs.

In our architecture the roaming agreements have as goal the establishment of mutual trust between heterogeneous network operators by ensuring the respect of SLAs (*Service Level Agreement*) and providing efficient access control mechanisms (*authentication and authorization*) for cellular users roaming into UWNs. The key-elements (*the HN, the VN, and the RB*) in our architecture communicate through an enhanced version of SIP (*Session Initiation Protocol*) (Rosenberg & Schulzrinne, 2006) also proposed in this thesis.. We decided to use SIP as signaling protocol due to its simplicity and the fact that it already plays an important role in the IMS (*IP Multimedia Subsystem*) in 3G (*Third Generation*) networks (Garcia-Martin, May 2005) (Soininen, August 2003).

By identifying, studying, and analyzing the issues and challenges of heterogeneous multi-operator roaming and providing an efficient SIP-based roaming platform, we made the following contributions:

- We analyzed the impact of heterogeneous multi-operator roaming on the FMC. We identified the issues and challenges of this type of roaming. Among the identified issues we found: mobile user's authentication, user profile authorization, and accounting. Additionally, we also studied and evaluate po-

tential issues such as network security, and additional signaling overhead.

- An extensive analysis of trust relationships and roles of the different elements of our architecture. We introduced a trust infrastructure that allows binding trust relations between the cellular operators and UWNs through a new element called the Roaming Broker.

- We studied the impact of AAA (*Authentication, Authorization, and Accounting*) mechanisms on the vertical handover process. Through this analysis we identified *the authentication and authorization issues* as a potential impairments for seamless roaming.

- We conceived a SIP-based Roaming Architecture to enable interoperability and seamless service mobility under heterogeneous multi-operator wireless environments.

- We enhanced the authentication and authorization process in our SIP-based roaming architecture through the utilization of distributed PKI and PMI mechanisms.

- Finally, we also proposed a technique that allows local validation (*in the VN*) of the user service profile (*authorization*) without compromising the security of the mobile user. Our method relies on SIP-embedded XML Attribute Certificates.

2.4 Thesis Overview

This thesis focuses on the roaming issue in heterogeneous multi-operator wireless environments and explicitly provides a solution in this subject matter. The structure of this thesis is organized as follows:

- Chapter 3 provides a general introduction to NGWNs and presents the issues and challenges of roaming in such environments. Moreover, we state why the current solutions in this area do not fulfill the requirements for a seamless heterogeneous roaming and underline the differences with our research work. Additionally, we concentrate on next generation mobile services. We analyze the issue of service mobility in heterogeneous wireless networks and present the driver technologies towards seamless service mobility. In this chapter, we also provide the rationale behind our roaming archictecture.

- Chapter 4 is dedicated to our first contribution. We present our SIP-based roaming architecture and define the architectural elements along the description of all their functionalities. A detailed description of the SIP-based roaming signaling protocol is also provided. Finally, we present an evaluation of our proposed architecture through computer simulation and a feasibility study we performed in a testbed platform.

- In chapter 5, we address important issues of our architecture: the trade-off in terms of signaling triangulation introduced by the RB and the enhancement of the authentication and authorization mechanisms. An important contribution of our research work is also presented in this chapter, the SIP-embedded XML Attribute Certificates.
- Finally, chapter 6 concludes this thesis and provides an overview of the perspectives and future work in this research area.

The complete bibliography and the terminology index are found at the end of the document. Implementation details on concepts and testbeds are found in the appendices of this thesis.

Chapter 3

Heterogeneous Wireless Inter-working

A goal of NGWNs (*Next Generation Wireless Networks*) has been the support of seamless migration of applications and services across heterogeneous wireless networks. Nevertheless, we currently observe that although service transparency is important for mobile environments, wireless inter-working between cellular and UWNs has not been fully achieved. Manual setups, the reluctance of UWNs to accept external visitors, and the lack of service agreements between heterogeneous network operators prevent seamless service migration. In this respect, we have identified a research area to contribute to the enhancement of heterogeneous roaming.

Currently, the literature provides interesting approaches concerning service mobility under heterogeneous wireless environments. In this chapter, we present the access technologies and the wireless inter-working architectures that make wireless inter-operability possible. In addition, we present the specific case of heterogeneous multi-operator wireless roaming and explain why current solutions are not enough to completely address this issue. Furthermore, we also state the differences between current solutions and our research work.

The outline of this chapter is organized as follows: section 3.1 provides an outlook of existing wireless access technologies. Section 3.2 provides an introduction to existing wireless inter-working architectures and highlight their advantages and trade-offs. Section 3.3 explains the issues and challenges of service mobility in heterogeneous multi-operator wireless environments and states why existing solutions have not fully addressed such issues. In section 3.4, we provide our vision towards a seamless heterogeneous multi-operator roaming. Section 3.5 states the underlying assumptions upon which our research work was based and finally, section 3.6 concludes this chapter.

3.1 Wireless Access Technologies

There is no doubt that wireless networking plays an important role in the new information society. With the use of wireless networks, information can be sent overseas easily, quickly, and efficiently. Certainly, these and other factors have

impacted on the continuous development and rapid adoption of wireless access technologies. We rely on wireless technologies on daily basis, people use their cellular phones to communicate with one another, we transmit information through the use of satellites, etc. Moreover, services such as TV, radio, and GPS (*Global Positioning System*) are also examples of wireless communications.

Every day people and businesses rely on WLANs (*Wireless Local Area Networks*) to exchange information in a small office or a large corporate business. At the same time, developing countries rely on wireless technologies as a cheaper alternative to address the *last mile problem* and get connection to the Internet.

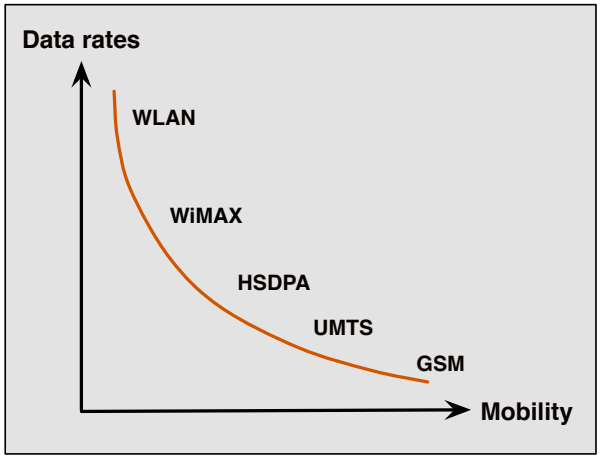


Figure 3.1: Mobile Broadband Wireless Access

Wireless networks free people from messy and expensive cable-based solutions, however each wireless technology comes with its own advantages and trade-offs. The main trade-off of wireless communications relates to data rates and mobility, the higher the mobility the lower the data rates and vice versa, as depicted by Figure 3.1.

In this section, we present the most popular wireless access technologies and describe their technical characteristics in terms of coverage and data rates. For simplicity, we divided this section in: WLANs, WMANs (*Wireless Metropolitan Area Networks*), and 3G cellular networks.

3.1.1 Wireless Local Area Networks

We consider that the success of WLANs relates to their convenience, their cost efficiency, the ease of deployment, and the relatively rapid integration with other networks and devices. These networks provide easy and efficient access to network resources as they do not require cabling and complex infrastructure. Thus, infrastructure-based wireless architectures only require wireless access interfaces in mobile devices and a wireless access point to operate, compared to wired networks that introduce additional costs and complexity due to the deployment of physical cables. Moreover, mobile users can access the Internet from outside of

their current network environment (*though, this particularity is restricted to a coverage area $<$ 150 meters for 802.11b and $\approx$ 150 meters for 802.11n in outdoor environments*). Finally, WLANs are relatively easier to scale than wired networks, where additional clients mean additional wiring.

Certainly, in addition to the multiple technical advantages offered by wireless networks, we also find some trade-offs that must be taken into account before deploying a wireless solution. Security in WLANs is a delicate issue, the use of a shared medium such as the wireless spectrum allows other people to intercept the transmitted signals in a relatively "easy" way. In wired networks, any potential eavesdropper would first have to physically break into the place where the network is located and tapping into the actual wires, this is clearly not an issue with wireless information. To overcome this security issue, WLAN standards rely on encryption technologies to assure the privacy of the data i.e (*WPA Wireless Protected Access*). Other encryption mechanisms such as WEP (*Wired Equivalent Privacy*) are known to have weaknesses that a well trained eavesdropper can compromise (Cam-Winget et al., 2003).

The trade-off between mobility and data rates is also present in WLANs, being the coverage area an important constraint of WLANs. For example the typical range of IEEE 802.11g-based networks is in the order of tens of meters, a condition that can be improved by adding directional antennas under line-of-sight conditions. Notwithstanding, WLAN's data rates are reasonably slow compared to the data rates currently supported by wired networks (100 *Mbps up to several Gbps*). In spite of the several attempts to provide efficient solutions to address the aforementioned trade-offs, we consider that the WLAN landscape is dominated by few wireless standards. In this section, we present the two most popular WLAN technologies to provide wireless connectivity: the IEEE 802.11 and HiperLAN/2.

3.1.1.1 The IEEE 802.11

The IEEE 802.11 working group was founded in 1987 to begin standardization of WLANs for use in the ISM (*Industrial, Scientific and Medical*) band. In 1997 IEEE 802.11 was finally standardized and provided interoperability standards for WLAN manufacturers using the same configuration (*11 Mcps DS-SS spreading and 2 Mbps data rates*). Commonly known as WiFi, the IEEE 802.11 specifies physical and data link layer as well as two configurations for wireless networks: PCF (*Point Coordination Function*) and DCF (*Distributed Coordination Function*) (IEEE 802.11 Standard, 1999). PCF uses an access point to coordinate medium access and packet routing. The access point polls each node in the network for packet transmission. If a polled node has a packet to send, this node is granted channel access and can transmit the packet within an interval of time. This access technique is known as contention free because the nodes do not have to compete for the channel thus collisions are non- existent. In DCF, the nodes use CSMA/CA (*Carrier Sense Multiple Access / Collision Avoidance*) combined with RTS (*Request to Send*) / CTS (*Clear to Send*) to overcome the hidden terminal problem. The principal characteristic of this kind of network is that the nodes have to compete for the channel using CSMA/CA and different spacing intervals called: DIFS (*DCF Inter-Frame Spacing*) and SIFS (*Short Inter-Frame Spacing*), depending on the priority of

Table 3.1: The IEEE 802.11 Roadmap

Protocol	**Release Date**	**Frequency**	**Throughput**	**Max. Data Rate**
Legacy	1997	2.4-2.5 GHz	1 Mbps	2 Mbps
802.11a	1999	5.1-5.8 GHz	25 Mbps	54 Mbps
802.11b	1999	2.4-2.5 GHz	6.5 Mbps	11 Mbps
802.11g	2003	2.4-2.5 GHz	20 Mbps	54 Mbps
802.11y	March 2008 (est.)	3.65-3.7 GHz	25 Mbps	54 Mbps
802.11n	March 2009 (est.)	2.4 and/or 5 GHz	74 Mbps	248 Mbps

the transmission.

The physical layer specifies the transmission medium and the multiple access technique that must be used. IEEE 802.11 operates in the ISM frequency band, which comprises 902-928 MHz, 2.4-2.483 GHz, and 5.725-5.825 GHz. This standard supports DS-SS (*Direct Sequence Spread Spectrum*) or FH-SS (*Frequency Hopping Spread Spectrum*) as multiple access technologies. Data rates from 1 to 2 Mbps are supported in the original standard. In 1999, the 802.11 High Rate standard called IEEE 802.11b was approved, thereby providing new data rates of 11 Mbps and 5.5 Mbps in addition to the original 2 Mbps and 1 Mbps rates (*in 2.4 GHz band*). Data rates of 54 Mbps are achieved in IEEE 802.11a in the 5 GHz band and in IEEE 802.11g in the 2.4 GHz band

The evolution of the IEEE 802.11 has been mainly driven by the industry, with the aim of providing higher throughput and data rates and increased coverage area. In this respect, Table 3.1 depicts the evolution of the IEEE 802.11.

802.11a

The 802.11a uses the same core protocol as the original standard, operates in 5 GHz frequency band, and uses a 52 sub-carrier OFDM (*Orthogonal Frequency Division Multiplexing*) with a maximum theoretical data rate of 54 Mbps, which provides realistic net achievable throughput in the order of 20 Mbps, as described in Table 3.1. The data rate is reduced to 48, 36, 24, 18, 12, 9 then 6 Mbit/s if required. 802.11a originally had 12/13 non-overlapping channels, 12 that can be used indoor and 4/5 of the 12 that can be used in outdoor point to point configurations.

Since the 2.4 GHz band is heavily used, using the 5 GHz frequency band provides 802.11a a significant advantage in terms of interference. Nevertheless, this high carrier frequency also brings a slight trade-off: the effective range of 802.11a

is slightly less than that of IEEE 802.11b/g; as 802.11a signals cannot penetrate as far as those for 802.11b because they are absorbed more readily by walls and other solid objects in their path. On the other hand, OFDM has fundamental propagation advantages when in a high multipath environment, such as an indoor office, and the higher frequencies enable the building of smaller antennae with higher RF system gain which counteract the disadvantage of a higher band of operation.

802.11b

The IEEE 802.11b standard was released in 1999. This technology supports a maximum theoretical data rate of 11 Mbps and a realistic achievable throughput of 6.5 Mbps. 802.11b uses the same CSMA/CA media access method defined in the legacy standard.

802.11b is normally used in point-to-multipoint configuration, where an access point communicates via an omni-directional antenna with one ore more clients that are located within the coverage area around the access point. Typical indoor range is 30 meters at 11 Mbps and 90 meters at 1 Mbps. The overall bandwidth is dynamically shared across all the users on a channel. With high-gain antennas, the protocol can be also used in fixed point-to-point communications, achieving ranges up to 10 kilometers where line-of-sight can be established.

802.11g

The 802.11g standard was ratified in June 2003. This technology defines a third modulation standard that operates in the 2.4 GHz frequency band, offering a maximum theoretical data rate of 54 Mbps, or about 20 Mbps net throughput. A great advantage is that 802.11g hardware is backwards compatible with 802.11b hardware. However, in a 802.11g network, the presence of a single 802.11b participant does significantly reduce the speed of the overall 802.11g network.

The modulation scheme used in 802.11g is OFDM for the data rates of 6, 9, 12, 18, 24, 36, 48, and 54 Mbit/s, and reverts to CCK (*like the 802.11b standard*) for 5.5 and 11 Mbit/s and DBPSK/DQPSK+DSSS for 1 and 2 Mbit/s. Even though 802.11g operates in the same frequency band as 802.11b, it can achieve higher data rates because of its similarities to 802.11a. The maximum range of 802.11g devices is slightly greater than that of 802.11a devices.

802.11n

The 802.11n builds upon previous 802.11 standards by incorporating MIMO (*Multiple Input Multiple Output*) support. MIMO uses multiple transmitter and receiver antennas to allow for increased data throughput via spatial multiplexing and increased range by exploiting the spatial diversity. Antennae are listed in a format of 2x2 for two receivers and two transmitters. A 4x4 is four receivers and four transmitters. The number of antennae relates to the number of simultaneous streams. The standards requirement is a 2x2 with two streams. The standard does optionally allow for the potential of a 4x4 with four streams. This standard can

operate in the 2.4 and/or 5 GHz frequency bands, offering a maximum data rate of 284 Mbps with the 2x2 configuration and a coverage area up to 150 meters.

802.11e

The 802.11e, release in 2005, defines a set of Quality of Service enhancements for LAN applications, in particular the 802.11 WiFi standard. 802.11e is considered of critical importance for delay-sensitive applications, such as Voice over Wireless IP and Streaming Multimedia. The protocol enhances the IEEE 802.11 Media Access Control layer through a new coordination function: the HCF (*Hybrid Coordination Function*). Within the HCF, there are two methods of channel access, similar to those defined in the legacy 802.11 MAC: HCCA (*HCF Controlled Channel Access*) and EDCA (*Enhanced Distributed Channel Access*. Both EDCA and HCCA define Traffic Classes. For example, emails could be assigned to a low priority class, and Voice over Wireless LAN could be assigned to a high priority class. With EDCA, high priority traffic has a higher chance of being sent than low priority traffic: a station with high priority traffic waits a little less before it sends its packet, on average, than a station with low priority traffic. On the other hand, HCCA is generally considered the most advanced (*and complex*) coordination function. With the HCCA, QoS can be configured with great precision. QoS-enabled stations have the ability to request specific transmission parameters i.e. *data rate, jitter, etc.* which should allow advanced applications like VoIP and video streaming to work more effectively on a Wi-Fi network.

3.1.1.2 HiperLAN/2

High Performance Local Area Network (*HiperLAN*)/2 is the European standard for WLANs. This network operates in 5.7 and 17.1 GHz providing data rates of 1 to 20 Mbps. Mobility is considered and it can operate up to vehicle speeds of 35 km/hr. HiperLAN/2 has emerged as the next generation of wireless network in Europe and will provide up to 54 Mbps to a variety of networks. HiperLAN/2 supports ad-hoc network mode to meet the requirements needed in future generation wireless networks.

The physical layer of HiperLAN/2 is very similar to IEEE 802.11a wireless local area networks (ETSI BRAN, 2000b). However, the media access control (*the multiple access protocol*) is Dynamic TDMA (*Time Division Multiple Access*) in HIPERLAN/2, while CSMA/CA is used in 802.11a (ETSI BRAN, 2000a). In this context, basic services in HIPERLAN/2 are data, sound, and video transmission with an emphasis in the quality of these services (*Quality of Service*). HiperLAN/2 standard covers physical, data link control and convergence layers. Convergence layer takes care of service dependent functionality between DLC and network layer. Furthermore, convergence sublayers can be used also on the physical layer to connect IP, ATM (*Asynchronous Transfer Mode*) or 3G cellular networks. This feature makes HIPERLAN/2 suitable for the wireless connection of heterogeneous networks. Good security measures are offered by HIPERLAN/2 as the data are secured with DES or Triple DES algorithms (NBS, 1977). Additionally, the access

point and the wireless terminal can authenticate each other.

Some people consider that IEEE 802.11-based wireless networks have already occupied the niche that HiperLAN/2 was designed for, albeit with lower performance but higher market penetration, and that the *network effect* of existing deployment is preventing the adoption of HiperLAN/2.

3.1.2 Wireless Metropolitan Area Networks

A WMAN is a wireless communication network that extends the WLAN coverage to a larger area (*upto 50 or 60 miles*) such as a suburb or an entire city. Typically, such networks are used to provide broadband Internet access to fixed or mobile devices. Subscriber stations communicate with base-stations that are connected to a core network. Among the wireless technologies used to provide a metropolitan-scale access solution, we find: the IEEE 802.16, HiperMAN, and WiBro.

3.1.2.1 The IEEE 802.16

A well-known example of long distance wireless technologies is the IEEE 802.16 commonly known as WiMAX. This technology is defined by the WiMAX Forum, formed in June 2001 to promote conformance and interoperability of the IEEE 802.16 standard, officially known as WirelessMAN (*Wireless Metropolitan Area Network*). This technology aims at the provisioning of fixed and mobile data over long distances (*in km*), in several ways, from point-to-point links to full mobile cellular-type access. In this respect, WiMAX can be seen as a standards-based technology that enables last mile wireless broadband access.

Originally, the IEEE 802.16 standard (IEEE 802.16-2001, 2001) defined the 10 to 66 GHz band range for communications, however this was later updated and added the specifications for the 2 to 11 GHz bands (*3.5, 2.3/2.5, or 5 GHz are the most commonly used*) (IEEE 802.16 Standard, 2004). For multiple access, IEEE 802.16d relies on OFDM with 256 sub-carriers (IEEE 802.16 Standard, 2004) and IEEE 802.116e uses SOFDMA (*Scalable OFDMA*) (IEEE 802.16-2005, 2005). The support of MIMO antenna technology provides additional benefits in terms of coverage, power consumption, frequency reuse, and bandwidth efficiency hence supporting better QoS levels.

In WiFi the media access controller (*MAC*) uses contention access whereas WiMAX relies on a sophisticated scheduling algorithm for which the subscriber station need to compete once for initial entry into the network. After the access is gained, the base station allocates an access slot for the subscriber. The time slot can become larger or shorter however it remains assigned to the subscriber station for all the time it stays in the network, that means that other subscribers cannot use it.

The future of WiMAX is quite promising as it has been accepted as IP-OFDMA for inclusion as the sixth wireless link under the IMT-2000 (*International Mobile Telecommunication*) (ITU-R WP8F, 2006). Furthermore, the goal for the long term evolution of both WiMAX and LTE (*Long Term Evolution*) is to achieve 100 Mbit/s

mobile and 1 Gbit/s fixed-nomadic bandwidth as set by ITU (*International Telecommunications Union*) for 4G NGMN (*Next Generation Mobile Network*) systems through the adaptive use of MIMO-AAS (*Adaptive Antenna Systems*) and smart, granular network topologies. In this connection, 3GPP LTE and WiMAX-m are concentrating much effort on MIMO-AAS, mobile multi-hop relay networking and related developments needed to deliver 10X and higher co-channel reuse multiples.

3.1.2.2 HiperMAN

HiperMAN (*High Performance Radio Metropolitan Area Network*) is a standard created by the ETSI (*European Telecommunications Standards Institute*) to provide wireless metropolitan area network communication in the 2-11 GHz bands (ETSI BRAN, 2002). Originally created to provide broadband wireless DSL (*Digital Subscriber Line*) services in the 3.5 GHz band, HiperMAN architecture is optimized for packet switched networks with support for both fixed and mobile applications. Nevertheless, it focuses primarily on delivering data services to residential and small business. Among other characteristics, HiperMAN aims at achieving to become an interoperable broadband fixed wireless access sytem. Thus, this standard uses the basic MAC (*DLC and CLs*) of the IEEE 802.16-2001 standard as it has being developed in very close cooperation with IEEE 802.16 (IEEE 802.16-2001, 2001). Another feature of this technology is the ability of providing various service categories, full QoS (*Quality of Service*) support, fast connection control management, strong security, fast coding adaptation, modulation and transmit power to propagation conditions and the support of non-line-of-sight operation.

Currently, the ETSI BRAN (*Broadband Radio Access Networks*) group has had numerous contacts with the WiMAX forum, expecting that the HiperMAN test specifications are used in the WiMAX certification testing. WiMAX and HiperMAN are expected to cooperate strongly to achieve the desired level of validation of test specifications.

3.1.2.3 WiBro

WiBro (*Wireless Broadband*) is the Korean name of the IEEE 802.16e (*mobile WiMAX*) standard (IEEE 802.16-2005, 2005), currently being developed by the Korean telecommunication industry. This wireless MAN technology aims at the provisioning of wireless broadband Internet access. To achieve this, WiBro relies on TDD (*Time Division Duplex*) as multiplexing technology and OFDMA (*Orthogonal Frequency Division Multiple Access*) for multiple access in a 8.75 MHz channel bandwidth. It is expected that WiBro base stations offer an aggregate throughput of 30 to 50 Mbps and provide a coverage area radius of 1-5 km. Moreover, WiBro provides mobility for high speed moving devices as it supports mobile environments up to 120 km/hr. QoS is also supported as WiBro defines three QoS classes: real time polling service (*MPEG video, video telephony*), non real time polling service (*web browsing, file transfer*), and best-effor traffic (*e-mail*).

WiBro-based solutions are being tested along the world, countries such as Korea, Italy, and Venezuela are planning to provide commercial service after test service around 2006-2007.

Table 3.2: Mobile Internet Wireless Access Technologies

Standard	**Family**	**Primary Use**	**Radio Access**	**Downlink**	**Uplink**
802.16e	WMAN	Mobile Internet	MIMO-SOFDMA	70 Mbps	70 Mbps
HIPERMAN	WMAN	Mobile Internet	OFDM	56.9 Mbps	56.9 Mbps
WiBro	WMAN	Mobile Internet	OFDMA	50 Mbps	50 Mbps
iBurst	WMAN	Mobile Internet	HC-SDMA	64 Mbps	64 Mbps
UMTS HDSPA	Cellular	Mobile phone	CDMA/FDD	384 Kbps - 14 Mbps	384 Kbps - 5.76 Mbps
UMTS-TDD	Cellular	Mobile Internet	CDMA/TDD	16 Mbps	16 Mbps

3.1.3 3G Cellular Networks

3G Systems are intended to provide a global mobility with wide range of services including telephony, paging, messaging, Internet and broadband data. The ITU (*International Telecommunication Union*) started the process of defining the standard for third generation systems, referred to as International Mobile Telecommunications 2000 (ITU, 2000a). The ETSI (*Europe European Telecommunications Standards Institute*) was responsible of UMTS standardisation process. In 1998, the 3GPP (*Third Generation Partnership Project*) was formed to continue the technical specification work. 3GPP has five main UMTS standardisation areas: radio access network, core network, terminals, services and system aspects and GERAN (*Generic Radio Access Network*) (3GPP Rel. 99, 2000) .

The competition for providing mobile Internet access is reflected in a cluttered wireless marketplace. 3G cellular systems such as CDMA2000 (*Code Division Multiple Access*) and UMTS (*Universal Mobile Telecommunication System*) compete with wireless MANs (*WiMAX, WiBro*) to offer better mobility and data rates for DSL-class Internet services. This competition has given as a results a plethora of wireless access networks with different technical characteristics, as illustrated in Table 3.2 (*the data rates presented in the table are theoretical maximums and may vary by a number of factors, including the use of external antennae, distance from the tower and the ground speed*). Thus, 3G systems are integrating these different access networks to enhance cellular services. Consequently, cellular standards are evolving to so-called 4G, high bandwidth, low delay, all-IP-based network core with the support of VoIP (*Voice over IP*) applications.

In this section, we present the cellular standard that already plays an important role towards heterogeneous wireless inter-working: the UMTS (*Universal Mobile Telecommunication System*).

3.1.3.1 The Universal Mobile Telecommunication System

UMTS is a 3G wireless telecommunication system primarily designed for mobile phone communications but also supporting data transmission at high data rates. Standardized by the 3GPP, UMTS is the the european answer to the ITU IMT-2000.

Although UMTS relies on W-CDMA (*Wideband-CDMA*) as underlying air interface (3GPP Rel. 99, 2000), this technology also supports TD-CDMA or TD-SCDMA air interfaces. In the UMTS architecture, the subscribers can expect data rates in the order of 384 kbps for R99 handsets and 3.6 Mbps for HSDPA (*High-Speed Downlink Packet Access*-enabled handsets (3GPP Rel. 5, 2002). The specific frequency bands originally defined by the UMTS standard are 1885-2025 MHz for uplink and 2110-2200 MHz for downlink. Nevertheless, they can change based on the country's spectrum regulations i.e. *in the US, the 1700 MHz band will be used instead of 1900 MHz for the uplink.*

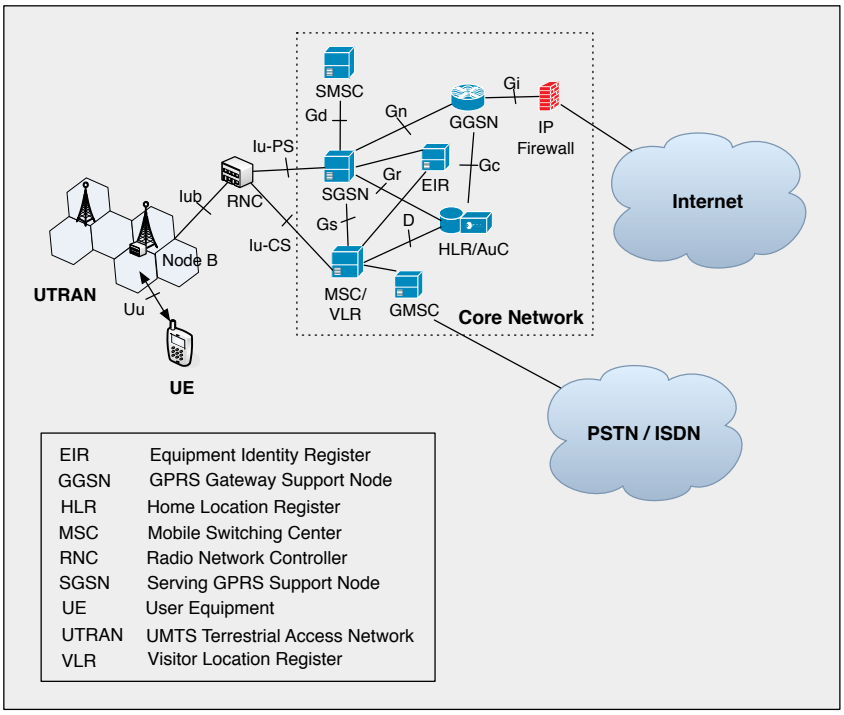


Figure 3.2: UMTS Network Architecture

UMTS is an evolution of 2G mobile phone systems such as GSM (*Global System for Mobile Communications*) and GPRS. A major innovation is the addition of the GERAN that allows the connection to various backbone networks such as the Internet, ISDN (*Integrated Services Digital Network*), GSM or another UMTS network. This functionality allows inter-network cooperation and the support of legacy mobile phone systems.

UMTS network services have different QoS classes for four types of traffic: conversational class (*voice, video telephony, video gaming*), streaming class (*mul-*

timedia, video on demand, webcast), interactive class (*web browsing, network gaming, database access*), and background class (*email, SMS, downloading*). Additionally, UMTS will also provide a VHE (*Virtual Home Environment*, which is a concept for personal service environment portability across network boundaries and between terminals (ETSI, June 1997). Personal service environment means that users are consistently presented with the same personalised features, User Interface customization and services in whatever network or terminal, wherever the user may be located. In this perspective, UMTS also has improved network security and location based services.

The UMTS network architecture, as depicted by Figure 3.2, comprises three interacting domains: the CN (*Core Network*), the UTRAN (*UMTS Terrestrial Radio Access Network*) and the UE (*User Equipment*). The main function of the core network is to provide switching, routing and transit for user traffic. Core network also contains the databases and network management functions.

The basic Core Network architecture for UMTS is based on GSM network with GPRS hence all equipment has to be modified for UMTS operation and services. The UTRAN provides the air interface access method for User Equipment. The Base Station is referred as Node-B and control equipment for Node-B is called the RNC (*Radio Network Controller*).

The CN is divided in circuit switched and packet switched domains. Some of the circuit switched elements are the MSC (*Mobile services Switching Centre*), the VLR (*Visitor Location Register*) and and the Gateway MSC. Packet switched elements are the SGSN (*Serving GPRS Support Node*) and the GGSN (*Gateway GPRS Support Node*). Additional network elements such as EIR (*Equipment Identity Register*), HLR, VLR and AUC (*Authorization Center*) are shared by both domains. The ATM (*Asynchronous Transfer Mode*) technology has been defined for UMTS core transmission. The architecture of the Core Network may change when new services and features are introduced. NPDB (*Number Portability DataBase*) will be used to enable user to change the network while keeping their old phone number. Furthermore, the GLR (*Gateway Location Register*) may be used to optimise the subscriber handling between network boundaries. Thus, the MSC, VLR and SGSN can merge to become a UMTS MSC.

The main functions of Node-B are: air interface transmission and reception, modulation and demodulation, CDMA physical channel coding, micro-diversity, error handling, and closed loop power control.

The functions of RNC are: radio resource control, admission control, channel allocation, power control settings, handover, control, macro-diversity, ciphering, segmentation and reassembly, broadcast signaling, and open loop power control.

The UMTS standard does not restrict the functionality of the UE in any way. As consequence, terminals might work as an air interface counter part for Node-B and have many different types of identities. Most of these UMTS identity types are taken directly from GSM specifications. The identity types currently supported by UMTS are the following: IMSI (*International Mobile Subscriber Identity*), TMSI (*Temporary Mobile Subscriber Identity*), P-TMSI (*Packet Temporary Mobile Subscriber Identity*), TLLI (*Temporary Logical Link Identity*), MSISDN (*Mobile station ISDN*), IMEI (*International Mobile Station Equipment Identity*), and IMEISN (*Inter-*

national Mobile Station Equipment Identity and Software Number). In addition, the UE can operate in one of three modes of operation:

1. PS/CS mode: the MS is attached to both the PS domain and CS domain, and the MS is capable of simultaneously operating PS services and CS services.
2. PS mode: the MS is attached to the PS domain only and may only operate services of the PS domain. However, this does not prevent CS-like services to be offered over the PS domain (*VoIP*).
3. CS mode: the MS is attached to the CS domain only and may only operate services of the CS domain.

UMTS IC card has same physical characteristics as GSM SIM card. It has several functions: the support of one USIM (*User Service Identity Module*) application, support of one or more user profile on the USIM, the update of USIM specific information over the air, security functions, user authentication, optional inclusion of payment methods, and optional secure downloading of new applications.

In a near future, the 3GPP Long Term Evolution project plans to move UMTS to 4G speeds of 100 Mbit/s down and 50 Mbit/s up, using a next generation air interface technology based upon OFDM. Currently, UMTS supports mobile video-conferencing, although experience has shown that user demand for video calls is not very high. Other potential uses for UMTS include the downloading of music and video content, as well as live TV.

3.2 Wireless Inter-working Architectures

Heterogeneous wireless inter-working refers to the integration and interoperability of wireless networks with different access technologies. Nevertheless, this integration cannot be fully accomplished by only adding different wireless interfaces under the same mobile terminal. The integration of heterogeneous wireless access network can be achieved in two difference scenarios, mainly determined by the integration architecture. In this respect, these scenarios can be classified in: *single operator* and *multi-operator*

The single-operator scenario provides a straightforward integration for heterogeneous wireless networks as the UWNs are connected to the cellular networks through a private IP-based packet switching network, this configuration is also known as (*tightly coupled*), as illustrated in Figure 3.3. In addition, single-operator scenarios also support the interconnection of UWN through a public IP network such as the Internet, following the *loosely coupled* approach. Generally, in a single-operator heterogeneous network the cellular operator deploys its own UWNs, however it can also establish contractual agreements wireless hotspot aggregators or WISPs (*Wireless Internet Service Provider*) to increase coverage. Another characteristic of this wireless integration architecture is that the cellular network architecture assumes a packet switching core network-based, thereby allowing a single operator to fully control the cellular and the UWNs. This approach allows, the cellular operator to provide new services supported by the UWN in addition to the traditional cellular services. Nowadays, it is easy to

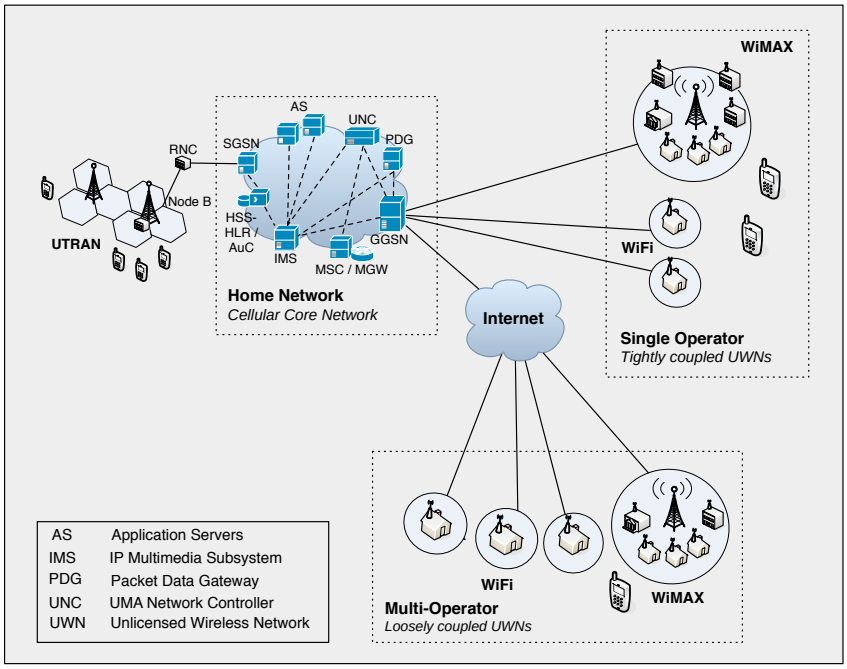


Figure 3.3: Service Mobility in Single and Multi-Operator Environments

find such scenarios as the top mobile operators in certain countries also own ISPs or WISPs.

Heterogeneous multi-operator wireless networks operate in most of the cases under a *loosely coupled architecture* as the UWNs are connected to the cellular network through an Internet backbone, as illustrated in Figure 3.3. Here, the UWNs are controlled by independent individuals and by no means are completely operated by the cellular network operator. We are aware that both single and multi-operator scenarios have their own advantages and trade-offs. Nevertheless, our research work focuses on loosely-coupled multi-operator scenarios, as there is an important number of UWN deployed by independent individuals that can be used to enhanced cellular services.

3.2.1 3GPP/WLAN TS. 23.234

The 3GPP TS. 23.234 defines an inter-working architecture to enable seamless cooperation between WLANs and 3G cellular networks with the objective of achieving the so-called *access independence* (3GPP TS 23.234, 2006) with minimal changes in current wireless architectures. An interesting concept proposed in the 3GPP/WLAN is the differentiation of *access authentication* and *service authentication*, in this integration architecture the authentication method does not influence the solution for service authentication. According to the 3GPP TS. 23.234 the 3GPP/WLAN main goals are the WLAN access authentication and authorization through the mobile core network using the AAA server and the HSS (*Home Subscriber Service*). To achieve this, this wireless integration architecture relies on

two elements: the IMS and the PDG (*Packet Data Gateway*), see Figure 3.4.

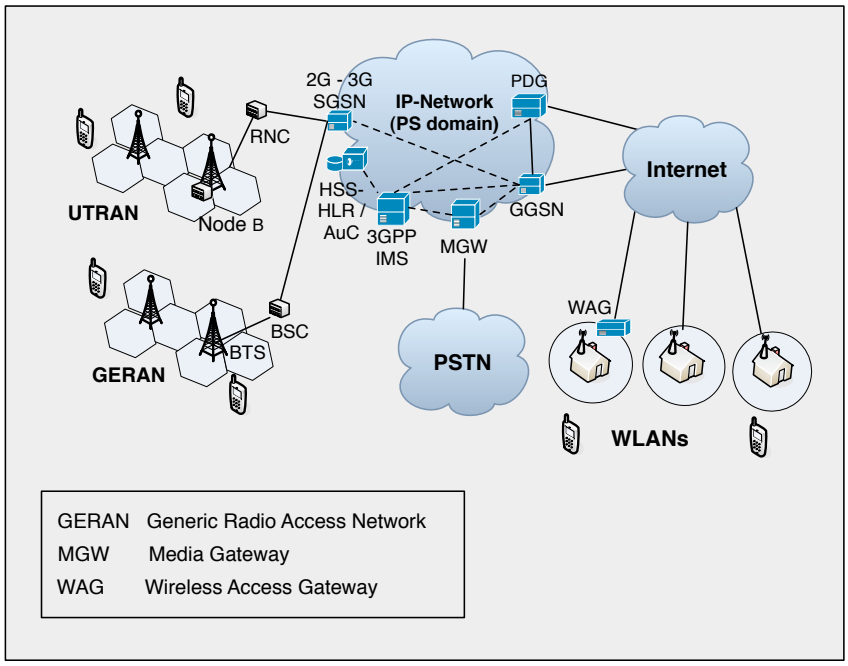


Figure 3.4: WLAN access to 3GPP IMS

Based on the specifications of the 3GPP TS 23.234 the IMS registration process through a WLAN, as illustrated in Figure 3.5, is the following: the UE entering into the WLAN coverage area initiates the WLAN association (*flow 1*). This process involves switching wireless networking interface and layer two protocols. Once the WLAN association accomplished, the UE attempts USIM-based secure access authentication in the WLAN's AAA server (*flow 2*). If the authentication is successful, the UE will be able to obtain a valid IP address, service authentication and authorization, and policy enforcement and charging. To to this, the UE queries the WLAN's DHCP (*Dynamic Host Configuration Protocol*) to obtain a valid IP address (*flow 3*). The next step is for the UE to update its location in the cellular network. Consequently, the UE retrieves the PDG's address by querying the WLAN's DNS (*Domain Name Server*) (*flow 4*). Then, the UE establishes a tunnel to the PDG through the Internet (*flow 5*). Upon the establishment of the tunnel, the UE obtains the remote P-CSCF's IP address and discovers the P-CSCF (*Proxy-Call Session Control Function*) which is the initial interface (*SIP server*) between the UE and the IMS (*flow 6*). Now, with the P-CSCF's IP address the UE establishes a security association between UE and IMS. Finally, the UE performs the IMS registration and service set-up.

Although the TS. 23.234 provided one of the first standardized cellular-wireless LAN inter-working, it assumed that the WLAN has full access to the mobile core network's databases. Thus, the WLAN's access point can verify the identity of the user and triggers the PDP (*Packet Data Protocol*) context activation process for further data transmission from the WLAN. This configuration operates effi-

ciently in wireless integration scenarios where the cellular network deploys their own WLAN-based access network or are in full control of both wireless access networks.

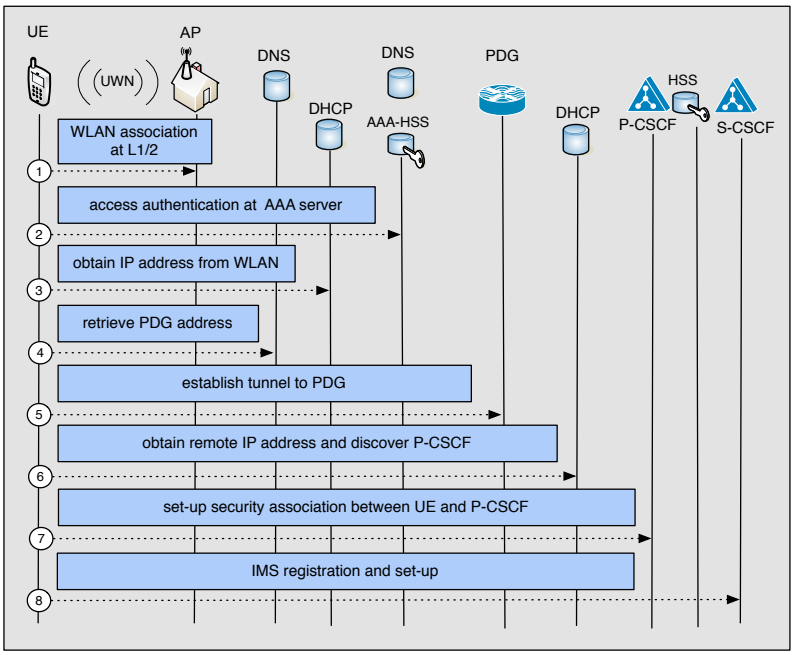


Figure 3.5: IMS registration through WLAN

3.2.2 Unlicensed Mobile Access

UMA (*Unlicensed Mobile Access*) is the 3GPP standard for FMC. This technology enables access to mobile services such as voice, data, and IMS services over IP broadband access and UWN technologies (3GPP Rel. 6, 2004) . By deploying UMA technology, service providers and network operators can offer seamless roaming or handover between heterogeneous wireless networks using UMA-enabled, dual-mode mobile handsets.

UMA wireless inter-working, as depicted in Figure 3.6, is possible through the addition of a new element called the UNC (*UMA Network Controller*) and associated protocols that provide the secure transport of GSM/GPRS signaling and user plane traffic over IP. Consequently, the UNC interfaces into the mobile core network via the existing 3GPP A/Gb interfaces. UMA-based roaming comprises the following steps:

1. A mobile user with a UMA-enabled, dual-mode handset enters within the coverage of an UWN to which the handset is allowed to connect.

2. Upon establishing the connection, the handset contacts the UNC (*UMA Network Controller*) over the IP access network to be authenticated and authorized to access GSM voice and GPRS data services via the UWN.

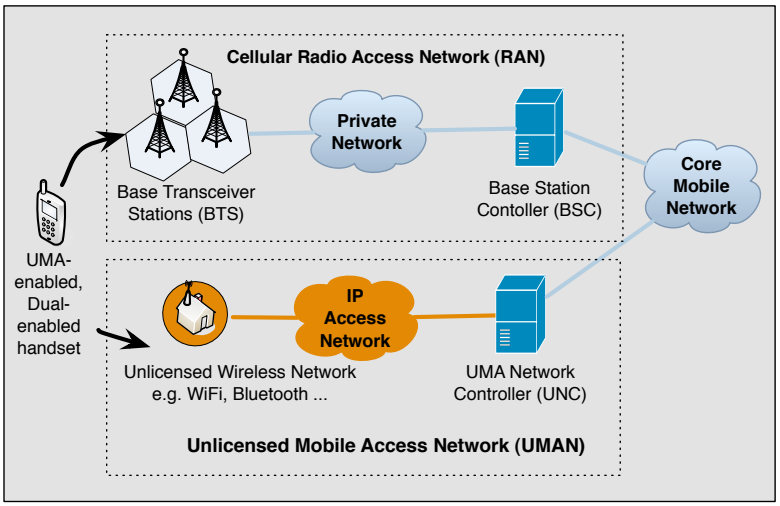


Figure 3.6: UMA technology

3. If approved, the mobile user's current location information stored in the core network is updated, and from that point on all mobile voice and data traffic is routed to the handset via the UMAN (*Unlicensed Mobile Access Network*) rather than the cellular radio access network (*RAN*).

UMA technology provides a clear definition for roaming and handover. In this connection, roaming is performed when UMA-enabled mobile users move outside of the range of an UWN to which they are connected hence they roam back to the cellular network or vice versa. On the other hand, handover occurs when a subscriber is on an active GSM voice call or GPRS data session when they come within range (*or out of range*) of an UWN, the voice call or data session are automatically handover between different access networks with no discernible service interruption.

3.2.2.1 The Generic Access Network

The GAN (*Generic Access Network*) was formerly known as UMA, until its adoption by the 3GPP in April 2005. This access network provides a telecommunication system that allows seamless roaming and handover between WLANs and cellular networks using a dual-mode mobile phone (3GPP Rel. 6, 2004). Nevertheless, the term GAN remains unknown outside the 3GPP community, and the term UMA continues to be used.

GAN allows cellular network operators to deliver voice, data, and IMS/SIP-based applications to cellular phones on wireless local area networks. Its main goal is to enable the 3GPP/WLAN inter-working to achieve efficient convergence of mobile, fixed, and Internet telephony. On the cellular network, the UE communicates over the air with a base station (*Uu interface*), through a BSC, to servers in the mobile core network. In the GAN system, when the UE detects a WLAN, it establishes a secure IP connection through the Internet to a server called a GANC

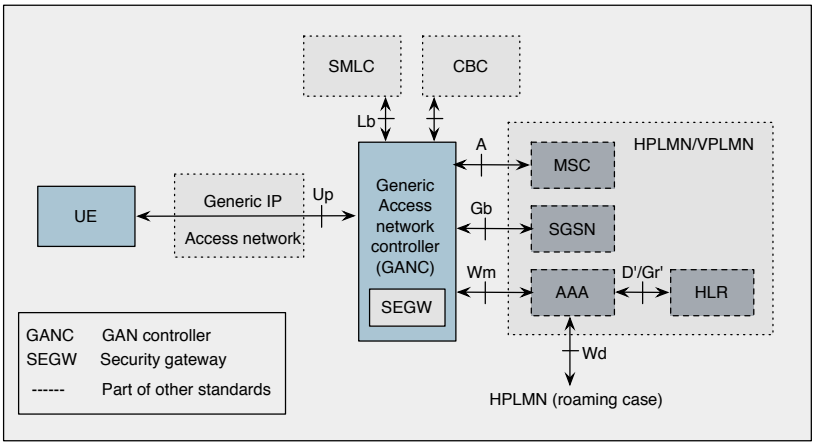


Figure 3.7: Architecture of the Generic Access Network (GAN)

(*GAN Controller*) on the cellular network. The GANC translates the signals coming from the UE to make it appear to be coming from another BTS. Thus, when a mobile moves from a GSM to a WiFi network, it appears to the core network as if it is simply on a different BTS.

The main objective of GAN is that the UE seamlessly establishes communication with the GANC through the WLAN. As illustrated in Figure 3.8, such transparency is achieved through the use of the Up interface, which is the core of the standard. In this respect, interconnection between the WLAN and the GAN through a IP network is required, since the UE exchanges IP packets with the GANC over the Up interface.

The GANC interacts with the mobile core network through the A and Gb interfaces in the same way that BSCs (3GPP TS 44.318, 2005). Thus, to the mobile core network, the GANC looks like any other BSC in the cellular network. Additionally, an IPsec tunnel between the UE and the SEGW (*Security Gateway*) protects the information exchanged over the Up interface. The SEGW uses the Wm interface for authentication, authorization, and accounting purposes (*in the AAA server in the mobile core network*). A subset of the Wm interface is used to authenticate mobile users (*through SIM or USIM credentials*) once the IPsec tunnel is established. IETF specifications define protocols for this purpose, they include IKEv2, EAP-SIM, and EAP-AKA. The GAN standard defines functions and procedures required in the Up interface to support seamless mobility (*handover and roaming*) between the GAN and GSM and between the GAN and WCDMA, and to provide access to services in the mobile core network (3GPP Rel. 6, 2004).

In this perspective, the GAN standard defines the term *rove in* to describe roaming between WiFi and GERAN/UTRAN. *Rove-in* means the UE has initiated communication through the Up interface. On the other hand, *rove out* indicates that the UE has stopped communication with the Up interface. Consequently, relevant GERAN/UTRAN protocols are used to serve the upper OSI (*Open System Interconnection*) layers in the UE.

Figure 3.8 illustrates an example of GAN *rove in*, that is, of a UE entering and

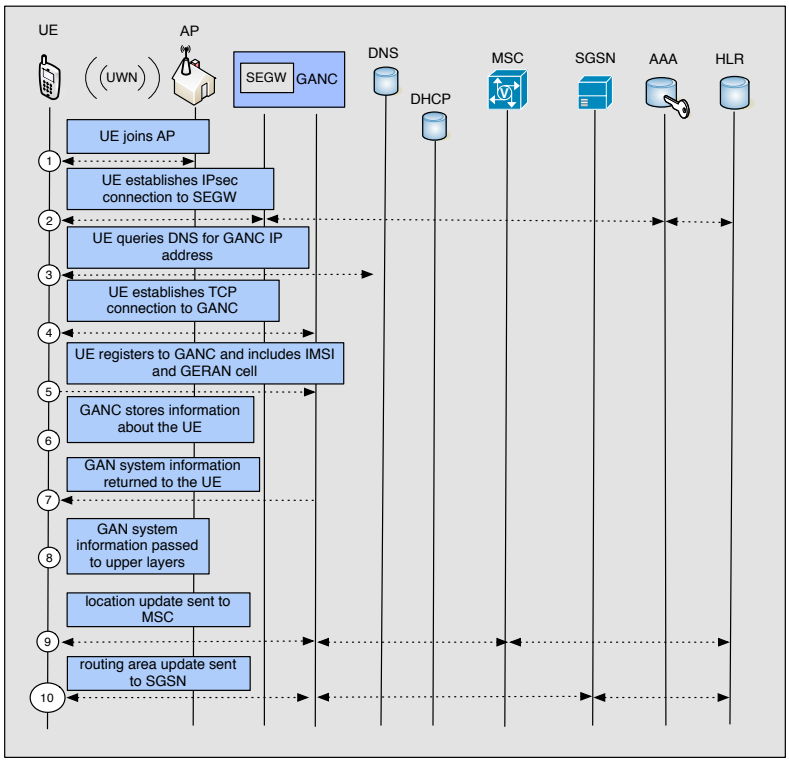


Figure 3.8: Rove-in example

registering into a serving GANC. The process is as follows:

1. The UE joins the AP to gain IP connectivity.
2. The UE establishes an IPsec tunnel through the SEGW and the Up interface. Here, the SEGW also authenticates the UE through the Wm interface to the AAA server in the mobile core network.
3. The UE queries a DNS to obtain the GANC IP address.
4. The UE establishes a TCP connection to the GANC.
5. The UE registers to GANC and provide its credentials, among the information it provides we find the IMSI and the GERAN cell identifier.
6. The GANC accepts registration and stores all necessary information.
7. The GANC informs the UE that it has accepted the registration and transmits GAN system information.
8. The UE passes relevant system information to its upper layers.
9. The upper layers in the UE initiate location area update in the MSC.
10. The upper layers in the UE initiate a routing area update in the SGSN.

Once the step ten is reached the UE has roved in to the WLAN and will be able to initiate services through the Up interface. Thus, the UE will be also reachable through the same interface from another UE.

Currently, GAN architecture has been successfully and commercially deployed in England, France, Poland, Spain, The Netherlands, and Italy. Not surprisingly, in most of these cases, the cellular network operator also owns or manage an ISP, this facilitates the interaction between the WLAN and the mobile core network. UMA technology relies on a trusted WLAN or an open WLAN to initiate the *rover in* process. Unfortunately, under heterogeneous multi-operator wireless environments UMA-enabled UEs will face the network access problem, hence they will not be able to roam freely into such networks.

3.2.3 Ambient Networks Project

The IST Ambient Network Project is an integrated project co-sponsored by the European Commission within the Sixth Framework Programme (*FP6*). This project presents a network integration solution to the modern-day problems of switching from one network to the other in order to keep in contact with the outside world (Ahlgren et al., 2005). In short, it aims to develop a network *software-driven* infrastructure that will run on top of all current or future network physical infrastructures to provide a way for devices to connect to each other, and through each other to the outside world, as illustrated in Figure 3.9.

The project works at a new concept called Ambient Networking, to provide suitable mobile networking technology for the future mobile and wireless communications environment. Ambient Networks provides a unified networking

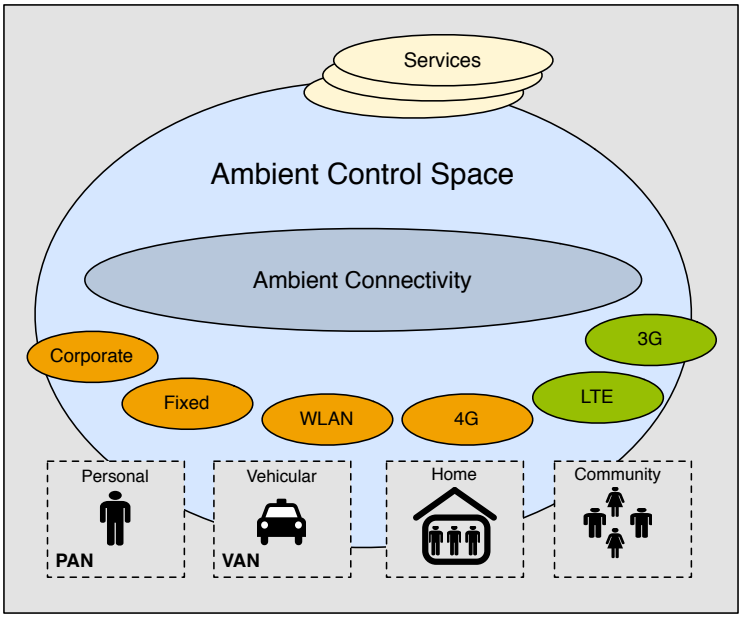


Figure 3.9: The Ambient Network Idea

concept that can adapt to the very heterogeneous environment of different radio technologies and service and network environments. Special focus is put on facilitating both competition and cooperation of various market players by defining interfaces, which allow the instant negotiation of service agreements. This approach goes clearly beyond interworking of well-defined protocols and is expected to have a long-term effect on the business landscape in the wireless world (Ahlgren et al., 2005). Central to the project is the concept of composition of networks, which is an approach to address the dynamic nature of the target environment. The approach is based on an open framework for network control functionality, which can be extended with new capabilities as well as operating over existing connectivity infrastructure. The current development phases of the AN project are:

- Phase 1 of the project (*2004-2005*) has laid the conceptual foundations. The Deliverable D1-5 *"Ambient Networks Framework Architecture"* summarizes the work from phase 1 and provides links to other relevant material (Project, 2005).
- ANs Phase 2 (*2006-2007*) focuses on validation aspects. One key result of phase 2 is an integrated prototype that will be used to study the feasibility of the AN concept for a number of typical network scenarios. The Ambient Control Space (*ACS*) prototype will be used to iteratively test the components developed by the project in a real implementation. In parallel, the top-down work is being continued which will lead to a refined System Specification. Furthermore, standardization of the composition concept is addressed in 3GPP (3GPP TR 22.980, 2006).

The main characteristics of Ambient Networks (*AN*) are:

- **all-IP network-based:** the concept of Ambient Network is based on all-IP-based wireless architectures. This assumption allows the clear separation between transport and control tasks. In this context, the functions concerned with either of these tasks are divided in two layers to ease the independent development in both areas. Consequently, the AN concept adapts these premises and assumes the presence of a connectivity layer that assures basic connectivity between heterogeneous wireless access networks.
- **Heterogeneity:** AN are based on a federation of multiple networks with different operators and access technologies.
- **Mobility:** the AN mobility solution aims at providing service migration across business and administrative boundaries. To achieve this, a security platform for inter-domain security issues is required.
- **Self-Management:** AN elements and networks are able of managing and configuring themselves automatically with none or minimal human intervention.
- **Modularity:** AN can be dynamically composed of several networks to create a new AN. Such networks could tentatively belong to different administrative or economic entities. Thus, ANs will provide network services in a cooperative and a competitive way. This cooperation and modularity are feasible due to the AN interfaces.

3.2.3.1 Ambient Network Interfaces

The concept of Access Control Space is introduced to group all control functions in a certain network domain. In this perspective, the ACS provides a set of modular control functions that can coexist and interoperate. These functions include basic functionalities such as management, security, and connectivity. Moreover, ACS is based on *plug&play* concepts that enable bootstrapping and the dynamic discovery of sets of supported functions.

As illustrated in Figure 3.10, there are three interfaces present to communicate with an ACS as depicted by Figure 3.10, these are:

- ANI (*Ambient Network Interface*): if a network wants to join in, it has to do so through this interface.
- ASI (*Ambient Service Interface*): provides uniform access to the AN functionality from higher layers hence if a function needs to be accessed inside the ACS, this Interface is used.
- ARI (*the Ambient Resource Interface*): encapsulates the capabilities of the underlying infrastructure, and a flow level abstraction provides the basic building block of data transfer between network addresses. This interface is used when a resource inside a network needs to be accessed.

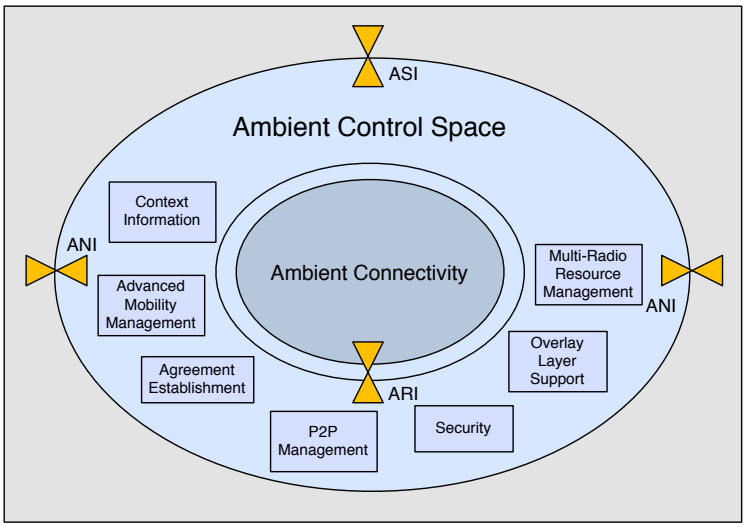


Figure 3.10: The Ambient Control Space

Interfaces are used in order to hide the internal structures of the underlying network. Consequently, If one network is discovered by another and they decide to merge, a new ACS will be formed of the two (*though the two networks will have their own ACS along with the interfaces inside this global, new ACS*). Thus, the newly composed ACS will of course have its own ANI, ASI and ARI, and will use these interfaces when merging with other ANs. Other options for composition are to not merge the two ANs, this is called *network inter-working*, or to establish a new virtual ACS that exercises joint control over a given set of shared resources, this technique is known as *Control Sharing*.

3.2.3.2 The Ambient Network Scenario

A typical AN scenario is the following: Alice has a PAN (*Personal Area Network*) composed by her bluetooth-enabled PDA (*Personal Digital Assistant*), a mobile phone, and a laptop. All the devices are turned on and forming a network. Her laptop has a WLAN interface, and her mobile provides GPRS (*General Packet Radio Service*) data access, though GPRS data rate is slower and much more expensive for Alice to use. She is now on her way to work, and her laptop is downloading her e-mails through the GPRS connection on the mobile.

$$laptop \rightarrow (bluetooth) \rightarrow mobile \rightarrow (GPRS) \rightarrow cellular\ network$$

While walking, she enters into a coffee-shop that is covered by a free WLAN hotspot. Her laptop discovers this *new* network and her PAN immediately initiates a connection through the WLAN. This is called *network merging* (*that of the hotspot and her PAN*) in the AN jargon. Once this merging is accomplished, her e-mail downloading seamlessly continues (*without noticing any change of network interface*). The difference is that now the e-mail is being downloaded using a faster

and cheaper wireless interface (*WLAN*). If she now wants to access the web with her PDA, her PDA will use the WLAN interface on the laptop.

$$PDA \rightarrow (bluetooth) \rightarrow laptop \rightarrow (WLAN) \rightarrow \textit{WLAN hotspot}$$

3.2.4 Application-Layer Mobility Management using SIP

Important part of research work includes the study of SIP-based roaming architectures for heterogeneous wireless inter-working. In this perspective, we carefully analyzed the approaches that address service mobility in heterogeneous networks from an application-layer perspective. Through their work, Schulzrinne and Wedlun introduce the concept of *Application-Layer Mobility* (Schulzrinne & Wedlund, 2000). In this perspective, they describe how SIP can be used to provide terminal, personal, session and service mobility to applications ranging from Internet telephony to presence and instant messaging. Furthermore, the work of (Wei et al., 2005) and (Banerjee et al., 2005) provides an interesting vision of SIP-based terminal mobility support in heterogeneous wireless environments.

3.2.4.1 Review of SIP

SIP is an application-layer control protocol (*signaling*) that allows the creation, modification, and termination of multimedia session with one or more participants. The media streams can be audio, video, or any other Inter-based communication mechanism i.e. *VoIP*. Currently, the protocol has been standardized by the IETF (Schulzrinne & Handley, 1999) and is being implemented by number of vendors, including cellular network operators (3GPP Rel. 5, 2002).

SIP end points are referred by SIP URLs that have the form of e-mail addresses, such as *alice@example.com*. SIP requests contain a source address and two destination addresses, with one identifying the original, logical destination of the request (*the To header*), and, in the request URI the current destination, see Figure 3.11 (*flow 1*). The current destination is determined by SIP routers, or *proxies* as explained below. SIP requests also indicate in the *Contact* header, where future request should be sent.

SIP defines four logical entities, namely user agents, redirect servers, proxy servers, and registrars. UAs (*User Agents*) originate and terminate requests. Typically, they are the only elements in the architecture where signaling and media converge. Redirect servers receive requests and return a response indicating where to send the next request. This server keeps track of user's location and returns a list of one or more SIP or other URLs indicating the current location of the host. A proxy server can be stateless or stateful. The former server simply forwards incoming requests to another server without ensuring the request's reliability. On the contrary, the later, maintains a state for a transaction i.e. *a request an all responses that belong to that transaction*. A stateful proxy can also fork a request i.e. *send copies of the request to different destinations*. This forking method is useful when proxy servers are not aware of the final destination of the request and need to try different alternatives.

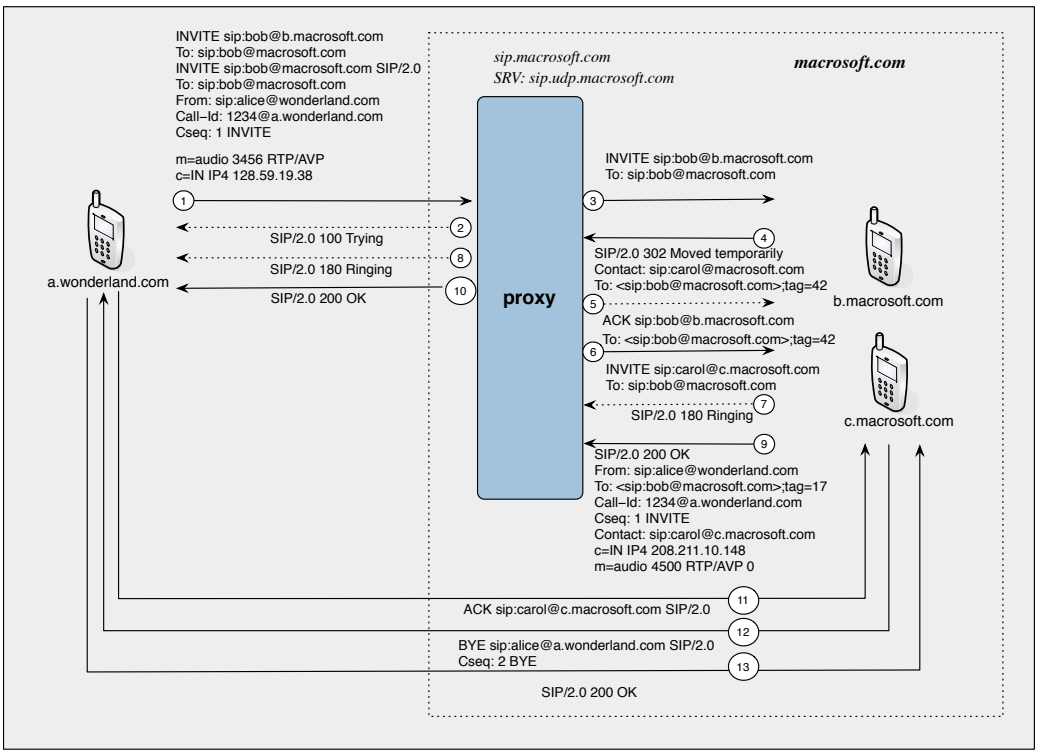


Figure 3.11: SIP session set-up, (Schulzrinne & Wedlund, 2000)

Traditionally, a SIP server implements a redirect and proxy server, with information provided by a built-in registrar. Depending on the configuration, and a specific request, the server can act as either a proxy or redirect server or a registrar. The SIP process, as illustrated in Figure 3.11 is as follows: consider the example of alice@wonderland.com request addressed to bob@macrosoft.com (*flow 1*). Typically, Alice would send all her requests to a local server at wonderland.com. The server in wonderland.com identifies that the request is not meant for it and forwards it to the server in macrosoft.com (*flow 3*), meanwhile it informs Alice that it is trying to contact the other part (*flow 2*). The SIP server in macrosoft.com first tries to contact Bob through his address bob@b.macrosoft.com, based on his registration on host b.macrosoft.com. Nevertheless, Bob is out of office and he is temporally forwarding his calls to Carol and thus has his Internet phone issuing a redirect response to the proxy server (*flow 4*). The proxy server without Alice's intervention, sends an invitation to Carol that is successfully accomplished (*flows 5 to 12*). The response contains the network address of Carol's computer, which is then used to directly exchange the acknowledgment and later tear down the call via the BYE request (*flow 13*).

3.2.4.2 Handling Terminal Mobility Using SIP

SIP is suitable to handle mobility due to its ease of deployment. SIP-based applications does not required to add complex network infrastructure nor the installation of home agents or dynamic location update techniques (Schulzrinne & Wedlund, 2000). Additionally, SIP provides the necessary support for *pre-call* and *mid-call* terminal mobility.

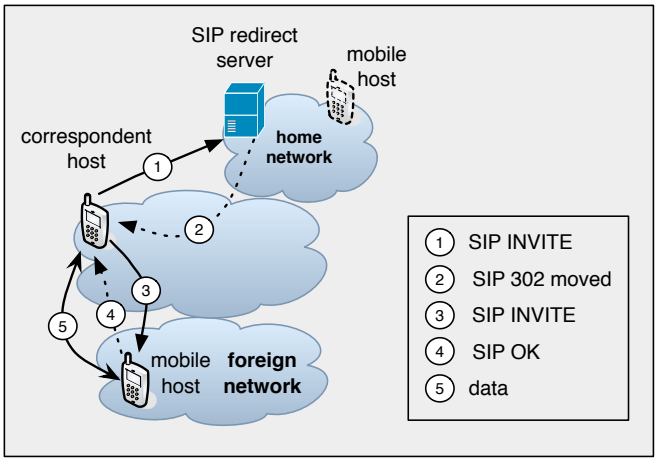


Figure 3.12: SIP-based pre-call terminal location, (Schulzrinne & Wedlund, 2000)

Terminal mobility requires SIP to establish a connection either during the start of a new session (*when the terminal has already moved to another location*), or in the middle of a session (*hand-over*). The former case is known as pre-call mobility, the latter as mid-call or in-session mobility. For pre-call mobility, the MH (*Mobile*

Host) updates its new IP address with its home network by sending a SIP REGISTER message as illustrated in Figure 3.12. On the other hand for mid-call mobility, the terminal requires to establish a relation with the CH (*Correspondent Host*) by sending an SIP INVITE message informing about the terminal's new IP address and updated session parameters. As depicted by Figure 3.13, the CH starts data transmission to the new location upon the reception of the SIP INVITE message. Thus, the location update takes one-way delay after the application in the MH recognizes that it has acquired a new IP address. For wideband access, the delay is probably equal to the propagation delay plus a few milliseconds, but narrowband systems may impose delays of several tens of milliseconds. This without taking in to account the delay introduced by the DHCP and AAA process.

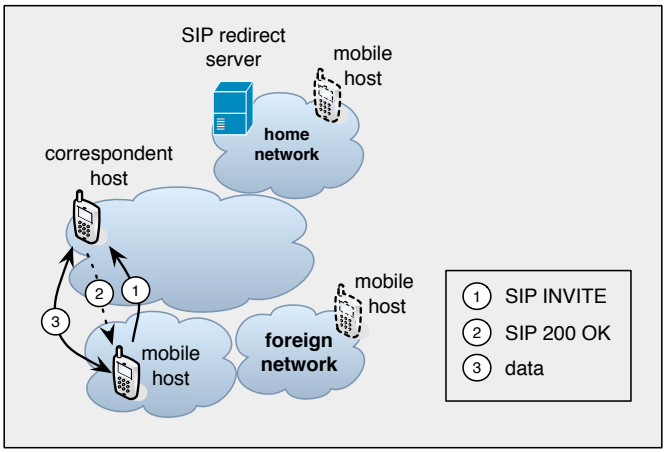


Figure 3.13: SIP-based hand-over in mid-call, (Schulzrinne & Wedlund, 2000)

SIP-based Mobility Management in 4G Wireless Network

In their work *SIP-based Mobility Management in 4G Wireless Network* (Banerjee et al., 2005), the authors take advantage of the application-layer protocol abstraction provided by SIP to support seamless mobility in next generation heterogeneous wireless networks (Banerjee et al., 2005). They propose an architecture that supports handover for IP-centric wireless networks while alleviating the problem of packet loss. In this context, the authors consider SIP suitable for handling mobility in heterogeneous wireless network due its transparency to lower-layer characteristics, ease of deployment, and greater scalability (Wei et al., 2005). Their integration architecture assumes a loosely-coupled approach due to its broader application. Figure 3.14 depicts a logical view of the author's WLAN-UMTS inter-working architecture. An important assumption is the selection of an all-IP mobile network core and the use of the Internet as an interface between the two access networks: UMTS and WLANs. Their architecture is largely based on the 3GPP standard Release 5, as they use GERAN as access network and GPRS

as access technology. Furthermore, the IMS plays an important role in this inter-working architecture as provide the elements and interfaces to support SIP-based services and applications.

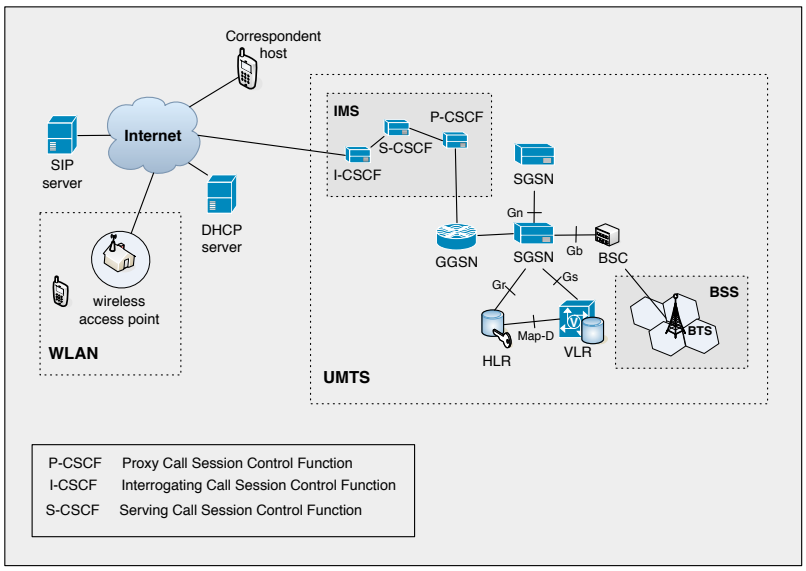


Figure 3.14: SIP-based 4G architecture

The alternative access technology in this wireless inter-working architecture is IEEE 802.11-based WLAN. As illustrated in Figure 3.14, the APs are connected to the Internet with an ethernet switch and a DHCP server is used to assign a dynamic IP address to the visiting UE. An important assumption of this proposal is that the UE roams into open APs as no authentication mechanism are proposed or evaluated. Thus, this proposal is not reliable to provide heterogeneous roaming support in UWNs with access under multi-operator environments, as under this scenario the UWNs are usually protected with network access control mechanism.

WLAN to UMTS Vertical Handover

Based on their integration architecture, when a UE attempts to roam from a UMTS to a WLAN network, it performs two key functions to trigger the handover process. Firstly, the UE initiates the data connection set-up, including the GPRS attach and the PDP context activation and lastly, the SIP message exchange that re-establishes the connection as depicted in Figure 3.15. As part of the GPRS Attach procedure, the UE transmits the attach request message (*flow 1*) to the SGSN. The SGSN uses the UE's IMSI to validate the identity of the UE with the HLR (*flows 2,3,4, and 5*). Successful authentication is followed by the SGSN sending a location update to the HLR (*flows 6 and 7*). A successful GPRS attachment is followed by the SGSN sending an *attach completed* message to the UE (*flow 8*). Thus, a logical association between the UE and the SGSN is established. The next

step now is to activate a PDP address (*or IP address*) to initiate data communication. The activation of such PDP address logically associates the UE's SSGN and GGSN. PDP context transfer is initiated by the UE transmitting the PDP content activation message (*flow 9*) to the SGSN. The SGSN, upon the reception of this message, discovers an appropriate GGSN that supports SIP-based applications and services (*flows 10 and 11*). Then, the SGSN and GGSN select special paths for the transfer of SIP messages to the P-CSCF. The P-CSCF's IP address is sent along with the activation accept message (*flows 12-16*). Finally, the UE enters into the SIP-exchange phase. Here, the UE invites the CH to its new temporary address by sending a SIP INVITE message (*flow 17*) through the CSCF servers. The SIP INVITE messages contains the same call identifier as in the original set-up and contains the new IP address at the new location in updated SDP content. Once the CH updates the location information about the UE, it sends a SIP OK message (*flow 18*) while starting data transmission.

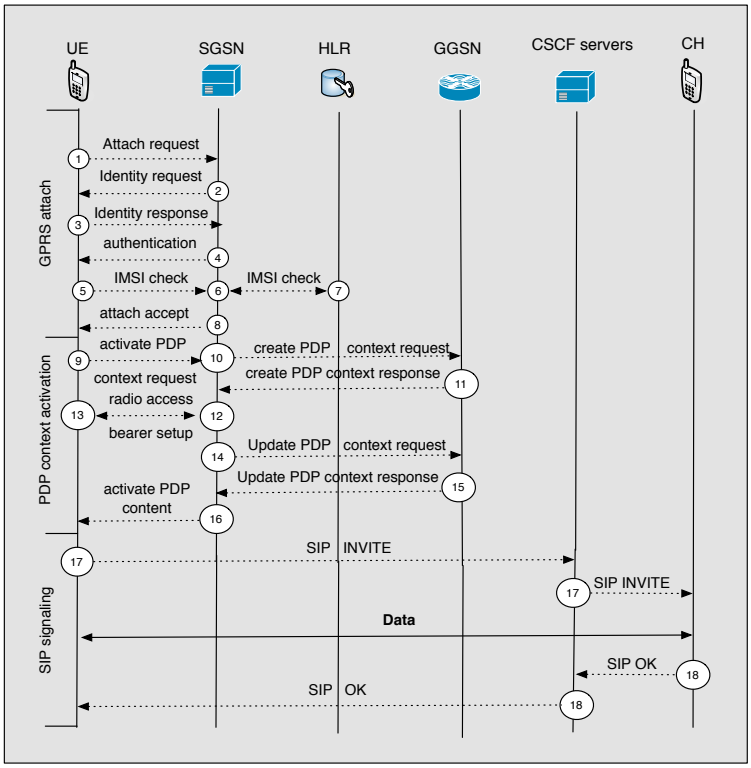


Figure 3.15: Signaling for WLAN-to-UMTS handover

UMTS to WLAN Vertical Handover

When a UE enters into a WLAN coverage area, it goes through the following steps to update its new location with the other communicating peer (*the CH*). The first step is the DHCP registration procedure which is the dynamic assignation of a new IP address for UE's new location. The messages exchanged during

this process are illustrated in Figure 3.16. When the UE detects the presence of a WLAN, it broadcast a DHCP discover message (*flow 1*) to find a DHCP server willing to provide the registration service. To achieve this, the authors assume an open network and they do not further discuss the related security issues such as mobile user authentication, authorization or accounting. In the next step, an appropriate DHCP server sends a DHCP offer message (*flow 2*) to the UE. Upon the reception of this message, the UE sends a DHCP request (*flow 3*) to confirm the offer just made.The DHCP server sends an ACK message with the adequate IP information to be assigned to the UE. Once the DHCP phase is accomplished and the UE has a valid IP address to operate under the WLAN, the UE initiates the SIP exchange process. Thus, the UE re-establishes the connection in a similar way than in the UMTS network, where the UE re-invites the CH by exchanging the SIP INVITE and SIP OK messages (*flows 5 and 6*).

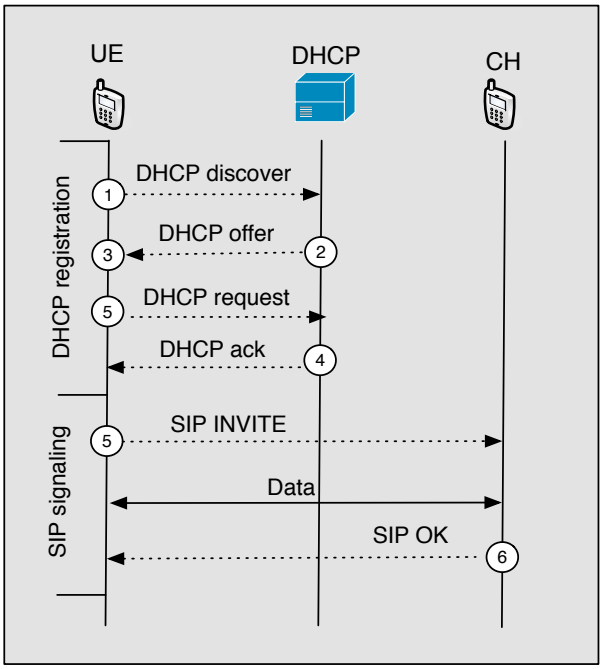


Figure 3.16: Signaling for UTMS-to-WLAN handover

Although, their analysis present acceptable results in terms of roaming delay, they do not take into account the potential delay component introduced by authentication and authorization process. Additionally, the assumption of an unsecured and open UWN willing to provide roaming services is not quite realistic.

3.3 Service Mobility in Heterogeneous Multi-Operator Wireless Networks

Service mobility is the capability of seamlessly transfer services and applications from one network to another. Yet, in spite of the numerous wireless inter-working

architectures, the seamless migration of services under a multi-operator environment yields not only technical but also other types of challenges.

Generally, roaming occurs when a subscriber of one wireless provider uses the facilities of another service provider. Moreover, roaming can be classified in *horizontal roaming* when switching network operators and *vertical roaming* when switching wireless access technologies. The combination of horizontal and vertical handover in a converged wireless world yields a new form of roaming: *the heterogeneous multi-operator roaming*. This type of roaming occurs when a mobile user switches not only network operator but also access network, as depicted by Figure 3.17. Heterogeneous multi-operator roaming poses several challenges for network operators and service providers, some of them can be solved by technical means but some others require redefining business models, new wireless regulations, standardization and most important the adaptation of applications to the new wireless world. Traditionally, terminal mobility is controlled from the core network of the cellular network however it must also be locally managed by the UWN once the user roam in. Current inter-working architectures rely on the assumption that both wireless access networks are controlled by the same operator, or in the best case scenario, both network operators have established roaming agreements. However, in the absence of such agreements the presence of NAT (*Network Address Translation*) and firewall rules cause problems to Internet telephony and slow down the deployment of transparent heterogeneous wireless access solutions. In this respect, we have identified two major potential preventers of service mobility in this scenario, these are: network access and seamless service mobility.

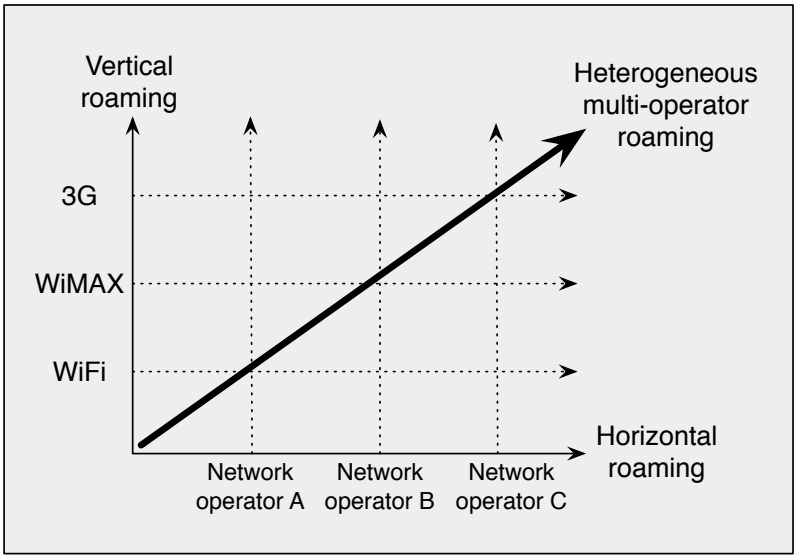


Figure 3.17: Roaming in Heterogeneous Multi-Operator Environments

In this section, we describe the challenges of seamless service mobility in heterogeneous multi-operator wireless networks. Moreover, we provide our vision of *SIP-based Seamless Service Mobility* and introduce a broker-based model for the

establishment of trust relations between heterogeneous network operators. Additionally, the underlying assumptions upon our research work is based are also presented in this section.

3.3.1 Technical-related Issues

The advances in mobility enabler technologies can make heterogeneous roaming a reality, however there are still some technical issues to overcome. After a detailed analysis, we consider necessary to rely on efficient network access mechanism, and efficient service provisioning to achieve seamless service mobility.

3.3.1.1 Network Access

Traditionally, before accessing a resource, a mobile user must go through authentication and authorization. In a heterogeneous multi-operator wireless environment, a mobile user must be granted with network access before roaming towards a UWN. To achieve this, UWN operators validate mobile user identity and verify (*in the service profile*) if they are authorized to roam in. This validation must be performed among both operators i.e. *WiFi - UMTS*. To increase the security of the transaction, some UWN relies on EAP-SIM to validate user identity (3GPP Rel. 6, 2004). Consequently, network access is only granted if both the cellular and the UWN have validated the mobile user.

From the proposed solutions, we observe that the performance of SIP-based handover is clearly influenced by the potentially large delay component introduced by the DHCP server, namely if the DHCP server attempts to use ICMP echo requests to determine if the new address has already been assigned (Banerjee et al., 2005). Additionally, the roaming process is also impacted by the AAA delays that would be incurred on inter-domain handoffs (Schulzrinne & Wedlund, 2000). Existing approaches such as UMA or the SIP-based architecture proposed in (Banerjee et al., 2005) do not address the issue of roaming into UWN that require authentication and authorization of mobile users and services.

For simplicity, cellular operators addresses network access by integrating the UWNs to their own architecture. In this way, they can apply their own AAA mechanisms and their own policies. Nevertheless, in a multi-operator context where each UWN applies their own security policies, network access becomes a real challenge to overcome.

For our research work, we use the concept proposed in the 3GPP concerning the differentiation between user authentication and service authentication (3GPP TS 23.234, 2006). In this perspective, our proposal in terms of heterogeneous network access primarily focuses on service authentication rather than user authentication. That is, the assurance that a specific service is really initiated or terminated by a specific mobile device (*service authentication*). This concept will be explain in details in the following chapter.

3.3.1.2 Seamless Service Mobility

Service mobility allows users to maintain access to their services even while moving or changing devices and network operators or service providers. One solution for service mobility is to have the mobile user carry this information on his PDA or cellular phone (*in a memory chip or the SIM card*). Nevertheless, even with local storage, updates made on any of the user's end systems still needs to propagate to the other devices, even if the device performing the update and the other devices are never in the same place. This requires network storage (Schulzrinne & Wedlund, 2000). The other issue refers to authorization, in this matter, even if the service profile is placed on the mobile device, the home network still has to validate such user profile before allowing any service to be used. Without the proper roaming agreements between network operators the remote validation of a user profile is not a trivial task. Furthermore, a concern of heterogeneous wireless operators refers to efficient service provisioning to mobile users once they roam into a UWN. The SIP architecture is predicated on having a *home server*, that can be physically located in the cellular network, associated with the user's SIP address. Under this architecture, the home sever stores service-related information and this element is in charge of propagating service information among all domains (Schulzrinne & Wedlund, 2000). Although, this architecture assures an efficient update of user's service profile it lacks of the security mechanisms to verify the mobile user's and rights to perform the services indicated in the service profile. In this connection, our research work provides efficient mechanisms for *service authorization*, a process that comprises the verification of the mobile user profile validity and the right of a subscriber to perform certain service. All this, without compromising the privacy of the user service profile. More details are available in the next chapter.

3.3.2 Business-related Issues

Heterogeneous wireless networks present technical differences but also differ in operational processes, organizational structure, and business models. While cellular networks are designed with the objective of providing voice and data to mobile users and charge for these services, private UWN are mainly used to provide data services to home/office users. From this, we can observe that cellular networks are entirely commercial solutions, thus having a robust business model hence a business oriented architecture. On the contrary, despite of the several UWN currently deployed around the world, UWN operators have not yet found an efficient business model that encourages rapid adoption of current connectivity technologies.

We are aware that some questions rise, when attempting to integrate heterogeneous networks, some of them are: can the business model of cellular networks hold in UWNs? can both networks be integrated in a business-related way? The answer to these questions may be found with an adequate business mode that can take advantage of the characteristics of UWNs to enhance cellular service. Although, business-related issues are out of the scope of this thesis, we present three issues that we consider as potential inhibitors for heterogeneous multi-operator

wireless inter-working.

3.3.2.1 Payment models

Nowadays, the most common payment model offered by WiFi hotspots is *pay-per-use* pricing as opposed to subscription-based (*prepaid*) pricing. In this context, the lack of a mechanism for a potential customer to compensate the UWN owner, this has no reason to accept the increased network traffic and security risk that would come from allowing external users to access her network. Payment models motivate UWN owners to open their networks to the public and enhance cellular service by enabling cellular-UWN roaming. Thus, the simplest way for an external user to compensate a UWN operator is to pay the UWN directly (*direct model or pay-per-use*). The other model is the aggregator model. In this model, the aggregator uses its brand name for all the host spots under its federation. Most of the time, the aggregators use a subscription-based pricing that applies to all the branded hot spots. A great challenge is to find a payment model that suits both UMTS and UWN operators, considering that the latter could or could not have deployed the UWN for business purposes.

- **Direct Model:**

  In spite of its simplicity, the direct or pay-per-use model poses certain issues for its deployment in great scale. The main issue is trust, in many cases the cellular operator, the mobile user and the UWN operator do not know each other and might not be able to trust each other to enable seamless heterogeneous roaming.

- **Aggregator Model:**

  Wireless aggregators, as stated previously relies on a subscription-based service, for a monthly or daily fee. The user gets a username and password (*credentials*) that allow her to tap into several hot spots under the aggregator domain. The main value proposition to customers is clearly convenience. They would otherwise have to sign up for multiple micro-carriers or UWNs depending on where they are and also be aware of who is servicing which areas. To provide this value, the aggregators rely on partnerships with micro-carriers or UWNs.

3.3.2.2 Billing

Among the main concerns in the heterogeneous wireless integration are who is going to manage *billing* and how users are going to see it reflected in their monthly bills. Most of the costumers do not want to get independent bills every time they roam in to a UWN. Independent UWN operators mean independent bills, thus even if the UWNs are federated by a single provider it is still not clear how the mobile users can get a single bill. In this respect, some authors propose a *central billing entity* (Balachandran et al., 2005), under this scheme a single

provider charges mobile users for their monthly hotspot usage with just another entry in their monthly bill. Nevertheless, there is till a challenge to overcome, how can operators can implement such unified pricing mechanism? The other billing approach is a *third-party billing contracts*, under this scheme UWN operators establish billing contracts with multiple ISPs that have already invested in a billing infrastructure (Haverinen et al., 2002). As a consequence, this approach yields challenges that are mainly related to *trust* among network operators.

3.3.2.3 Usability

Usability relates to how the end user perceives and interacts with the services provided. In this respect, some of the biggest challenges in terms of usability are: the legal consequences of UWN operators re-selling their Internet connection when offering roaming service, tax payments handling for service charged in the UWN, QoS guarantee and provisioning, customer support, coping with illegal actions performed in a UWN, and availability of UWNs. Most of these issues should be addressed by the mean of regulations.

Regulations motivate competition, they can improve services but most important they guarantee fair service provisioning. Thus, cellular-UWN integration needs urgent regulations to avoid unfair competition and provide acceptable quality of service. We consider that these regulation can be endorsed through inter-operator roaming agreements. Here cellular and UWN operators can settle the legal and business-related framework to enable efficient heterogeneous network interoperability, the challenge now is how to establish such agreements between within heterogeneous environments.

3.4 Towards Seamless Heterogeneous Multi-operator Roaming

Along this chapter, we have introduced the current solutions for heterogeneous wireless cooperation. We have also described the advantages and trade-offs of each solution, and their impact under the context of heterogeneous multi-operator roaming. In this section, we present our SIP-based vision towards a seamless service mobility. We provide the justification of why SIP would be suitable to handle roaming under this particular environments, and explain how the role that SIP currently plays in wireless architectures i.e. *UMTS* supports our decision.

3.4.1 SIP-based Seamless Service Mobility

Mobility is important for NGWNs as it has an enormous impact on how communication services and applications are evolving into the future. At the same time, mobility in heterogeneous multi-operator wireless networks requires new levels of mobility support as compared to traditional approaches.

Among the several proposals to achieve wireless integration, the all-IP core network architecture remains the most popular solution. The IP protocol plays an important role in achieving wireless networks convergence and most important,

compatibility with the Internet. IP-centric architectures and the adoption of SIP in the IMS of 3G wireless networks support this argument.

The SIP protocol design follows an IP state model which means most of the intelligence and state are located in the end devices. The network core maintains at most transactional state. This design provides benefits such as optimization in memory and CPU consumption and high reliability. Nevertheless, we consider network core intelligence necessary when operating with heterogeneous networks. In this context, by enabling network core intelligence i.e. *roaming decision making*, while establishing a session we can provide important benefits for converged wireless networks such as improved QoS levels, efficient user location management and a Firewall / NAT traversal.

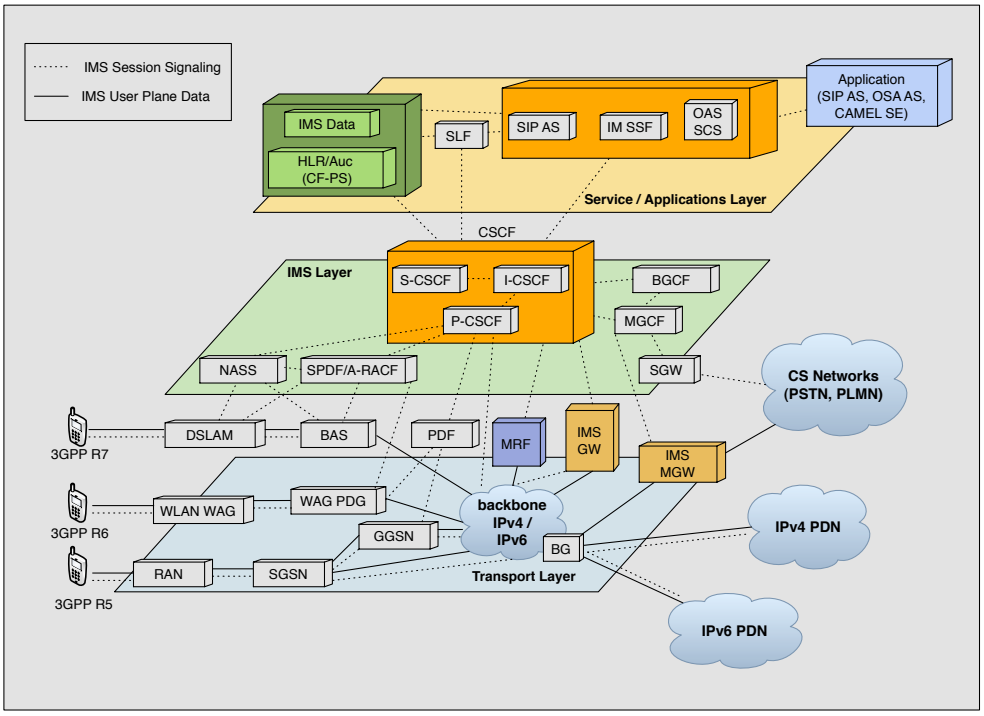


Figure 3.18: SIP integration in current wireless architectures

SIP-based cellular-UWN integration can be accomplished with no major changes in current wireless architectures. As illustrated in Figure 3.18, SIP allows a relatively easy integration of heterogeneous access networks. Different elements such as the PDG or the GAN enables the interaction of mobile devices attached to different wireless access technologies through an IP backbone. For this interoperability, in the UWN network, SIP relies on three logical entities: a proxy server, a redirect server and a local registrar. The proxy server is in charge of service negotiation with the CSCF in the UMTS network. Redirect server handles location management in the UWN network. The local registrar is a database within the redirect server. UWN local registrar handles local mobility within a wireless distribution system i.e. *an enterprise wireless local area network*. In the UMTS network no major changes are needed as the IMS core elements P-CSCF,

I-CSCF, S-CSCF and the HSS support SIP. Some of the advantages of a SIP-based implementation are: the UMTS architecture does not require major changes, SIP can be used to address mobility issues, and SIP extensions are easily supported through software updates.

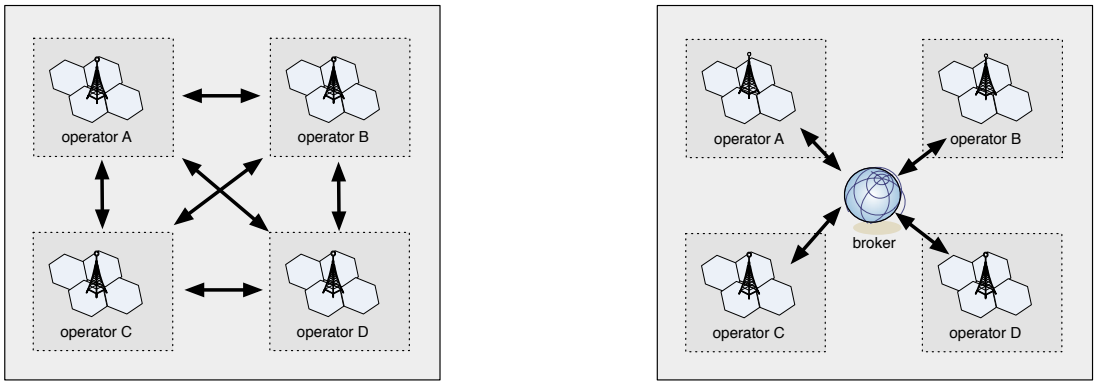
3.4.1.1 SIP-based Location Management

Mobility management in homogeneous network has been addressed extensively, currently there also are certain approaches for mobility management in heterogeneous wireless networks. In this context, we have mobility management from four different layers of the OSI (*Open System Interconnection*) model: link layer mobility which is specific for each wireless access network. This relates to the capability of roaming between access networks of different kinds i.e. *vertical handover*. The network layer also addresses mobility issue from the IP layer, being Mobile IP (Perkins, August 2002) and IPv6-based solutions among the most popular. The transport layer also contributes to mobility management, currently solutions have been proposed by using SCTP (*Stream Control Transmission Protocol*) t (Stewart et al., 2000). Finally, application layer mobility, specifically with SIP has being widely accepted and nowadays it has being adopted by the 3GPP as fundamental part in the IMS (*IP Multimedia Subsystem*) (Garcia-Martin, May 2005). Cross-layer solutions are mainly proposed for handover management, they aim at providing layer 3 handover with the support from layer 2 by obtaining link layer statistics such as signal strength reports and movement detection information. The rationale behind this is that by obtaining information from lower layers the system can better prepared for the network layer handover hence eliminating packet loss and reducing handover latency (Akyildiz et al., 2004).

Recent research efforts attempt to design general location management mechanisms for the integration and inter-working of heterogeneous networks.These research activities can be grouped into two scenarios: location management for adjacent heterogeneous networks with partially overlapping coverage at the boundaries and location management in fully overlapped heterogeneous networks . We consider that under real circumstances and due to the penetration of cellular technologies, the most common scenario is the latter. When the service areas of heterogeneous wireless networks are fully overlapped, the mobile user is reachable through multiple networks. In this context, we consider that the utilization of network core entities that are in charge of performing address translation, packet filtering, access control, will play an important role. Thus, firewalls and NAT traversal do not pose a problem to mobility. End-users and core network entities such as WiFi access points and the elements in the IMS should exchange user location information to provide acceptable QoS levels and efficient packet delivery

3.4.2 A Broker-based Roaming Model

Future mobile communications systems will be characterized by multi-operator / service provider environments. At the same time, this will be followed by dynamic interactions and relationships (Edwards et al., 2004). In this context, we require a scalable security and trust establishment infrastructure. Trust can be



(a) Independent Trust Relations **(b) Broker-based Trust**

Figure 3.19: Trust Relationships

defined in different ways, a definition provided in (Grandison & Sloman, 2000) defines trust as:

> *"the firm belief in the competence of an entity to act dependably, securely and reliably within a specified context"*.

On the other hand, in their work (Wilhelm & Butty, 1998), define four general foundations of trust: blind trust, trust based on reputation, trust based on control and penalties, and trust based on policy enforcement. In this respect, the WWI (*World Wireless Initiative*) also mention an additional foundation for trust:

> *"Trust hat a device or process will behave in a particular way based on its design"* (Edwards et al., 2004).

This definition is also the foundation of our research work as all the elements integrating our roaming architecture have trust policies implicit in their architecture as we proposed the use of devices with trust policies implicit in their conception. This will be explained in details in the following chapter.

The principle of transitivity states that if A trust B and B trust C then A trust C. In this regard, The best example of transitivity in mobile communications is *roaming*, when a mobile user travels abroad it gets connectivity via a local operator with whom it might had no prior relationship. Nevertheless, the mobile user gets access to its services thanks to a roaming agreement establish between the home and the visiting network. In this perspective, we see roaming agreements as instruments towards trust establishment. Trust in a device is already common in mobile communications where the SIM card provide a predefine service environment. Consequently, we consider roaming agreements as efficient mechanism to endorse trust relations between network operators.

The establishment of bilateral trust relationships between entities can be achieved independently or through a broker-based model, as illustrated in Figure 3.19. The main difference of both approaches is the number of relationship for users, in the former we have $n(n-1)/2$ relationships for n users and in the later we have n

relationships for n users. On the other hand, trust modeling can be addressed from a *quantitative* view, through a trust computational model used by each entity to manage its own trust relations or through a *qualitative* view that provides a global view of trust relations between each pair of entities. In case of using the latter approach a high level design of a system is required.

For our research we decided to focus on the qualitative approach as we developed an architecture that seamlessly integrate cellular and UWNs by providing an efficient roaming platform. Thus, by choosing SIP as core signaling protocol and adopting a broker based approach to efficiently reduce the trust relationships among operators, we designed a SIP-based roaming architecture for heterogeneous multi-operator wireless networks. With this model, our SIP-centric roaming broker aims at establishing *mutual trust* between operators not only with the objective of verify the integrity of the elements of the architecture. Our architecture is presented in details in chapters 4 and 5.

3.5 Underlying Assumptions

We are aware that the presence of multiple independent network operators or service providers raise new and significant issues. In this case, in order to provide an efficient solution for heterogeneous wireless networks integration, we consider important to provide the underlying assumptions. As the differences between network operators or service providers are not merely technical we classify our assumptions in technical and network-related. Although, wireless interoperability also raises business related issues, they are out of the scope of our research.

3.5.1 Technical assumptions

When working with multi-operator wireless networks we are likely to face technical differences, thereby for the further development of our proposal we state the following technical assumptions:

- The mobile station should have dual wireless interface i.e. *WiFi-UMTS*.
  Justification: for the support of heterogeneous wireless roaming, the mobile device will be physically able to switch from different access networks.

- We rely on an efficient vertical handover algorithm in the mobile station.
  Justification: Our research focuses mainly in the heterogeneous roaming issue hence, vertical handover is out of the scope of this thesis. Nevertheless, to avoid dropped calls, packet loss and large delays while roaming from one wireless access network to another, an efficient vertical handover algorithm is required.

- The cellular network has an all-IP core network.
  Justification: We assume that the all the calls performed are VoIP calls, hence the need of an all-IP core network.

- The cellular network and the UWN are interconnected through an IP packet switching network such as the Internet.
  Justification: The rationale of this assumption is that most of the WiFi hotspots deployed are connected through the Internet. Thus, this configuration creates a loosely coupled architecture that new Visiting Networks can easily join.

- The access points in the UWNs support monitoring, logging and statistics for SLA monitoring and enforcement.
  Justification: Some access points in the market provide acceptable computing and storage capacity, hence can provide network statistics on demand.

- The UWNs allows network attachment to all the external visitors without requiring authentication (*mobile users obtained valid IP addresses*). Nevertheless, all the network services are blocked until the mobile user performs service authentication and authorization.
  Justification: With this assumption we separate the user AA (*Authentication and Authorization*) from service AA.

- All the key-elements of our architecture support SIP as signaling protocol.
  Justification: As our roaming signaling protocol is SIP-based, we assumed that the access points in the VN, the cellular core network and the broker rely on SIP as roaming signaling protocol.

3.5.2 Network-related Assumptions

In order to cope with multiple heterogeneous network operators we state the following assumptions:

- The cellular network is always considered the HN (*Home Network*).
  Justification: The cellular network provide certain levels of QoS and SLAs to the users that UWN cannot support all the time. Hence, in order to roam into UWNs the cellular network must approve and verify that such networks can offer acceptable levels of QoS to the roaming users.

- The UWNs are always considered the VNs (*Visiting Networks*).
  Justification: In our architecture the UWNs are extensions of the cellular network, hence are considered the VNs.

- Mutual trust between the broker and the HN is endorsed by contractual agreements.
  Justification: Through these agreements the broker assumes responsibility of the VNs under her domain. Consequently, the broker is also responsible of verifying the functionality and the performance of the VNs.

- Mutual trust between the broker and the VNs is endorsed by SLAs that clearly state on the conditions of service supplying.

3.6 Conclusion

Nowadays, existing architectures support service migration across heterogeneous wireless access network. For example, a mobile user can perform a voice call from GSM and continue such service from a UWN. Along this chapter, we have presented interesting wireless inter-working architectures and efficient methods to handle terminal mobility under heterogeneous environments. Nevertheless, the main assumption of current solutions is that a single-operator owns or managed different access network or in certain cases that cellular subscribers roam into open WLANs. While these assumptions holds under certain circumstances, there are still a great number of independent UWNs that could be used as extensions of cellular networks.

Roaming across heterogeneous wireless access networks raises several technical, business, and network-related challenges. This process becomes more complex when roaming combines switching wireless access network and network operator. The dynamic establishment of roaming agreements between cellular networks and UWN is still in early development. Notwithstanding, existing technologies such as vertical handover techniques and SIP-based terminal mobility can be used to build an efficient roaming platform that enables seamless service mobility in heterogeneous multi-operator wireless environments.

Our research considers such technologies as a path towards a world where heterogeneous multi-operator wireless networks are capable of interacting and providing uninterrupted services to mobile users. Thus, combining existing roaming technologies, wireless inter-working architecture, and interesting concepts such as roaming agreements, broker-based trust foundation, and the differentiation of user and *service authentication and authorization*), we consider that the support of service mobility between networks with different business models and technical characteristics is feasible.

Chapter 4

A SIP-based Roaming Architecture

The SIP-based roaming architecture aims at the creation of a suitable environment for the support of seamless service mobility in heterogeneous multi-operator wireless networks, this without major changes in current wireless network architectures. To achieve this, we rely on an element called the Roaming Broker (*RB*). The goal of the RB is to establish *mutual trust* between the cellular and the UWNs. As trust can be built by different means (Wilhelm & Butty, 1998), for our research we consider that mutual trust between the cellular network and the RB is endorsed by contractual agreements. On the other hand, trust between the RB and the UWNs is endorsed by SLA (*Service Level Agreements*) and efficient access control mechanisms. Furthermore, we assume that roaming between heterogeneous wireless networks, primarily in the case cellular-UWN, follows economic rather than technical reasons. Consequently, heterogeneous roaming must be allowed only if the UWN respects the accorded SLAs.

This chapter is organized as follows: section 4.1 describes the architectural elements of our SIP-based roaming architecture. Section 4.2 explains the RB's mechanisms to bind mutual trust between heterogeneous multi-operator wireless networks. Moreover, section 4.3 presents in detail the proposed SIP-based Roaming Signaling in our architecture. Section 4.4 describes the simulation environment used to evaluate the performance or the proposed signaling protocol. Section 4.5 presents the results obtained through computer simulation. In addition, section 4.6 describes a feasibility study and presents the testbed outputs. Finally, section 4.7 concludes this chapter.

4.1 Architectural Elements

The key-elements integrating our architecture are: the Roaming Broker, the Home Network (*HN*) and the User Equipment (*UE*), and the Visiting Networks (*VN*).

4.1.1 The Roaming Broker

As stated in section 3.5 a broker-based model simplifies service deployment since the UWNs members of the roaming broker's domain do not need to agree among themselves, only with the broker (Raman et al., 2002). Additionally, the UWNs

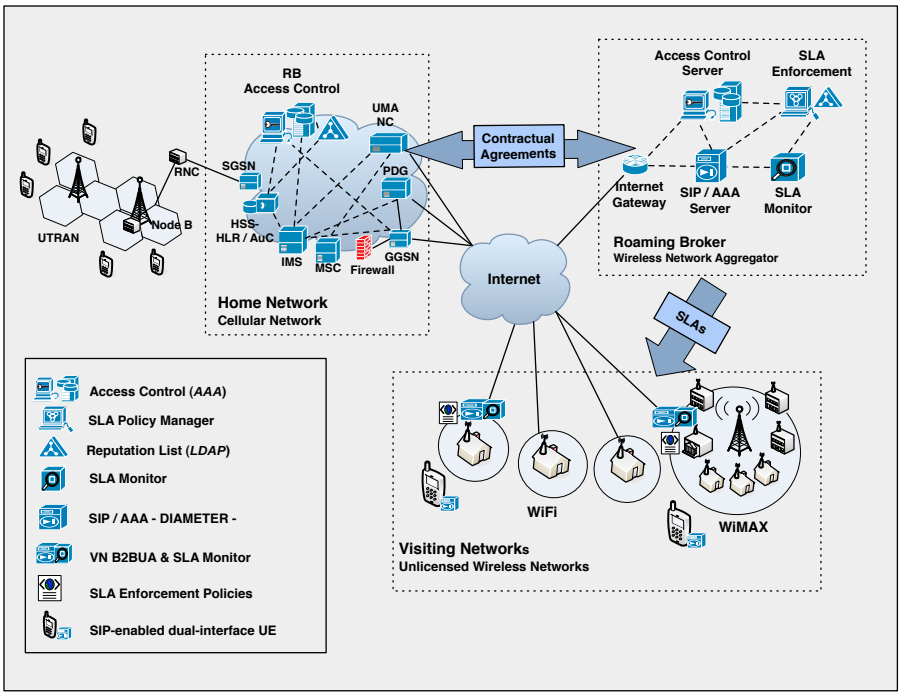


Figure 4.1: SIP-based Roaming Architecture

could or could not be aware that are participating under a large-scale roaming architecture. In this context RB binds mutual trust between the cellular (*HN*) and the UWN (*VNs*). In particular, the RB establish mutual trust with the HN by a *contractual agreement*, in this contract the RB assumes responsibility about the VNs under her domain. In this perspective, we can imagine the RB entering into contracts with network operators and service providers, and using these contracts to offer roaming services between cellular and independent UWNs. On the other hand, mutual trust between the RB and the VN cannot be established through contractual agreements as it is no efficient to enter into contracts with every UWN. Instead, the RB establishes mutual trust with the VNs through SLAs and a verification mechanism that ensures the functionality of each VN federated by the broker. Here, we present the key elements that integrate our proposed roaming broker, these are: the access control server, a SLA policy manager, a reputation list, a SIP/AAA Server and a SLA monitor, see Figure 4.1.

4.1.1.1 Access Control Server

The RB manages access control for VN's under its domain and collaborates with the HN in UE's access control. This is accomplished through the Access Control Server in the RB, this element contains the access control database (*RADIUS*) (Rigney et al., 2000). This server allows the RB to authenticate and authorize VNs to offer roaming services. For authentication, the RB assign special credentials i.e. *VN id and password* to each VN in the broker domain, this information is stored in

the access control database in the server. On the contrary, authorization is managed by the access control server in collaboration with the reputation list.

4.1.1.2 SIP/AAA Server

The SIP/AAA Server in the RB analyses the SIP-INVITE and SIP-REGISTER messages and interacts with the Access Control Server in the authentication and authorization process. This element is the SIP server in the RB, at the same time it interacts with the SIP Outbound Proxy server in the VN and the IMS in the HN. The SIP/AAA server also interacts with the SLA Enforcement & SLA Monitoring elements. Here, the SIP/AAA extracts SLA information from SIP messages and passes this information to the related entity. From this, we can imply that the SIP/AAA server is the core entity in the RB.

4.1.1.3 SLA Policy Manager

In many scenarios, one of the parties i.e. *the HN* defines most of the content of the SLA and the service provider such as the VN simply agrees to such information. However, in a more flexible scenario, while the HN may define many of the aspects of a service, including the definition of specific SLA metrics and measurement directives, it may offer a choice to the VN operator on the details of the guarantees such as SLA violation threshold, punishment in case of violation, etc. For our architecture, we assume that the SLA Policy Manager in the RB, based on the requirements provided by the HN, defines the SLA content, measurement metrics and actions to be invoked in case of SLA violation. In this context, we define a SLA as a contract between the RB and the VN of one or more technical features, that rules the supply conditions and that defines constraints of quality levels of such features i.e. *overall capacity/throughput, reporting mechanism, authentication methods, etc*. In this perspective, the SLAs can be composed of a business-legal part and a technical part however our research work only focuses on the technical part. Upon the subscription of the VN in the RB's domain and based on its technical capabilities i.e. *Internet bandwidth, wireless bandwidth, etc* the RB assisted by the SLA Policy Manager creates a SLA for every VN. Thus, the SLAs policies are created by the RB from measurable parameters such as network capacity/throuhput, network delay, number of simultaneous VoIP calls, etc. These policies are created, managed and applied in the VN through the SLA policy manager. In order a SLA to become effective, the agreed constraints have to be placed in the real network, thus the SLAs in our architecture are placed in the form of SLA enforcement policies on the wireless routers in the VNs. These policies define a set of rules placed on the access point that make it behave as specified in the SLA.

4.1.1.4 The Reputation List

The RB relies on a LDAP (*Lightweight Directory Access Protocol*) (Zeilenga, 2006) based reputation list to store the SLA-related information provided by the VNs. The LDAP directories in the architecture follow the 1993 edition of the X.500 (ITU,

2005) model. In this respect, an LDAP-based directory can be seen as a tree of directory entities, at the same time these entities consists of a set of attributes, see Table 4.1. In the Reputation List each attribute has a name (*attribute description*), a unique identifier such as the VN's Distinguished Name (*DN*), object class, e-mail, manager (*the RB identification*), a user certificate, etc. In addition, the Reputation List contains three type of labels describing the VN's reputation: WHITE, GREY and BLACK. Such labels are assigned based on the VN's technical capabilities and their respect to SLAs. Thus, white indicates a reliable VN, grey an unstable VN and black a VN that has continuously broken the SLA.

Table 4.1: Reputation List's Directory Structure

dn: cn=VNid, dc=roamingbroker,dc=com
cn: VNid
objectClass: organizationalPerson
objectClass: person
displayName: Bob's VN
sn: VNid
mail: bob@example.com
manager: cn=brokerid,dc=roamingbroker,dc=com
userCertificate: optional
userPKCS12: optional
reputation: **WHITE**

4.1.1.5 SLA Monitor

The SLA Monitor in the RB is in charge of obtaining network statistics from the VNs and determining the SLA-based reputation of each VN under the broker domain. The SLA-related information is provided by the VNs and conveyed within SIP messages once an external user attempts to register in the network or initiates a session from a VN. This process is explained in detail in section 4.3. Upon the reception of such statistics, the SLA monitor creates or updates the Reputation List.

4.1.2 The Home Network and the User Equipment

The cellular network, as stated in the underlying assumptions in section **??**, is always considered the HN. To enable heterogeneous roaming the architecture relies on the IMS and on a SIP-enabled UE, as illustrated in Figure 4.2.

4.1.2.1 IP Multimedia Subsystem

The IMS, is an architectural framework originally developed for delivering IP multimedia services to end-users. In our architecture, we use it for roaming sig-

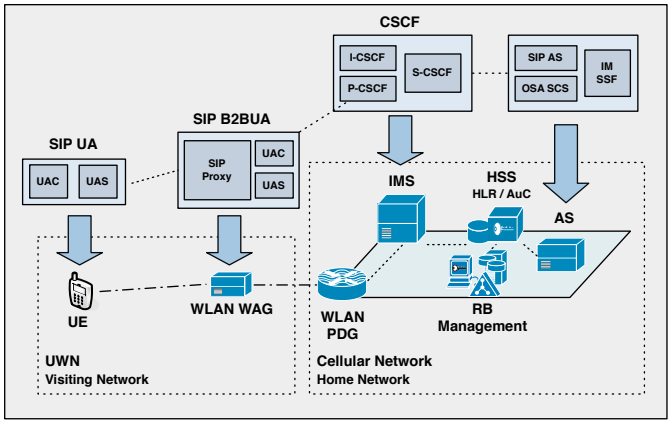


Figure 4.2: The Home Network and the UE

naling purposes. Among the tasks performed through the IMS are: session initiation, SLA-information exchange, location update, and access control. All of this without compromising the primary goal of the IMS which is providing Internet services to mobile users.

4.1.2.2 SIP User Agent

A SIP User Agent (*SIP UA*) is a logical entity that comprises a User Agent Client (*UAC*) and a User Agent Server (*UAS*). The UAC originates requests whereas the UAS generates responses. During a session, a UA will usually operate as both UAC and UAS. When integrating into a UE, our SIP UA enables the mobile device to originate and terminate calls with or without a SIP proxy server, see Figure 4.2. In our architecture, the UA advertises its capabilities and features in any requests it sends. Thus, the VN could provide the services or applications that better suit the user.

4.1.3 The Visiting Networks

As stated in section 3.5, the UWN are always considered the VNs. The objective of such networks is to provide connectivity to external users while respecting their service profile, all this within a multi-operator context. To achieve this, our architecture defines three elements within the Wireless Access Gateway (*WAG*) in the UWN: a SIP B2BUA (*SIP Back-to-Back User Agent*), the SLA Policies and a SLA Monitor.

4.1.3.1 The need for a SIP B2BUA

There are some cases in which SIP-based service mobility will not work. For example, if a local proxy server must be traversed in order to enable firewall or NAT traversal, or for some security reason, then a UA may have to use a different first

hop proxy when roaming. In this case, service mobility is still possible provided that:

1. The roaming UA is able to discover the necessary local proxy (*VN*).
2. Both incoming and outgoing requests are routed through the home proxy in addition to any local proxies (*VN*).

These SIP mobility capabilities are well suited to use over a wireless network such as 802.11 in a home, office, or public space. As roaming agreements allow such wireless hotspots to be linked up in metropolitan areas, this will provide a wireless service. However, commercial wireless providers plan a specific purpose wireless telephony network using SIP. Wireless SIP clients may also make use of voice codecs tolerant to packet loss, which may be experienced in a heavily loaded 802.11 network.

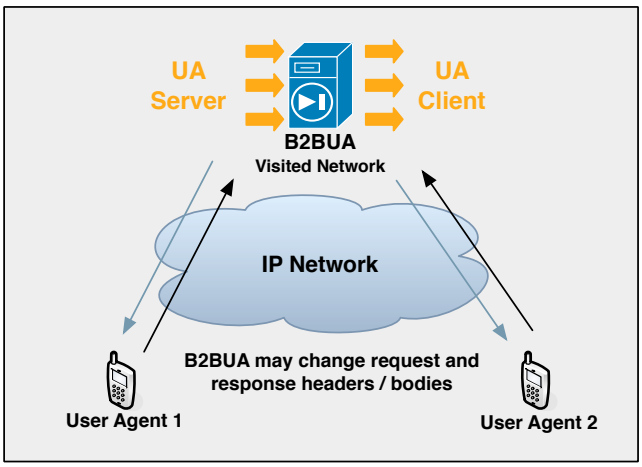


Figure 4.3: SIP Back-to-Back User Agent

A SIP proxy operates in a similar way to a proxy HTTP (*Hyper Text Transfer Protocol*) and other Internet protocols. A SIP proxy does not set up or terminate sessions, but it is placed in the middle of a SIP message exchange, receiving messages and forwarding them. In this perspective, a SIP proxy is an intermediary entity that mainly plays the role of routing. It decides about routing and forking and also applies policy and authorize certain calls to certain users. Nevertheless, based on the RFC 3261 recommendation, a SIP proxy may not alter SIP messages headers or body (*except routing related headers such as VIA*).

4.1.3.2 The SIP B2BUA

A SIP B2BUA (*Back-to-Back User Agent)* acts as a user agent to both ends of a SIP call. The B2BUA is in charge of handling all SIP signaling between both ends of the call, from call establishment to termination. To the SIP clients, the B2BUA behaves as a UA server on one side and a UA client on the other (*back-to-back*)

side, as illustrated in Figure 4.3. The basic implementation of B2BUA is defined in the RFC 3621 however, the B2BUA may also provide the following functionalities:

- call management (billing, automatic call disconnection, call transfer, etc).
- inter-working with alternative networks.
- management and monitoring of the entire call state.
- firewall or NAT traversal.
- codec translation between two legs.

In the architecture, the VN relies on a B2BUA (*Back-to-Back User Agent)* to include SLA-related information within the SIP messages before forwarding them. It is worth-mentioning that our design of the B2BUA respects the rules defined in the RFC 3621. Thus, we preserve the end-to-end transparency while still allowing our VN to perform valuable services and functions for users agents, such as providing SLA-related information.

4.1.3.3 SLA Enforcement Policies

The SLA Enforcement Policies are created by the RB's SLA Policy Manager from measurable parameters such as network capacity/throuhput, network delay, number of simultaneous VoIP calls, etc. These policies are then placed on the access points in the UWNs (*to define a set of rules that will make it behave as specified in the SLA*). Once the SLA policies are placed on the VN, certain services such as access control, QoS and bandwidth managers are governed by these policies. To address issues such as wireless bandwidth management, we recommend the utilization of call admission control (*CAC*) policies for UWNs. Some example of such policies can be found in (Garg & Kappes, 2003a) (Gao et al., 2005) (Wu & Bertsekas, n.d.).

4.1.3.4 SLA Monitor

The VN's SLA Monitor in the VN is in charge of gathering network information and calculating the SLA-related statistics. This element relies on different techniques to obtain or calculate network statistics i.e. *SNMP (Simple Network Management Protocol), Ping, and HTTP (Hyper Text Transfer Protocol) requests*. Then, it transmits these network statistics to the RB. Once the RB's SLA monitor receives the statistics, it creates or updates the *Reputation List*.

4.2 Roaming Broker's Mechanisms

The mutual trust bound between the RB and the VNs is valid while the VNs provide efficient access control mechanisms and respect the QoS levels specified in the SLA. In this context, the architectural elements presented in previous section play an important role in the two core-mechanisms in the RB: Access Control and SLA Enforcement.

4.2.1 Access Control

The RB manages access control for VN's under its domain and collaborates with the HN in UE's access control. This is accomplished through two elements located in the RB: the access control database (*RADIUS*) (Rigney et al., 2000) and the SIP/AAA DIAMETER gateway (Calhoun et al., 2003). The former allows the RB to authenticate and authorize VNs to offer roaming services. For authentication, the RB assign the credentials (*VN id and password*) to each VN in the network, this information is stored in the RADIUS server. The authorization is also controlled by the RADIUS server in collaboration with the LDAP reputation list. In this perspective, in addition to VN's identification the RB must verify whether or not the VN has acceptable reputation to offer roaming services.

The other type of access control relates to the UE roaming into the VN. In our architecture, we consider the HN as the roaming decision maker. Neither the VN nor the RB can authenticate or authorize a UE to roam into the VN, this privilege is reserved to the HN. The UE access control is triggered once the UE enters into the VN coverage area. At this point, the access point in the VN acts as an SIP gateway that sends the UE SIP-AAA request to the RB. The RB due to the contractual agreements with the HN is able to forward the SIP-REGISTER message to the IMS in the HN. Once the user is authenticated and authorized by the HSS-HLR/AuC (*Home Subscriber Server-Home Location Register/Authentication Centre*) in the HN, the RB authorizes the VN to accept the visitor. The architecture includes SIP/RADIUS and SIP/DIAMETER elements in the RB because we consider important, for backward compatibility, to support both AAA mechanisms.

4.2.2 SLA Enforcement

SLA Enforcement relies on the SLA Policy Manager, the SLA Enforcement Policy, the Reputation List, the RB's SLA monitor and the VN's SLA Monitor, as illustrated by Figure 4.1. The SLA Enforcement process is the following: The SLA Policy manager, based on the technical characteristics of the VN, creates and places the SLA Enforcement Policy in the VN. Once applied, the VN will follow the rules specified in the policies.

In the architecture, this SLA Enforcement is embedded in the network registration and session initiation process. Thus, before allowing a visitor to register into the network, the VN assures it has an acceptable reputation and that it can provide the QoS levels specified by the SLA Enforcement Policies. At the same time, once the UE attempts to initiate a VoIP call from the VN, the SLE Enforcement in the session initiation process assures the VN can allocate the session without affecting ongoing multimedia sessions. Although, this process remains similar to network registration, there are slight changes in the signaling messages. The SIP-based Roaming Signaling used in the architecture to provide efficient access control and SLA enforcement is explained in the next section.

4.3 SIP-based Roaming Signaling

The key-elements of our architecture communicate through an enhanced version of SIP. Our contribution to this protocol is the addition of extensions to enable reputation-based access control as well as SLA monitoring and enforcement support. This section describes the aforementioned extensions to SIP.

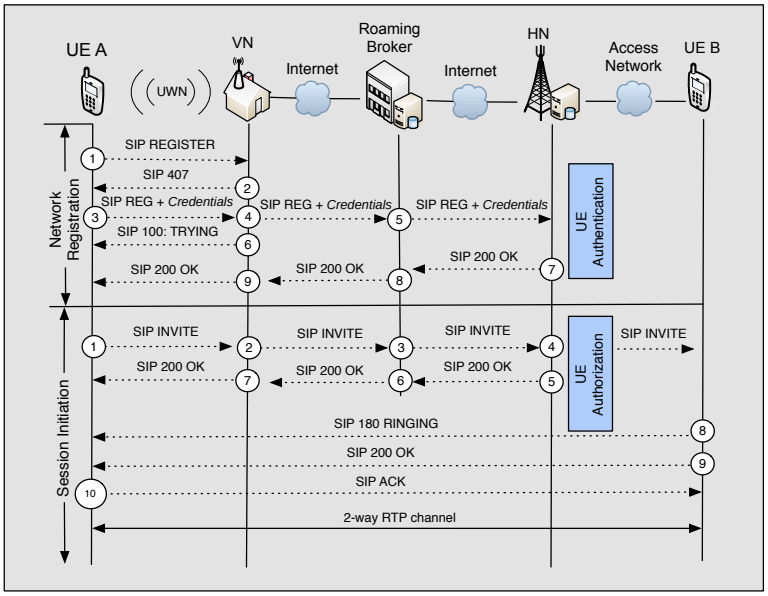


Figure 4.4: SIP-based Roaming Signaling

In most of the cases, before accessing a resource we are required to perform two mechanisms: authentication and authorization. In our architecture, authentication signaling is related to the network registration and authorization signaling to session initiation process. The RB acts as an authorized proxy to establish SIP dialogs with the HN on behalf of the VNs. The signaling message exchange starts when the UE initiates the network registration process in the VN or when the UE attempts to initiate a multimedia session i.e. *VoIP call*. In this section we describe our proposed extensions to SIP to support broker-based access control and SLA information exchange. It is worth-mentioning that our extensions to SIP follow the guidelines for authors of extensions to the SIP as described in (Rosenberg & Schulzrinne, 2006). SIP has been proposed as a solution for numerous problems, including mobility, configuration and management, QoS control, call control, etc. However we do not consider SIP as a solution itself but merely as the roaming signaling protocol of our architecture.

4.3.1 Reputation-based Network Registration Signaling

In SIP-AAA (Loughney & Camarillo, 2004), the authors propose the addition of the user's credentials within the **REGISTER** message upon the reception of a **407: Proxy Authentication Request** message, as illustrated in Figure 4.4. Further-

more, we add the VN broker-based access control and SLA information exchange, to include SLA Enforcement in the network registration process.

```
SIP/2.0 407 Proxy Authentication Required
Via:SIP/2.0/TLS client.mobile.com:5061;
branch=z9hG4bKnashds7
Cal-ID:311316842@mobile.com
From: <sips:bob@mobile.com;user=phone>
To: <sips:bob@mobile.com; user=phone>
CSeq: 1 REGISTER
Proxy-Authenticate: Digest realm="roamingbroker.com",
qop="auth",
nonce="ea9c8e88df84f1cec4341ae6cbe5a359",
opaque="", stale=FALSE, algorithm=MD5
Content-Length: 0
```

Figure 4.5: SIP Proxy Authentication Request Message

The network registration process as illustrated in Figure 4.4 is the following: the UE requests network access by transmitting a SIP REGISTER message (*flow 1*). This message includes a Content-type in the SDP (*Session Description Protocol*) (Handley & Jacobson, April 1998) section of the SIP message, specifying the minimum QoS requirements for the application. Once the VN receives the registration message it calculates the available resources. If the VN can fulfill the QoS requirements then the VN replies with a **Proxy Authentication Message: Error 407** asking the UE to provide its credentials (*username and password*) (*flow 2*), the structure of this message is depicted in Figure 4.5. Otherwise the VN transmits an error message indicating that the QoS cannot be guaranteed. Upon the reception of a 407 message, the UE re-transmits the registration message with its credentials (*flow 3*).

Our contribution to SIP initiates here and comprises the addition of the VN's credentials and current network statistics (*SLA-related information*) upon the reception of the SIP REGISTER message from the UE, see Figure 4.6. The VN's B2BUA includes such information in the SIP REGISTER message and forwards the message to the RB, as illustrated in Figure 4.4 (*flow 4*). Upon the reception of the SIP-REGISTER message, the RB extracts the VN identifier and the SLA-related information to perform a look up in the Reputation List. If the VN has an acceptable SLA reputation and the current network statistics fulfill the SLA requirements, then the RB is ready to forward the SIP-REGISTER message to the HN for network registration, otherwise it is discarded. Before the transmission of the SIP-REGISTER message to the HN, the RB includes its identifier and password (*flow 5*). Finally, once the RB credentials are confirmed and the UE is authenticated and authorized to roam in to the VN, the HN informs through a SIP-200 OK message that network registration has been successfully accomplished (*flows 6 to 9*). If there is an error concerning the UE credentials the VN responds with a **SIP 401 UNAUTHORIZED** message; the reception of such message indicates that the UE will not be able to roam into the VN.

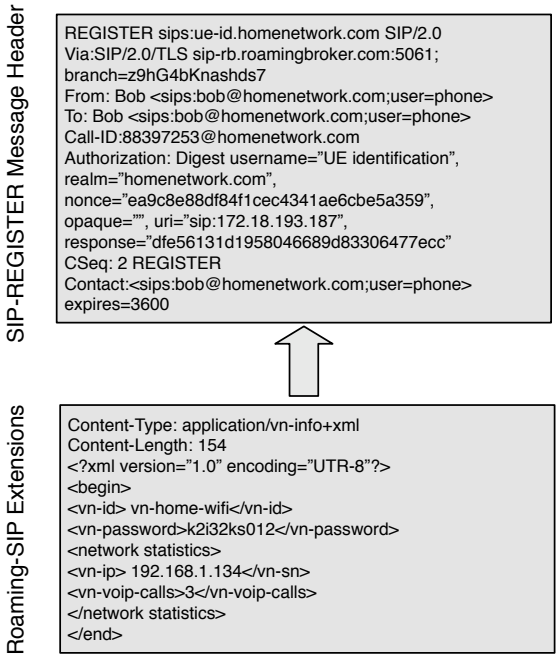


Figure 4.6: Roaming SIP-REGISTER Message

4.3.2 SLA-compliant Session Initiation Signaling

Session initiation signaling starts with the **SIP-INVITE** message once a UE attempts to establish a multimedia session i.e. *VoIP call* from a VN (*flow 1*). Likewise the network registration process, the application specifies within the SIP INVITE message, the minimum QoS requirements in order to establish the session. We have enhanced this process as follows, upon the reception of the SIP INVITE message, the VN based on its current capacity and the SLA policies determines whether or not can meet the SLA specified by the RB, as illustrated in Figure 4.7. Once verified the SLA requirements, the VN attaches into the SIP INVITE message current network statistics and forwards it to the RB (*flow 2*). Then, the RB receives a SIP-INVITE message from a VN, the RB processes this message and queries the reputation list. If the VN has acceptable SLA reputation it redirects the SIP INVITE message to the HN which is the entity that will deliver, upon authorization, the message to the other peer (*flow 3*). This can be performed through the cellular network (*flow 4*) or through a broadband IP network. Once the other peer responds to the invitation, both peers are able to start the multimedia session (*flows 5 to 10*). The data exchange between the peers do not pass through the RB.

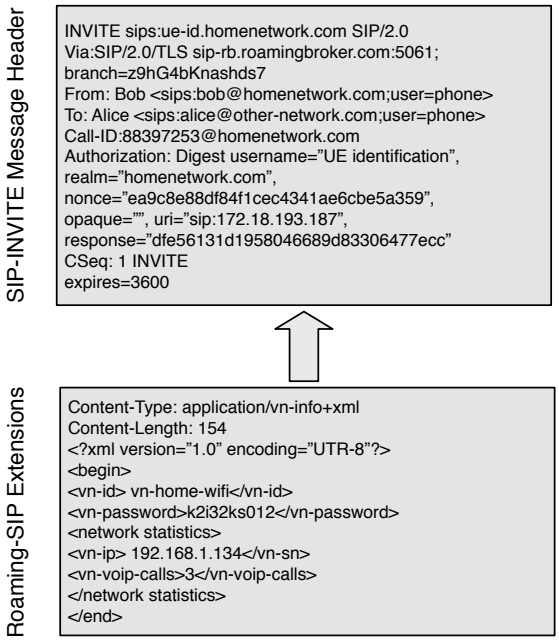


Figure 4.7: Roaming SIP-INVITE Message

4.4 Roaming Signaling Simulation Environment

We consider that in a future it will be common to find several mobile users attempting to roam into UWNs. Hence, we decided to create a simulation model to evaluate the impact of high levels of traffic on our SIP-based roaming architecture.

4.4.1 Simulation Objective

The main objective of the simulation is to evaluate the impact of high load scenarios on network registration process and session initiation in our architecture. Our main interests is to prove that our extensions to SIP do not impact the overall performance of the protocol. In addition, we are also interested on showing the feasibility of implemented the aforementioned architecture.

The simulation was performed on the Network Simulator (*ns-2*) (NS-2, 1995), a discrete event simulator targeted at networking research. For our analysis we enhanced the Rui Prior's SIP module for *ns-2* (Prior, 2006) to support our proposed SIP extensions. Our contributions to this module include the SLA-aware SIP message length, the processing delay due to authentication and authorization, and the processing delay introduced by the VN and the RB. Moreover, a VoIP model to simulate traffic in the VNs was also added. These additions are described in detail in the appendix of this thesis.

4.4.2 Simulation Model

Figure 4.8 illustrates the network model used in the simulations of the roaming architecture. This model takes into account the elements in the VN, the RB and in the HN, that participate in the roaming process.

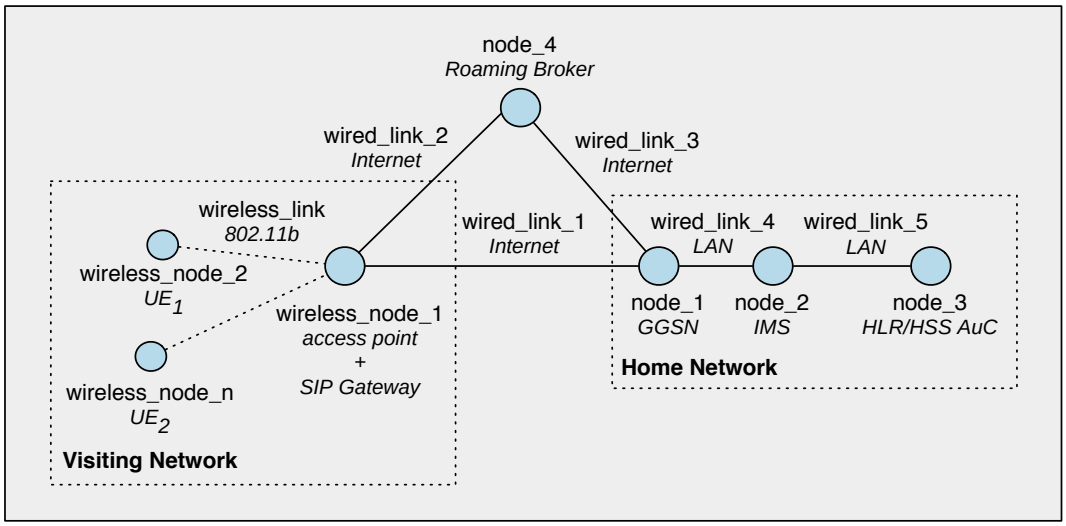


Figure 4.8: Simulation Model

To simulate the VN, we relied on the wireless model provided by NS-2. This model essentially consists of the MobileNode at the core,with additional supporting features that allows simulations of multi-hop ad-hoc networks and wireless LAN (NS-2, 1995). The MobileNode object is a split object. The C++ class MobileNode is derived from parent class Node. A MobileNode thus is the basic Node object with added functionalities of a wireless and mobile node like ability to move within a given topology, ability to receive and transmit signals to and from a wireless channel, etc. A major difference between them, though, is that a MobileNode is not connected by means of physical Links to other nodes or mobile nodes. The components supported in the NS-2 wireless model are Channel, Network interface, Radio propagation model, MAC protocols, Interface Queue, Link layer and Address resolution protocol model (ARP). Moreover, the support of SIP messages processing was added. Figure 4.9 describes our additions to the traditional wireless node.

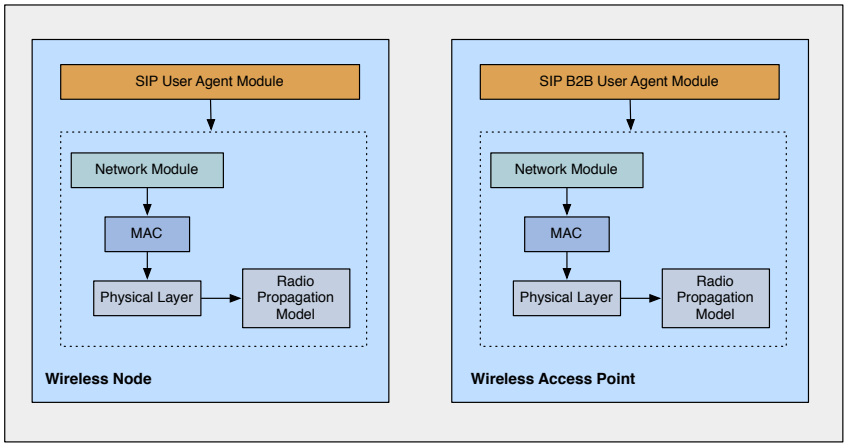


Figure 4.9: Wireless Nodes Model

4.4.3 Simulation Parameters

As depicted in Figure 4.8, the simulation environment comprises four types of nodes: the wireless node (*UE*), the VN, the RB and the HN. The HN includes the GGSN (*Gateway GPRS Support Node*), the IMS and the HSS-HLR/AuC. The initial inputs for our simulation model are divided in: wireless, network, and application layer parameters.

4.4.3.1 Wireless Parameters

The physical, medium access and network layers in the simulation control parameters such as channel type, radio propagation model, network interface type, antenna model, and MAC layer protocol. The parameters used in the simulation are depicted in Table 4.2.

Table 4.2: Wireless Parameters

Parameters	Values
Radio propagation model	Shadowing
Path loss exponent	4.0
Shadowing deviation	9.6
Reference distance	1 m
MAC Layer	802.11
Data-rate	11 Mbps
Antenna model	Omni-antenna

NS-2 implements three propagation models: Free Space, Two Ray Ground, and Shadowing (*Ricean and Rayleigh are not implemented in NS-2*) to predict the signal power received by the wireless nodes. The signal strength is used to determine whether the frame is transmitted successfully (NS-2, 1995). In NS-2 the Free Space model is used to simulate path loss of wireless communication when line-of-sight path exists between transmitter and receiver. Two Ray Ground model is used when line-of-sight path exists and reflection of ground is considered. Shadowing model simulates shadow effect of obstructions between the transmitter and receiver and it is recommended to simulate wireless channel in in-door environment. Thus, according to the findings in (Xiuchao, 2004), if we want to simulate a 802.11b 11 Mbps PC card in closed environment, the propagation model should be shadowing and shadowing factor should be selected to the specific environment simulated i.e. *office*. Here, an example of the TCL script used for the simulation:

```
#The propagation model is Shadowing model.
Propagation/Shadowing set pathlossExp_ 4
Propagation/Shadowing set std_db_ 0

//Data Rate
Phy/WirelessPhy set bandwidth_ 11Mb

//Tx Power (15dBm)
Phy/WirelessPhy set Pt_ 0.031622777

 //Collision Threshold
Phy/WirelessPhy set CPThresh_ 10.0

//Carrier Sense Pw (-94dBm)
Phy/WirelessPhy set CSThresh_ 3.1622777e-14

 //Rx Power Threshold
Phy/WirelessPhy set RXThresh_ 3.1622777e-13
```

In particular, to simulate the Physical Layer impairments in the VN we used the Shadowing radio propagation model and a path-loss exponent of 4 and a shadowing deviation of 9.6 to simulate an office environment with soft partition, as illustrated in Table 4.2. In addition, the antenna model used for the wireless nodes was omnidirectional with no diversity techniques implemented.

The MAC layer protocol used in the simulation was the IEEE 802.11 CSMA/CA (*Carrier Sense Multiple Access / Collision Avoidance*) operating in DCF (*Distributed Coordination Function*) mode, providing a data rate of 11 Mbps.

4.4.3.2 Network Parameters

As depicted in Table 4.3,we consider that the VNs are connected to the Internet through an ADSL (*Asymmetric Digital Subscriber Line*) link offering a bandwidth of 2 Mbps which is the average broadband speed offered in the US according to the last OECD report (OECD, October 1999). The Internet link connecting the RB to the HN was assumed of 20 Mbps as we consider the RB has faster Internet link than the VNs i.e. *a professional ADSL link* . Furthermore, in terms of transmission delay (Δ_{HN}) we assumed 10 ms in the links connecting the GGSN (*GPRS Gateway Node*), the IMS and the AAA server in the HN. The propagation delay was calculated based on the maximum packet size and the bandwidth supported by the local area network in the HN. Moreover, we also considered an Internet delay (Δ_i) of 150 ms based on (Paxson, 1997).

Table 4.3: Simulation Parameters

Parameters	Values
VN Internet bandwidth	2 Mbps
RB Internet bandwidth	20 Mbps
Δ_{HN}	10 ms
Δ_i	150 ms

4.4.3.3 Application Layer Parameters

Traffic is necessary when evaluating protocol performance. Thus, for this simulation we relied on CBR (*Constant Bit Rate*) sources generating VoIP traffic. The number of sources varied to create different traffic levels in the network. In this context, we created four simulation scenarios, each one with a variable number (*1,5,10, and 15*) of wireless nodes continuously transmitting VoIP traffic to the wireless network, as illustrated in Table 4.4. The maximum number sources was set to 15 because it has been shown that 802.11b cannot support more than 15 simultaneous VoIP sources (Hole & Tobagi, 2004) (Coupechoux et al., 2004).

The SIP message size depends on the nature of the SIP message i.e. *SIP-REGISTER, SIP-INVITE, etc.*, in this context the maximum SIP message size was set to 587 bytes.

Table 4.4: Simulation Parameters

Parameters	Values
Number of VoIP sources	variable [1,5,10,15]
SIP Pkt. Size	587 bytes
VoIP codec	GSM - 13.3 kbps

4.4.4 Simulation Outputs

The outputs produced by computer simulations show the performance of the SIP-based signaling protocol in our proposed roaming architecture. Simulations provide lots of data that is often hard to analyze thus, before analyzing the data extracted from the simulation we consider necessary to establish which parameters must be evaluated and then collect the statistics needed to initiate the analysis. For our analysis, the performance metrics are the following:

1. *Average Wireless Transmission Delay* is the average amount of time a SIP message spends since its departure from the wireless node or the access point until its arrival to the wireless node or the access point. This metric is expressed in seconds and consider re-transmissions due to collisions or errors in the wireless channel. It also take into account the back-off time introduced by the medium access algorithm in the wireless network.

2. *Average Wireless Processing Delay* indicates the amount of time the access point in the VN spend processing SIP signaling messages. This metric is expressed in seconds and takes into account buffering times, SIP message processing, and SIP message re-routing or re-direction.

3. *Overall Network Registration Delay* include all possible delays i.e. *wireless transmission and processing delay* from the moment a SIP-REGISTER message is generated to the moment a SIP-200 OK message is received (*indicating a successful network registration in the VN*. This metric is expressed in seconds and takes into account *wireless transmission delay, processing delay, buffering delay and Internet delay*.

4. *Overall Session Initiation Delay* include all possible delays i.e. *wireless transmission and processing delay* from the moment a SIP-INVITE message is generated to the moment a SIP-200 OK message is received (*indicating a successful establishment of a VoIP session*. This metric is expressed in seconds and takes into account *wireless transmission delay, processing delay, buffering delay and Internet delay*.

5. *Signaling Overhead* indicates the amount of roaming SIP signaling messages (*SIP-REGISTER, SIP-INVITE, SIP-200, SIP-407, etc.*) transmitted during the simulation. This metric is expressed in bytes and as a percentage of the total bandwidth consumption.

4.4.4.1 Verification of Simulation Outputs

Verification of simulation outputs is important to ensure that the simulation is performing as intended. In this context, the *ns-2* model developed for this thesis was verified by tracing approach. On-screen tracing verifies that the program is actually following the steps it is supposed to do. For example, signaling message transmission, data packet transmission, SIP registration process, SIP session initiation and termination, etc. All this processes can be visualized on the screen using the *NAM* (*Network Animator*) tool provided by the simulator.

Computer simulations are driven by random numbers (*assuming the same random seed*), several runs were performed (*same random seed but varying the run time*) to ensure the simulation is in steady state region. Thus, simulation runs of duration 50, 100, 150, 300, 600 and 1200 seconds were performed. As a result, it was found that the outputs from the simulation of lengths 600 seconds were within the $+/-10\%$ interval, which was fair enough to identify the steady state region in the simulation. Hence, a simulation time of 600 seconds after transient period of 600 seconds was chosen.

The results presented in this thesis are the average of twenty independent replications of the terminating simulation. In this respect, the independence of replication was accomplished by using different random number seeds for each simulation. The confidence interval of the simulation outputs presented in this chapter are 95%.

4.4.4.2 Validation of Simulation Outputs

The purpose of validation is to ensure the correctness of simulation outputs. In this perspective, we performed a comparison of Rui Prior's SIP model with our enhanced Roaming-SIP model to assure the validity of the simulation outputs. For this validation, two sets of simulations under the same scenarios were performed. The first set of simulations used the Prior's SIP model as signaling protocol whereas the second set of simulations used the enhanced Roaming SIP model. The results provided by both scenarios demonstrated, that under the same conditions, the signaling delay achieved by Prior's SIP was within the error interval of the results for Roaming SIP. This good agreement provided some confidence in the outputs of the simulation.

4.5 Simulation Results

This section presents the results obtained through computer simulation. As mentioned in section 4.4.4, we are mainly interested in the signaling delay a UE experiences when attempting to register into a VN, and the signaling delay when initiating a session from a VN. The results in this section illustrate the overall network registration and overall session initiation delay.

4.5.1 Network Registration Delay

As stated, the overall network registration delay takes into account the different delays encountered during the network registration process (*wireless transmission delays and processing delays*), as illustrated in Table 4.5. Moreover, we present the impact of simultaneous VoIP wireless sources in the network. For this, we assumed a registration rate proportional to the number of mobile nodes in the wireless network, this is, n registration per second. Where n varied from *1, 5, 10, and 15*.

Table 4.5: Network Registration Delay

Sources	Wireless Tx Delay	Wireless Proc. Delay	Overall Delay
1	0.0075	0.000151	1.13036
5	0.8200	0.000856	3.37595
10	1.3515	0.00155	6.86792
15	1.5133	0.00162	9.09095

The results illustrated in Figure 5.21 show that in the case a single UE in the VN attempts network registration, it only requires 1.13 seconds to register in the VN. Nevertheless, as the number of wireless sources increased, the network registration delay also increased driven mainly by the wireless transmission delay. The worst case of network registration delay was achieved when the UE attempted to register into a VN where 15 UE were performing VoIP calls; in this scenario the UE needed 9.98 seconds to register. It is worth-mentioning that the highest delay components were introduced by the wireless network where for the 15 UEs the VN experienced an average wireless transmission delay of 1.51 seconds compared to 7 ms in the single UE scenario. From this, we can infer that as the load in the VN increases the overall network registration delay increases. Thus, with an average wireless delay of 1.5 seconds in the network the UE experiences an overall network registration delay of 9.09 seconds.

4.5.2 Session Initiation Delay

Session initiation delay expresses the amount of time spent by the UE when attempting to initiate a VoIP call from a VN. The results obtained through the simulation are presented following the same approach as in section 4.5.1. Moreover, we present the impact of simultaneous VoIP wireless sources in the network. For this, we assumed a session initiation rate proportional to the number of mobile nodes in the wireless network, this is, n registration per second. Where n varied from *1, 5, 10, and 15*.

Similarly to the outcomes produced by the network registration simulation, session initiation is also affected by the increasing number of traffic sources in the VN. Thus, the session initiation delay was also impacted by the increasing wireless transmission delay in the VN.

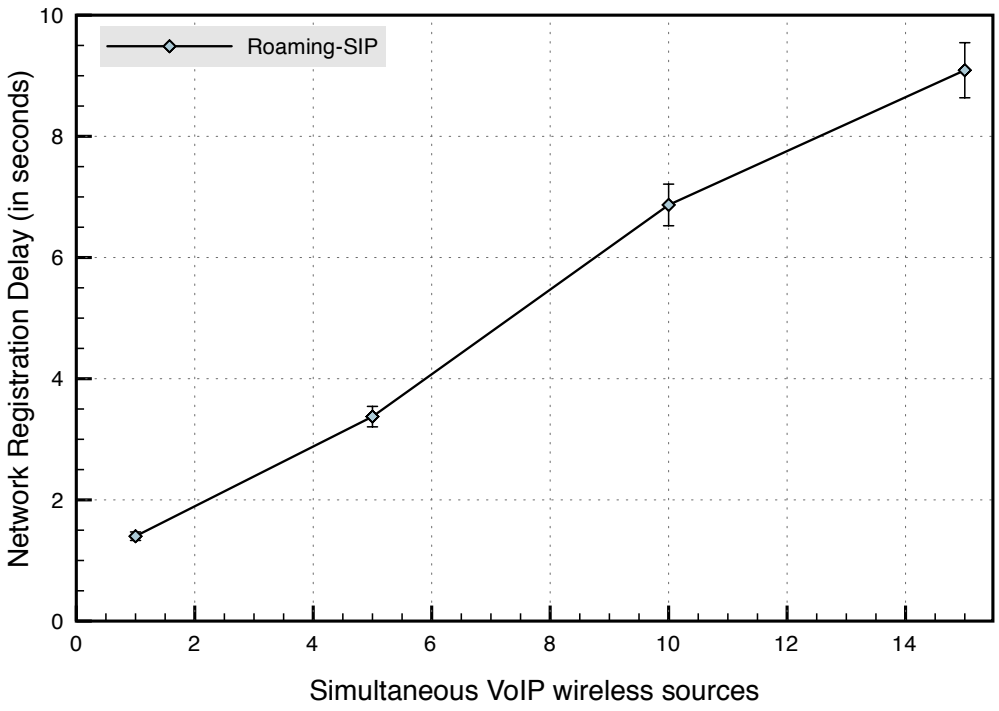


Figure 4.10: Network Registration Delay

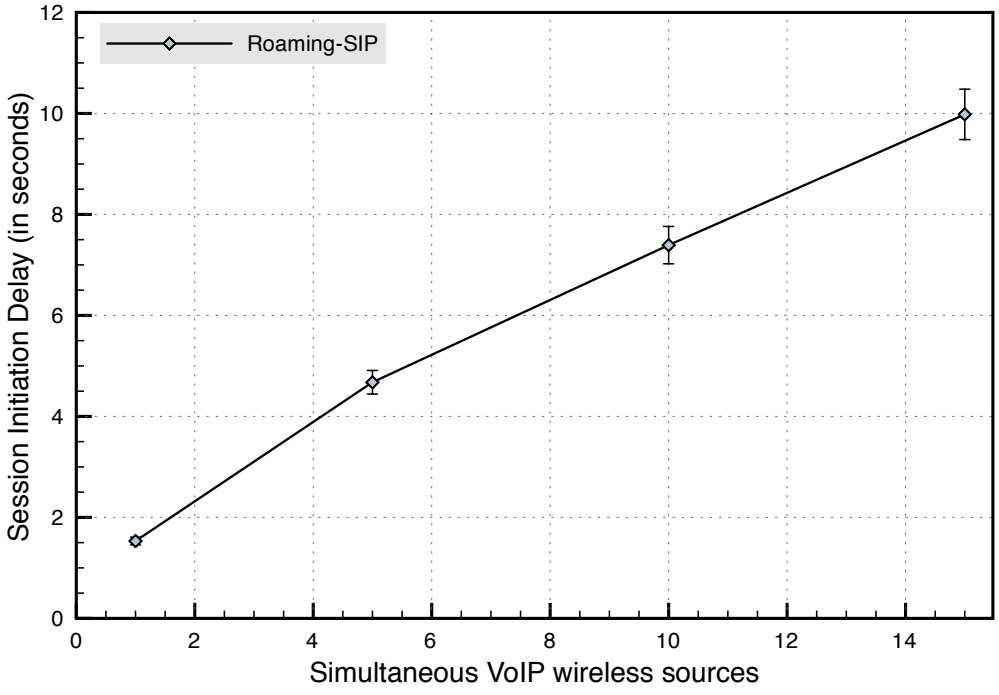


Figure 4.11: Session Initiation Delay

The results expressed in Figure 4.11 display that in the single traffic source scenario, the UE required 1.53 seconds to initiate a VoIP call from the VN. On the other hand, when sharing the wireless network with 15 traffic sources, the UE required 9.98 seconds to initiate the call. In this context, it is important to notice that session initiation delay is slightly higher than network registration delay, however this is due to the number of SIP messages required to complete the process (*session initiation requires an extra message compared to network registration*).

Table 4.6: Session Initiation Delay

Sources	Wireless Tx Delay	Wireless Proc. Delay	Overall Delay
1	0.0069	0.000151	1.5317
5	0.7859	0.000867	4.6762
10	1.3176	0.00161	7.3922
15	1.4973	0.00179	9.9809

From the results in Table 4.6, we can imply that again the overall delay is mainly affected by the wireless transmission delay. On the other hand, the size of the SIP messages contribute makes them relatively easy to process, however we can observe that as the number of wireless traffic sources increases the processing delay also increases. In the 15 source scenario, the UE experiences a processing delay approximately 10 times higher. Nevertheless, the processing delay is still negligible as it is in the order of $1ms$.

4.6 Feasibility Study

To validate the performance and the feasibility of implementing a SIP-based roaming architecture, we decided to analyze the performance of the Roaming-SIP and the SIP servers in the architecture under a more realistic scenario (*testbed*). Furthermore, we were interested on studying the feasibility of implementing the SIP-based roaming architecture using off-the-shelf equipment.

In this respect, we were mainly interested on evaluating the signaling performance in terms of network registration delay, and signaling overhead under heavy SIP traffic conditions. The rationale behind these metrics is that network registration delay could impact on roaming mechanisms *i.e. vertical handover* whereas signaling overhead has an impact on bandwidth consumption. In this respect, we consider signaling overhead as an important parameter when developing and implementing a signaling protocol, hence we decided to display the testbed signaling results as follows: bandwidth consumption in both wireless *(802.11b)* and wired domain *(Internet link)* expressed in percentage of the total bandwidth and Bandwidth consumption expressed in Mpbs.

4.6.1 Testbed Implementation

Our testbed was developed using the open source SIP server *OpenSER* (OpenSER, 2006) with the necessary modifications to support our extensions. The SIP server was installed on a computer running Linux Fedora Core 5, this server emulated the IMS in the HH. The HSS was emulated by a RADIUS+MySQL database. The SIP clients (*performing the registration*) were installed on five computers running SunOS 5.10, they represented the mobile users attempting to register in the VN. The SIP B2BUA emulating the VN was installed on a wireless router (*Linksys WRT54G*) running dd-wrt (dd wrt, 2006), a linux-based operating system. It will also followed the required modifications to support our SIP extensions. To simulate hundreds of clients we modified the SIP client *sipsak* (Sipsak, 2002) to support our extensions and multi-threading. Although the access point supports of theoretical data rate of 54 Mbps, to validate the simulation outputs against these results, the wireless clients were equipped with 802.11b wireless cards (*providing a theoretical data rate of 11 Mbps*).

4.6.2 Testbed Results

The main difference between our computer simulation analysis and our testbed is that in the latter we did not take into account the Internet delay. The rationale behind this is that we were interested on a performance evaluation of our Roaming-SIP protocol and the SIP servers under heavy SIP traffic conditions. To achieve this, as stated previously, we varied the number of simultaneous SIP sources (*mobile users*) attempting to register into the VN. Moreover, to increase the accuracy of the testbed, the results presented in this section are the average of thirty independent trials.

4.6.2.1 Network Registration Delay

We decided to test our SIP servers under two different scenarios. The first was a wired scenario, the SIP requests to the RB arrive from the Internet, with the clients and the server interconnected through a 100 Mbps link. The other scenario recreates the UE performing SIP requests through the wireless link thus the client and the server were interconnected through a 802.11b wireless network. We did not want to test the impact of the load in the wireless network hence we place only one client in the wireless network. The remaining SIP requests were performed through the ethernet interface of the wireless router.

For evaluation purposes we varied the number of SIP registration sources (*30, 60, 120, 240, 300*). From the results illustrated in Figure 4.12, we can infer that the network registration delay in the wireless network increases with the number of wireless nodes attempting to register their new SIP address. The wireless scenario (*in the VN*) exhibited an average network registration delay of ≈ 9 ms under the lowest signaling traffic conditions (*30 sources*) and ≈ 15 ms under the highest traffic conditions (*300 sources*).

On the other hand, as illustrated in Figure 4.13, the wired scenario also experienced higher delay under high signaling traffic conditions. Under low signaling

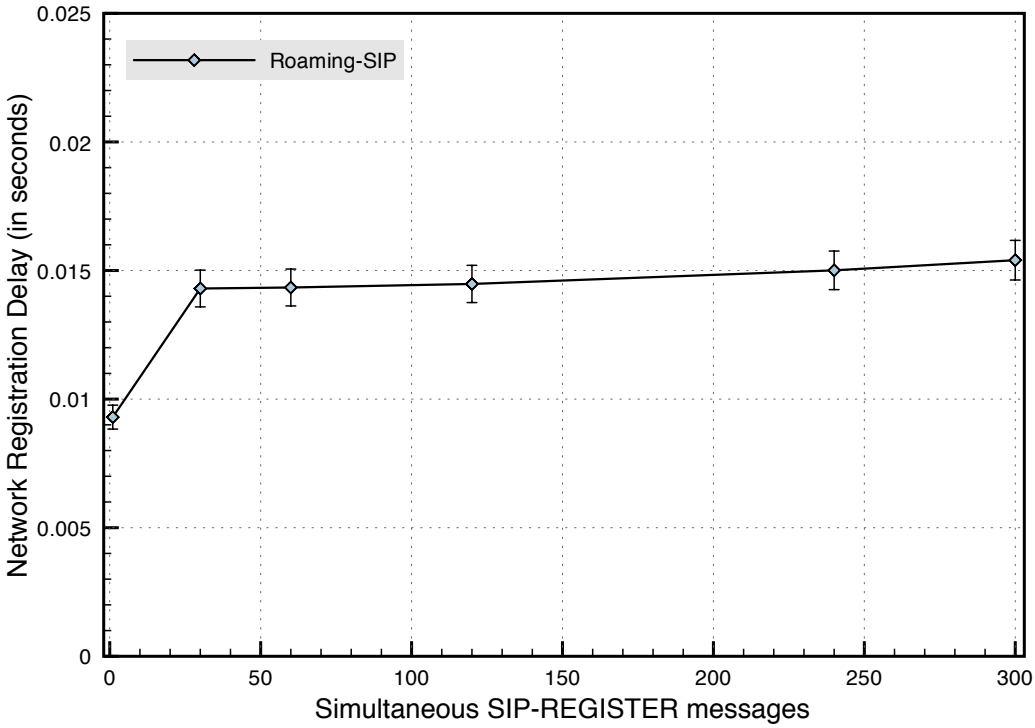


Figure 4.12: Network Registration Delay (in the VN)

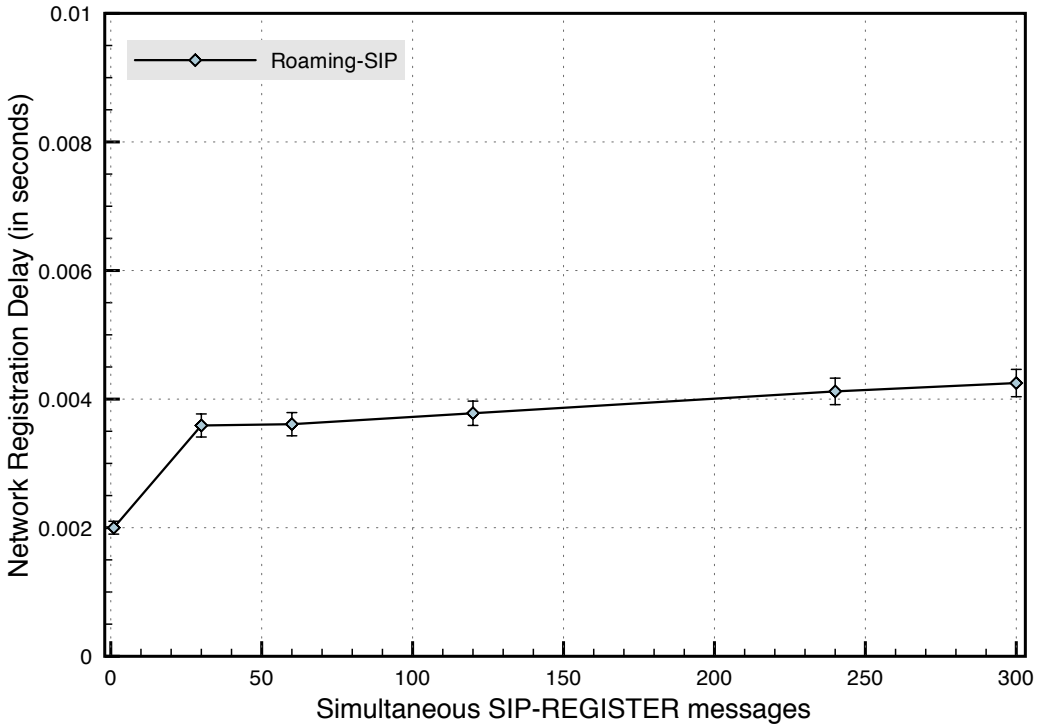


Figure 4.13: Network Registration Delay (in the HN)

traffic conditions (*30 sources*), we experienced an average network registration delay of ≈ 2 ms and a delay of ≈ 4 ms under high signaling traffic conditions (*300 sources*)

If we compare both results, as depicted in Figure 4.14, we can observe that the wireless scenario experienced under all traffic conditions a network registration delay approximately 10 ms above the delay in the wireless scenario. From this, we can imply that although processing delay has a direct impact on the network registration process, the transmission delay also plays an important role on the overall end-to-end delay.

From these results, we observe that two kinds of delay impact on the performance of the signaling protocol: the processing and the transmission delay. Nevertheless, we also observe that the processing delay introduced by three hundred sources it is slightly higher than the delay with thirty sources in both scenarios. Hence, confirming that processing delay do not have a great impact on the SIP signaling protocol as an increment of the SIP server capacity is enough to cope with this issue. However, transmission delay must be addressed carefully through the maintenance of an acceptable number of wireless node within the wireless networks.

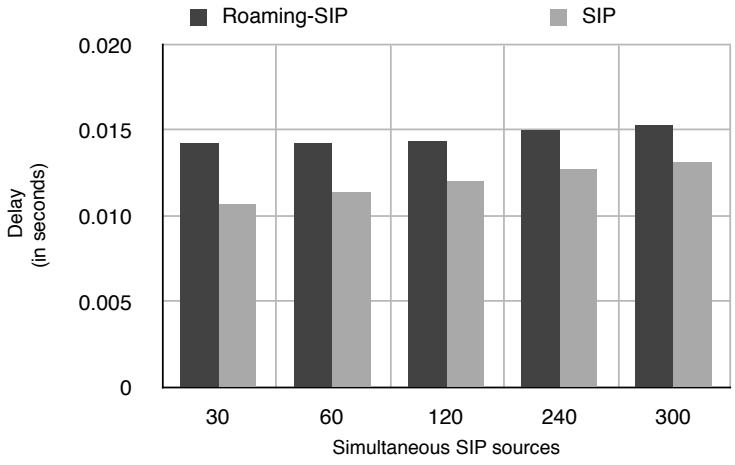


Figure 4.14: Network Registration Delay

4.6.2.2 Signaling Overhead

In terms of signaling overhead in the wireless network, our proposal displayed almost the same performance as SIP, as depicted in Figure 4.15. The reason is that our extensions only add few bytes to the SIP messages ($\approx 100 - 150$ *bytes*) thus the signaling overhead introduced by our protocol (*Roaming SIP*) and SIP remains comparable, as illustrated in Figure **??**.

As stated previously, the access point operating as VN relies on two network interfaces to provide connectivity: the wireless interface, providing data rates up to 11 Mbps and the wired interface which connects the UWN to the Internet. In

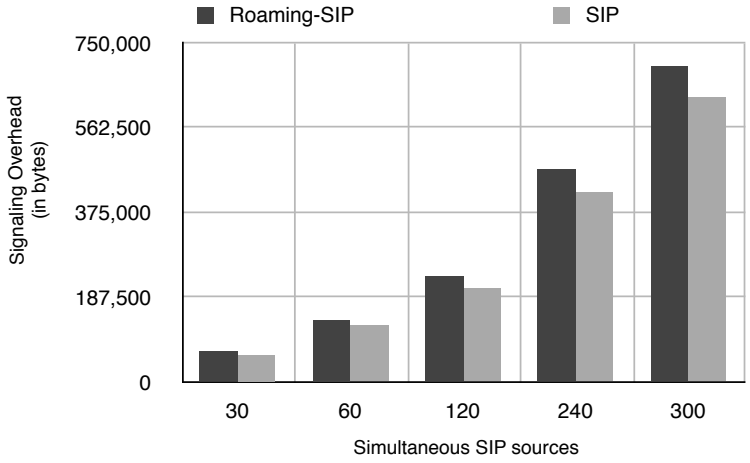


Figure 4.15: Signaling Overhead (in bytes)

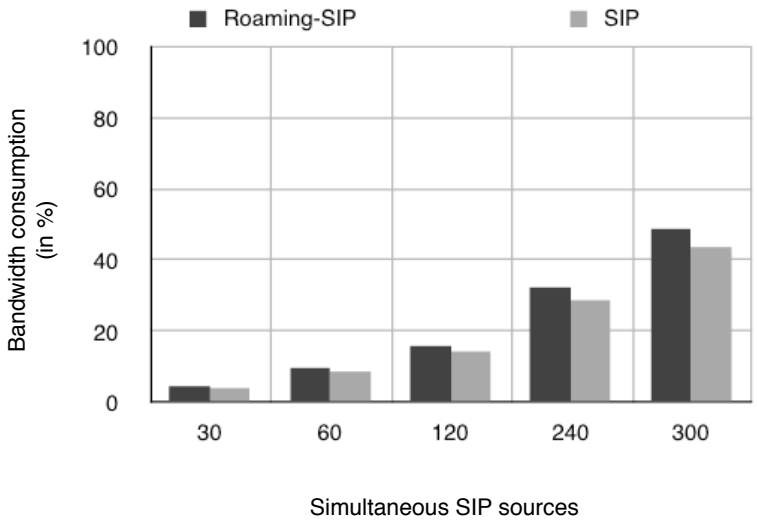


Figure 4.16: Wireless Bandwidth Consumption (in %)

this interface, the data rates are generally smaller than the offered by the wireless domain i.e. ($\approx$ 2 *Mbps in some ADSL links*). This characteristic makes the additional overhead introduced by our proposal more noticeable. Nevertheless, as depicted by Figures 4.15 and 4.16, a significant difference between SIP and our Roaming-SIP starts when more than a hundred simultaneous SIP sources attempt to register. Nevertheless, these scenarios are unlikely because wireless local area networks hardly support more than 15 simultaneous VoIP calls as shown in (Garg & Kappes, 2003b) (Hole & Tobagi, 2004) (Coupechoux et al., 2004). In conclusion, the testbed results indicate that our SIP-extensions do not diminish the efficiency of SIP protocol when enabling roaming in heterogeneous multi-operator wireless networks. We have also demonstrated that due to the simplicity and versatility of SIP, our SIP-based roaming architecture can be deployed under current wireless architecture without requiring major changes.

4.7 Conclusion

In this chapter we have presented our SIP-based roaming architecture to enable service mobility between cellular and UWNs. In our architecture, the RB is in charge of establishing mutual trust between the HN the VNs. In addition, other contributions of our work are the SIP extensions to enable broker-based access control and SLA information exchange.

For evaluation purposes, we analyzed the impact of traffic congestion in the VN on the roaming process through computer simulation. The results obtained in terms of SIP-based signaling delay showed that our architecture and our SIP-extensions do not have a negative impact on the performance of SIP. Furthermore, to verify the feasibility of implementing a SIP-based roaming protocol, we have deployed a testbed. The testbed outputs confirmed that the performance of our proposed Roaming-SIP is comparable to the performance of SIP. The testbed also provided an important insight regarding the wired domain (*Internet link*) of the access point operating as a VN. In this context, we must consider that, in most of the cases, the Internet bandwidth is smaller than the wireless bandwidth. Thus, it is necessary to rely on efficient access control mechanisms in the VN to provide and maintain efficient traffic balance between the wireless and the wired domain. Finally, we could also demonstrate that the implementation of a SIP-based roaming architecture under current wireless architecture is possible and do not require major changes in current wireless architecture (*a simple software upgrade could transform a wireless router into a VN*).

Next chapter presents our work on implementing and evaluating new SIP extensions to enhanced the authentication and authorization process.

Chapter 5

Enabling Fast & Secure Service Mobility

The SIP-based Roaming Architecture described in chapter 4 enables roaming in heterogeneous multi-operator wireless environments. In this architecture, the HN is always the roaming decision maker as it is determines, based on the mobile user's and VN's credentials, whether or not the mobile user can roam in to the VNs. That is mobile user authentication and authorization are always performed by the HN and from the HN. Furthermore, as VNs are not allowed to communicate directly with the HN, all these roaming signaling messages must pass through the RB. Thus, as confirmed by the simulation results, this process clearly contributes with an increment in the overall delay of the network registration and session initiation process.

Our objective is to reduce the roaming signaling exchange caused by the authentication and authorization mechanisms in the network registration and session initiation process while providing robust network security. To achieve this, we rely on the use of PKI (*Public Key Infrastructure)* and PMI (*Privilege Management Infrastructure*) techniques deployed on top of our architecture. Indeed, for authentication purposes we rely on X.509 PKCs (*Public-Key Certificates*) (ITU, March 2000b) to prove the identity of the key elements of our architecture: the UE, the VN, and the RB. Moreover, for authorization, we propose the use of PMI and XML Attribute Certificates to define the rights and roles of the key elements of our architecture.

The outline of this chapter is structured as follows: section 5.1 describes our motivation for enhancing our SIP-based roaming architecture and achieving fast & secure service mobility and our contributions to this subject matter. Then, section 5.2 provides essential PKI background. Section 5.3 describes how PKI is integrated in the roaming architecture. Section 5.4 presents the Privilege Management Infrastructure and its role in our architecture. Further, section 5.5 presents a new approach of XML Attribute Certificates. In addition, section 5.6 describes our PMI-enhanced roaming architecture while section 5.7 provides a realistic scenario of fast & secure roaming using the proposed architecture. The computer simulation environment used to evaluate the architecture is described in section 5.8, followed by the simulation results, which are presented in section 5.9. Finally,

section 5.10 concludes this chapter.

5.1 The need for Fast & Secure Service Mobility

Heterogeneous multi-operator roaming raises several challenges, among these we find the need for fast and secure service migration. Although, as stated in previous chapter, this type of roaming follows economic rather than technological reasons, for the mobile user's commodity it must be performed as fast as possible without compromising the network security. In this connection, robust security mechanism are required when authorizing a VN to become an extension of the cellular network or when a UE attempts to roam in to a VN.

The proposed SIP-based roaming architecture, through the RB and the use of user's credentials i.e. *username and password* grants access control mechanisms for both UEs and VNs. Nevertheless, as the decision maker is always the HN and the VNs are not allowed to communicate directly with the HN, all the signaling exchange must pass trough the RB. Due to the signaling triangulation issue introduced by the RB approach, as shown in the results in chapter 4, the roaming architecture introduces a slight delay in the network registration and session initiation process. Moreover, it is also shown that a high delay component is introduced by the Internet. In this context, in order to enhance the roaming architecture we propose:

- the reduction of the number of signaling message exchanged between the elements of the architecture.
- a new approach of mobile user's credential exchange.
- the delegation of authentication and authorization process to other elements, specifically the VNs. This under the supervision of the RB and the HN.

It is worth-mentioning that in spite of the enhancements, the HN still remains the roaming decision maker and the only entity can determine whether or not a mobile user belongs to its network and the rights of such user to perform any service. Here, you might ask how can we delegate authentication and authorization while maintaining the HN as the roaming decision maker. In this perspective, we rely on an enhanced Roaming Architecture that uses PMI & SIP-embedded XML Attribute Certificates to provide a suitable distributed authentication and authorization environment.

5.1.1 Contributions

The proposed roaming architecture enables mobile users to roam into wireless networks owned or managed by different providers. Nevertheless, core roaming process such as mobile user authentication and authorization are performed in a centralized manner from the HN. This, introduces certain delays and potentially yields considerable signaling overhead in the HN. Although, the overhead issue

can be overcome by relying on a distributed HN architecture, we efficiently delegate the authentication and authorization of mobile services to the VNs. In this context, our main contributions to heterogeneous multi-operator roaming are:

1. *The implementation of a trust infrastructure on top the proposed roaming architecture*. This, as described in section 5.3, is possible through the integration of PKI in our architecture.

2. *A PMI-Enhanced Roaming Architecture*. The use of PMI in our architecture allows the delegation (*decentralization*) of mobile service authentication and authorization. As explained in section 5.4, we relied on PMI to perform such process locally in the VN rather than remotely as in our original proposal. The main result of such enhancements is the improvement of the roaming process in terms of security and delay, for more details please refer to section 5.9.

3. *SIP-embedded XML Attribute Certificates*. The application of PKI and PMI mechanisms in the roaming architecture require the exchange of user's credentials and certificates between certain elements of the architecture. Thus, we propose an extension to the SIP protocol to enable the the transport of XML Attribute Certificates within SIP messages *i.e. SIP-REGISTER and SIP-INVITE*. The SIP-embedded XML Attribute Certificates along with the PMI-Enhancements are the most valuable contributions to the roaming architecture.

5.2 Public Key Infrastructure

A Public Key Infrastructure or PKI is a management system designed to administer asymmetrical cryptographic keys and public key certificates. It acts as a trusted component that guarantees the authenticity of the binding between a public key and security information, including identity, involved in securing a transaction with public key cryptography. Thus, PKI enables users of unsecure public networks such as the Internet to securely and privately exchange data (*and money*). Furthermore, PKI also provides the following benefits:

- Origin authentication: verification that a document or a message was created by the person or entity claimed.
- Content integrity: securely prove that a document has not been altered.
- Content confidentiality (*encryption*): protecting the content of a document or message so that it can be read only by specifically identified users.
- Non-repudiation: security service whereby the creation of a document or message cannot subsequently be denied.

Traditionally, the PKI authentication model comprises three entities: the CA, the subject, and the identity verifier. As illustrated in Figure 5.1, the CA certifies the subject's identity by issuing Public Key Certificates (*PKC*) for them.The

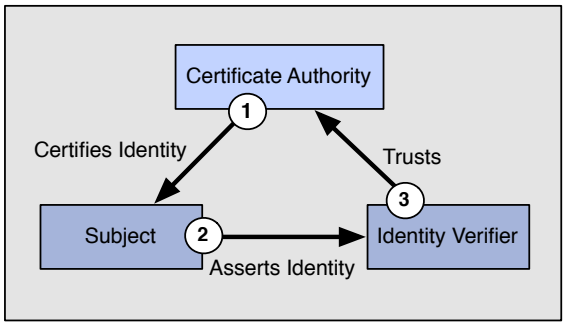


Figure 5.1: PKI Authentication Model

identity verifier determines as to whether or not the asserted identity is correct and finally, the subject , which is the certificate holder. The identity verifier *trusts* the CA as the authority certifying the subject's identity. If a subject's certificate is not issued by that CA, then the identity verifier must locate a certification path of certificates from that of the entity to one issued by the CA.

5.2.1 Creation of PKC

A PKC or identity certificate is an electronic document that incorporates a *digital signature* that binds together a public key with an identity i.e. *information such as the name of a person or an organization, IP address, and so forth*. Thus, such certificates are used to verify that a public key belongs to a specific individual.

In a classic public key infrastructure certificates are issued by CAs to other CAs or to end-entities i.e. *end-users, devices, web servers, processes, software, etc*. Additionally, CAs may also self-issue certificates to themselves. Thus, certificates issued to CAs are known as *CA certificates*, and certificates issued to end-entities are referred to as *end-entity certificates*. The difference between both CA and end-entity certificate is defined in the Basic Constraints certificate extension.

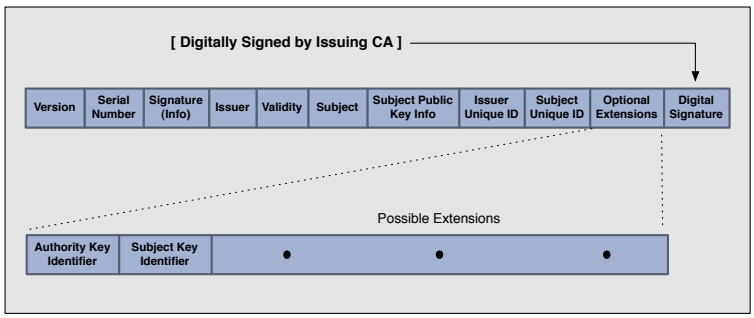


Figure 5.2: X.509 Version 3 Public Key Certificate

The basic structure of PKCs, as depicted by Figure 5.2 contains the following information:

- **Version:** the version number of certificate.
- **Serial number:** an integer uniquely assigned by the CA to each certificate.
- **Signature:** algorithm identifier for the algorithm and hash function used by the CA in signing the certificate.
- **Issuer:** the entity that has issued and signed the certificate.
- **Validity:** the time interval during which the CA warrants that it will maintain information about the status of the certificate.
- **Subject:** the entity associated with public key found in the subject public key field.
- **Subject public key info:** the public key being certified and the algorithm which this public key is an instance of.
- **Issuer unique identifier:** used to uniquely identify an issuer in case of name re-use.
- **Subject unique identifier:** used to uniquely identify a subject in case of name re-use.
- **Extensions:** allows addition of new fields to the structure.

5.2.2 Digital Signature

Digital signature is a type of asymmetric cryptography used to simulate security properties of a signature in digital, rather than written form. Digital signature schemes traditionally rely on two algorithms, one for signing which involves the user's secret or *private key* (SK), and one for verifying signatures which involves the user's *public key* (PK). The output of the signature process is called the *digital signature*.

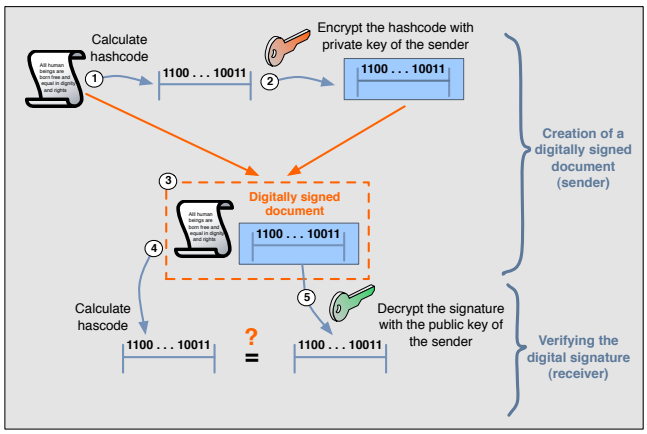


Figure 5.3: Digital Signature Creation and Verification

Digital signature, like written signatures, is used to provide authentication of the associated input, usually called a message. Consequently, digital signatures are used to create PKI schemes in which a user's public key is tied to a user by a digital identity certificate or PKC. A digital signature scheme typically comprises three algorithms:

1. A key generation algorithm G that randomly produces a *key pair* (PK, SK) for the signer. Where, PK is the verifying key, which is to be public, and SK is the signing key, to be kept private.

2. A signing algorithm S, that on input a message m and a signing key SK, produces a signature σ.

3. A signature verifying algorithm V, that on input a message m, a verifying key PK, and a signature σ, determines the validity of the digital signature.

For digital signature validation, two main properties are required. First, signatures computed honestly should always verify. That is, V should accept ($m, PK, S(m, SK)$) whenever SK is the secret key related to PK, for any message m. Secondly, it should be hard for any adversary, knowing only PK, to create valid signatures.

As illustrated in Figure 5.3, the process of creating and verifying a digital signature is the following. The sender, calculates the hash code of the document to sent, here the message becomes the *message digest*. Then, the sender uses the signing key SK to encrypt the message digest. At this point, the message becomes a digitally signed document. Now, the next step is for the receiver to verify the digital signature. To do this, the receiver calculates the hash code of the digitally signed document. Then, the receiver decrypts the signature using the public key, PK, of the sender and then compares the decrypted hash code with the calculated. If both codes match, then the digital signature is valid. Thus, the verifier can prove that the sender is the real author of the document (*no repudiation*).

5.2.3 Certification Path

A Certification Path is an ordered sequence of PKCs that enables a certificate user to verify signatures on the certificates along the path, and thus enables the user to obtain a certified public key of the entity that is the subject of the last certificate.

There are different path construction hierarchies, the more common is the strict hierarchy of CAs as illustrated in the example Figure 5.4. The rectangles represent CAs, the arrows certificate issuance, and the sectioned rectangles represent certificates. Here, we provide an example of what happens when User1 sends a digitally signed e-mail to User2. We assumed that User1's verification certificate is conveyed along with the message itself. In this example, User2 is the relying party. Since we are attempting to verify a digital signature using User1's verification certificate, we need to construct a certification path between User1's certificate and trust anchor recognized by User2. In this case, CA_0 is the root CA and is by definition the common trust anchor for all users within this strict hierarchy. Basically, User2 wants to know if CA_0 has established a trust relationship (*directly or indirectly*) with the issuer of User1's certificate (CA_1 *in this*

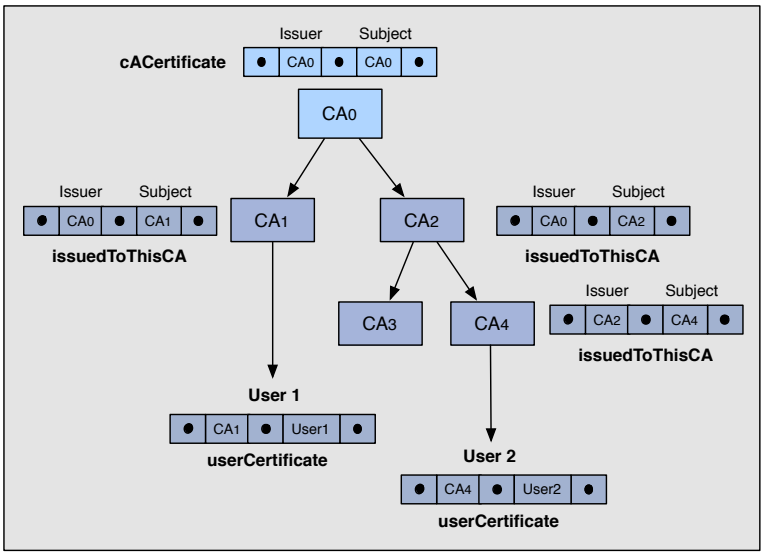


Figure 5.4: Path Construction - Strict Hierarchy

case). In other words, if the relying party is able to resolve the certification path $CA_0 \rightarrow CA_1 \rightarrow User1$, as consequence we would have a candidate certification path that could be submitted to the path validation logic.

The path certification construction within strict hierarchies is rather straightforward. As illustrated in Figure 5.4, paths are typically constructed in the forward direction, this means that we start with the target certificate and work our way to a recognized trust anchor. Thus, the relying party software will start with User1's verification certificate and works its way to CA_0. Since the relying party software knows CA_1 is the issuer of User1's certificate, this is accomplished by retrieving the *issuedToThisCA* element that is stored in the directory entry for CA_1. We then discover that CA_0 is the issuer of $CA_1's$ certificate, and we now have a complete candidate certification path between User1's certificate and trust anchor recognized by User2.

5.2.4 Certificate-based Network Access Control

Nowadays, digital certificates are the basis of most communication systems that require a high level of trust between the communicating parties, among such systems we find:

- VPN (*Virtual Private Network*) technologies. VPNs provide secure point-to-point or point-to-multipoint connections across the Internet.
- SSL (*Secure Socket Layer*) connections. Currently, SSL is the standard for Internet browser and server authentication as well as secure data exchange on the Internet.
- EAP (*Extensible Authentication Protocol*) negotiation schemes. EAP provides secure access to wired and wireless LANs.

This section focuses mainly on EAP as it is commonly used for wireless network access control. Mechanisms such as EAP-TLS (*Transport Layer Security*) (Aboba & Simon, 1999), EAP-TTLS (*Tunneled Transport Layer Security*), and PEAP (*Protected Extensible Authentication Protocol*) have been successfully implemented in wireless LANs. Not surprisingly, each of them presents advantages and disadvantages that can make them unreliable depending on the deployment conditions. Table 5.2.4 depicts some differences between such mechanisms:

Table 5.1: 802.1x EAP Types

802.1x EAP Types	TLS	TTLS	PEAP
Client certificate	required	optional	optional
Server certificate	required	required	required
WEP key management	yes	yes	yes
Authentication Attributes	mutual	mutual	mutual
Wireless Security	very high	high	high
Deployment Difficulty	difficult	moderate	moderate

For the enhancements of our SIP-based Roaming Architecture, we assumed a robust PKI-based access control in the HN, VNs and RB. Nevertheless, we do not rely on a specific mechanism. In this section we provide an example of EAP-TLS-based access control as it is a popular certificate-based access control mechanism currently deployed on certain wireless LANs such as the *Freephonie*, the SIP-based telephony service provided by the French provider *Free*.

5.2.4.1 Access Control with EAP-TLS

The EAP-TLS is based on TLS (*Transport Layer Security*) to provide protected cipher-suite negotiation, mutual authentication, and key management. Consequently, after the EAP-TLS negotiation is accomplished the two end-points can securely communicate within the encrypted TLS tunnel. During this process, the user's identity and password are not revealed.

TLS relies on the use of certificates for both user and server to authenticate each other. As a result a user, in addition to being authenticated, can also authenticate the network thus detecting forged UWNs. Both supplicant and authentication server need to have valid certificates when using EAP-TLS.

Figure 5.5 depicts the authentication process and EAP-TLS message exchange in a wireless LAN. Once the authenticator receives the supplicant's identity in EAP-response/id (*flow 3*), it initiates a RADIUS-access-Request, which also carries the supplicant's identity, to the authentication server. During this process the authentication server provides its certificate to the supplicant and requests the supplicant's certificate. Thus, the supplicant's validate the server's certificate and responds with an EAP-Response which contains the supplicant's certificate. The supplicant also initiates the negotiation for cryptographic material. Upon the supplicant's certificate validation, the authentication server responds with the cryptographic material for the session (*flows 5 and 6*). The session keys derived at both ends can be used for data encryption. The drawback of this approach is that

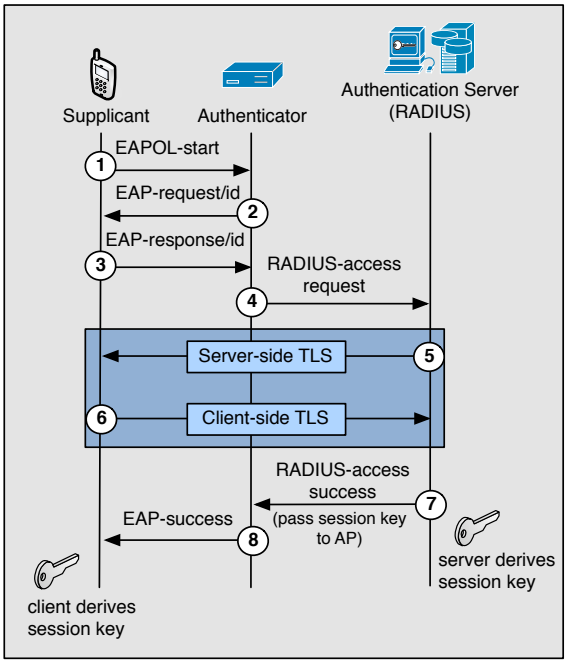


Figure 5.5: Message Flow of EAP-TLS

because both the supplicant and the authentication servers need to have valid certificates when using EAP-TLS, to some extend EAP-TLS is difficult to manage. Nevertheless, our architecture takes advantage of this characteristics and defines the UE as the supplicant, the AP in the VN as both the authenticator and the authentication server thus, the authentication is performed locally (*by the VN*). This is explained in details in the remaining of this chapter.

5.3 The role of PKI in our Roaming Architecture

Traditionally, the role of PKI is to bind individual, software or device identity to an electronic document. In our architecture, PKI plays an important role in the authentication process, primarily in the identification of entities providing or requiring roaming services. It is worth-mentioning that our research does not focus entirely on PKI but it relies on such technology to provide an efficient service mobility platform. Thus, we assume a robust and efficient PKI between the key-entities: the HN, the UE, the RB, and the VN. Here, we provide the guidelines for the implementation of the PKI and define the role of certain elements in our architecture.

As illustrated in Figure 5.6, four entities integrate the PKI in the roaming architecture. The cellular network, as specified in the assumptions in chapter **??**, is always considered the HN hence it plays the role of the CA root. On the other hand, the RB acts as an Intermediate CA. At the same time, the VN plays two roles as it acts as Authenticator and Authentication Server. Traditionally, for account-

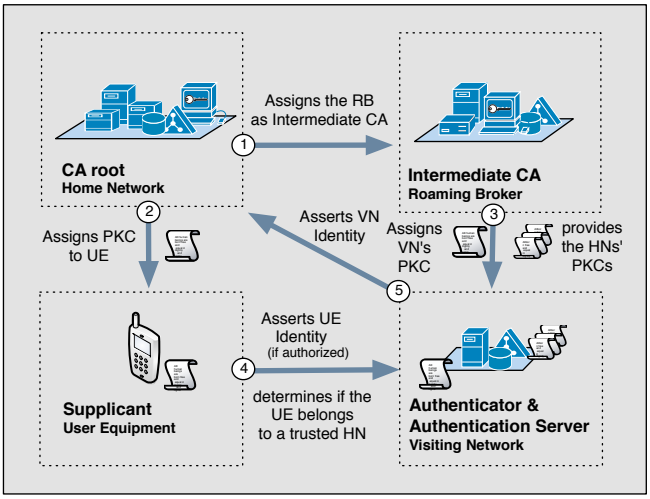


Figure 5.6: PKI in our roaming architecture

ing purposes the authenticator and the authentication server are independent devices, however in our architecture we decided to implement both functionalities on the access point as billing is left to the HN and out of the scope of our research. The implementation of an authentication server on an access point is feasible due to the increasing computing and storage capacity of current APs, to see more details about the implementation of a RADIUS server on an AP, please refer to the Appendix A. Finally, the UE acts as a holder of a digitally signed identity certificate that must present when attempting roaming into a heterogeneous wireless network.

The deployment of a PKI on top of the current roaming architecture requires the definition of certain roles for each of the key elements. This process is performed as follows:

1. Upon the establishment of a contractual agreement between the HN and the RB. The HN issues a PKC identifying the RB as an Intermediate CA. Thus, the RB can sign PKCs and be recognized by the HN as trusted.

2. The next step is for the HN to assign PKCs to the UEs. These PKCs can be stored on the phone memory or on the SIM card of the mobile device.

3. When building the VN's federation, the RB assigns PKCs to trusted VNs. These PKCs identify the VNs under the RB's domain. Furthermore, in addition to their own certificates, the VNs also have the digital certificates of the HNs with whom the RB has established contractual agreements. Thus, the VN can identify the UE's network operator.

4. The PKC in the UE allows the identification of the mobile device as well as the network operator it depends on.

5. Finally, the VN's PKC allows the HN to identify the RB that federates a specific VN.

The identification process is as follows: a UE attempting to join a VN establishes EAP based authentication process. Here, the UE provides its PKC to the VN. The VN verifies the digital signature with the HN's digital certificate. Upon such verification, the VN can determine if the PKC is valid and the network operator of such UE.

5.3.1 PKC validation

The utilization of a certificate revocation scheme allows the HN to invalidate the PKC in case of lost or exposure of the private key associated with a certificate i.e. *the lost of a mobile device such as a cellular phone containing the PKC*, hence any authentication using that certificate should be denied.

For PKC validation our architecture relies on the OCSP (*Online Certificate Status Protocol*) (Myers et al., 1999) to obtain the revocation status of a PKC. Such protocol requires less network bandwidth and enables real-time status checks for high volume or high value operations. The basic implementation of OCSP in our roaming architecture is the following:

1. The UE and the VN have public key certificates issued by the HN, the CA.
2. the UE requires wireless network access with the VN and sends its PKC.
3. the VN, concerned that UE's private key may have been compromised, creates an 'OCSP request' that contains a fingerprint of the UE's public key and sends it to the HN.
4. HN's OCSP responder looks up the revocation status of the UE's certificate (using the fingerprint the VN created) in its own CA database. If the UE's private key had been compromised, this is the only trusted location at which the fact would be recorded.
5. the HN's OCSP responder confirms that the UE's certificate is still OK, and returns a signed, successful 'OCSP response' to the VN.
6. the VN cryptographically verifies the signed response (it has HN's public key locally stored) and ensures that it was produced recently.
7. the VN grants network access to the UE.

By using certificates, the authentication process can be managed locally. Nevertheless, even though the UE is authenticated and it has gained access to the network, it still does not have the right to perform any service. To do so, it must be authorized and prove that it has the right of using certain services from the VN. This is normally assisted by the user profile, however we consider that the user profile contains confidential information about the user and the HN cannot share it with the VN. Thus, we proposed a PMI-based authorization platform to enable service verification without exposing the user's service profile. This will be explained in details in the remaining of this chapter.

5.4 Privilege Management Infrastructure for Mobile Service Authorization

This section introduces PMI as authorization mechanism, describing how PMI is combined with PKI to provide an efficient and robust authentication and authorization infrastructure. Furthermore, we present in this chapter the PMI-based authorization model and its role in our roaming architecture. Additionally, we also present a key-element of our architecture: the SIP XML-based Attribute Certificates.

Just as the way public key certificates in PKI prove the identity of the entities, in PMI the attribute certificates are used to specify what the entities can do. As a result, they become suitable for access control (*authorization*). An AC has a similar structure as a PKC, however they do not contain the subject's public-key, instead they contain the attributes (*privileges or rights*) of the holder. Similar to the PKC an AC binds the attributes such as group membership, roles, or other authorization information associated with the AC holder to that entity through the signature of a so-called AA (*Attribute Authority*). By definition (Chadwick, 2005) the issuer digitally signs the attribute certificate to assure the integrity. The main differences between PKI and PMI are illustrated in Table 5.2.

Table 5.2: A Comparison of PKIs with PMIs

Concept	PKI Entity	PMI Entity
Certificate	Public Key Certificate	Attribute Certificate
Certificate Issuer	Certification Authority	Attribute Authority
Certificate User	Subject	Holder
Certificate binding	Subject's Name to Public Key	Holder's Name to Attribute(*s*)
Revocation	Certificate Revocation List	Attribute CRL
Root of Trust	Root Certification Authority	Source of Authority
Subordinate Authority	Subordinate Certification Authority	Attribute Authority

5.4.1 PMI-based Authorization Model

The basic privilege management model, as defined by the X.509 attribute certificate framework (Chadwick, 2005), is integrated by the following elements: the SOA (*Source of Authority*), the privilege holder and the privilege verifier, as illustrated in Figure 5.7. In this model the SOA defines and assigns privileges and stores them in a digitally signed policy attribute certificate. The privilege holder is the entity that holds a particular privilege and asserts its privileges for a particular context of use. Thus, the privilege verifier determines whether or not asserted privileges are sufficient for the given context of use.

Privilege delegation is an optional aspect of the PMI framework. There are four components integrating the delegation model: the SOA, the intermediate

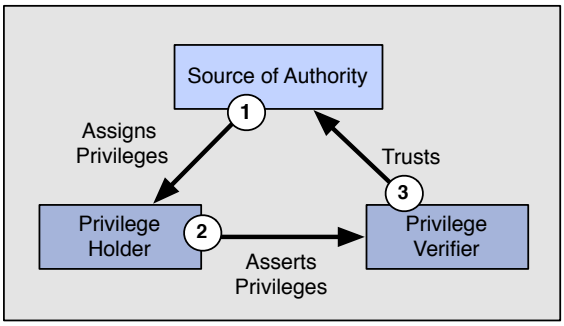


Figure 5.7: Authorization Model

AA, the privilege verifier, and the privilege holder, as depicted in Figure 5.8. The privilege verifier trusts the SOA as the authority given set of privileges for the resource. If the privilege holder's certificate is not issued by that SOA, then the privilege verifier must locate a *delegation path* of certificates from that of the privilege holder to one issued by the SOA. The validation of that delegation path must include verifying that each AA had sufficient privileges and was authorized to delegate those privileges.

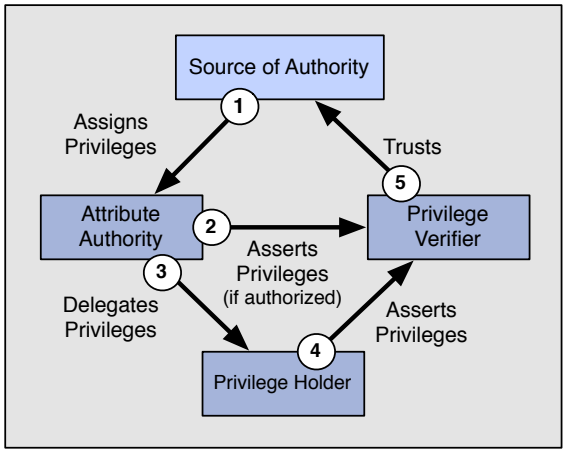


Figure 5.8: Authorization Model

5.4.2 Revocation of Attribute Certificates

The Attribute Certificates can be revoked, based on the chosen policy, in two ways: implicitly, by assigning short validity periods, or in the same way than in PKI, by using Certificate Revocation Lists, which are issued periodically according to the local policy, and contain the serial numbers of revoked certificates. In this context, within the AC data structure, there are two revocation extensions: the CRL distribution points extension and the *no revocation* extension.

The CRL distribution points extension indicates where the revocation list or lists for an specific AC will be found. This allows sets of AC to have their revocation information posted to different revocation lists. This mean, ensures that revocation lists do not become too large. The distribution of CRL information can be performed in two ways: by certificate serial number and by revocation reason. The former post in CRLs blocks of revoked certificates i.e. *serial numbers 1 to 500, 501 to 1000 etc.*, and each AC in a block will contain the same distribution point extension. In the latter, by revocation reason, the revoked certificates are posted to different CLRs according to their reason for revocation, and all ACs will hold the same set of distribution points.

The *no revocation* extensions tells the privilege verifier that this attribute certificate will never appear on an attribute CRL. This extension may be useful for short-lived ACs that will never be revoked.

5.4.3 Attribute Certificates

Under certain circumstances, the binding that requires a verification from an AA is not the binding of the name to a public key, but rather a binding of an identity to a set of attributes. This kind of certification is quite useful in distributed access control, where the AA issues these bindings and specifies the rights of the subject to specific resources. Attribute certificates ease the overall management burden needed for access control hence eliminating the need for specific Access Control Lists (*ACL*). This condition alleviates the management of ACLs in distributed environments where the objects themselves are responsible for the access control of their own resources. When the ACLs are largely spread around numerous objects, the task of maintaining consistency according to the chosen policy becomes difficult. When using such certificates, the object only has to verify the certificate given to the subject, and verify if the identity claimed in the AC is really the identity of the subject attempting to access the resource.

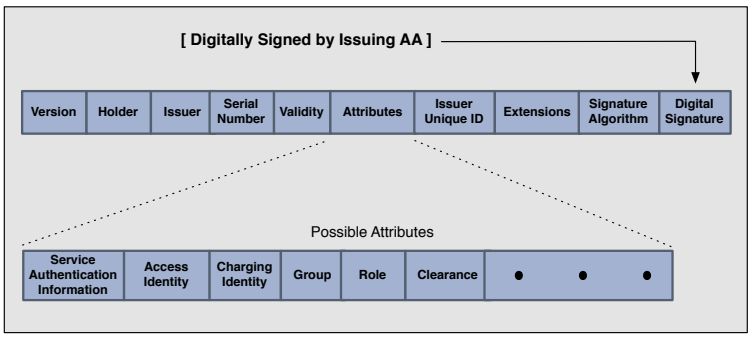


Figure 5.9: Attribute Certificate

An attribute is defined as the information that describes the qualifications and authorities granted to the target entity. By linking attributes to an ID, the qualifications and authorities of the person or entity in question are expressed. Furthermore, attributes also are information shared by target entities that have the

same authorities but generally do not have identifying qualities. As illustrated in Figure 5.9, in addition to the digital signature an AC contains the following fields:

- **Version number:** indicates the version (*1 or 2)*) of the AC format in use.
- **Holder:** is used to bind an attribute certificate to an X.509 public-key certificate. The Holder field identifies the client with which the attributes are being associated. Identification can be either by name or by reference to an X.509 public-key certificate. This field is a sequence allowing three different syntaxes: *baseCertID*, *EntityName* and *objectDigestInfo*. Only one option should be present. For any environment where the AC is passed in an authenticated message or session and where the authentication is based on the use of X.509 public-key certificate, the holder field should use the *baseCertificateID*. With the *baseCertificateID* option, the holder's PKC serial number and issuer must be identical to the AC's holder field.
- **Issuer:** identifies the AA that issued the AC.
- **Serial number:** a unique integer assigned by the user to identify the AC.
- **Validity period:** the time period during which the AC is assumed to be valid (*specified by a pair of time values: a start time and an expiration time*).
- **Attributes:** this field contains information concerning the rights or privileges of the AC Holder.
- **Issuer Unique Identifier:** is used to make the name of the issuing AA unambiguous, in the case where the same name was reassigned to different authorities through time.
- **Extensions:** allows the addition of new fields to the AC. The extensions defined for ACs provide methods for associating additional attributes with holders. This profile also allows communities to define private extensions to carry information unique to those communities.
- **Signature Algorithm:** specifies the cryptographic algorithm used to sign the AC.
- **Signature Value:** the signature value as the X.509 PKC.

The Attributes field of the AC can contain any data. Nevertheless, the standard types of attributes, as depicted by Figure 5.9, are the following :

1. Service Authentication Information: is used for providing legacy applications with needed credentials. Access identity can be used to identify the AC holder to the AC verifier or the larger system of which the AC verifier is a component. These attributes usually contains sensitive information and are thus encrypted.
2. Charging Identity: allows the identification of the AC holder for charging purposes. For example, the holder's company i.e. *the HN* may be the charging identity.

3. Group and Role: are used to present information of AC holder's role or group memberships.
4. Clearance: represents the clearance level of the AC holder, associated with security labeling.

Finally, the Extensions field in the AC can be used to further restrict the applicability of the AC into targets, and to specify distribution points for CRLs, or to declare that no CRL for the certificate is available. Within a heterogeneous roaming context, these field can contain the UE roaming profile, which is not the user service profile. This is explained in the remaining of this chapter.

5.5 SIP-embedded XML Attribute Certificates

Until now we have described PMI and ACs, however we have not specified yet the format of such certificates and how they are transported among the key elements of our roaming architecture. In this context, we present in this section a new approach of ACs: SIP-embedded XML (*Extensible Markup Language*) Attribute Certificates (*SIPXACs*).

The goal of SIPXACs is to transport the user credentials and a simplified version of the user profile within SIP messages so the mobile user can be authenticated and authorized. As stated previously, ACs are particularly well suited to control access to system resources and to implement access control. One of the advantages of ACs is that they are generally short-term, thus allowing for changes in role or security policy. Moreover, it has been shown (Chadwick & Otenko, 2002) that XML standard is particularly useful for the creation of ACs. As a XML-based AC is a string of plain text delimited by markups that defines the privileges of the holder it can be easily incorporated into a SIP message.

To enhance the network registration and session initiation process in terms of efficiency and security, we propose the addition of XML Attribute Certificates (*XAC*) within the SIP-REGISTER and SIP-INVITE messages. To achieve this, we enhanced our proposed Roaming-SIP protocol by adding the support of XML Attribute Certificates (*XAC*). Particularly, for our enhanced roaming architecture we propose the use of XACs in the UE (*UE SIPXAC*) and the Visiting Network (*VN SIPXAC*) and SIP as transport protocol of such certificates. The format of both certificates within a SIP-REGISTER message is illustrated in Figure 5.10(a) and Figure 5.10(b).

A SIPXAC defines the attributes and capabilities of the UE and the VNs. The UE SIPXAC is used to define the privileges of each UE based on their user profile i.e. *the right to roam and register, or the right to perform certain services from the VN*. In contrast, the VN SIPXAC allows the VN to prove that it is authorized to become an extension of the HN and interact with the HN in certain process i.e. *location database update*. If we look carefully Figure 5.10(a) and Figure 5.10(b) we can imply that the structure of both certificates is quite similar, the main and the only difference is the information contained in the certificate. The structure of XACs and the information contained in these fields are explained in detail in the remaining of this section.

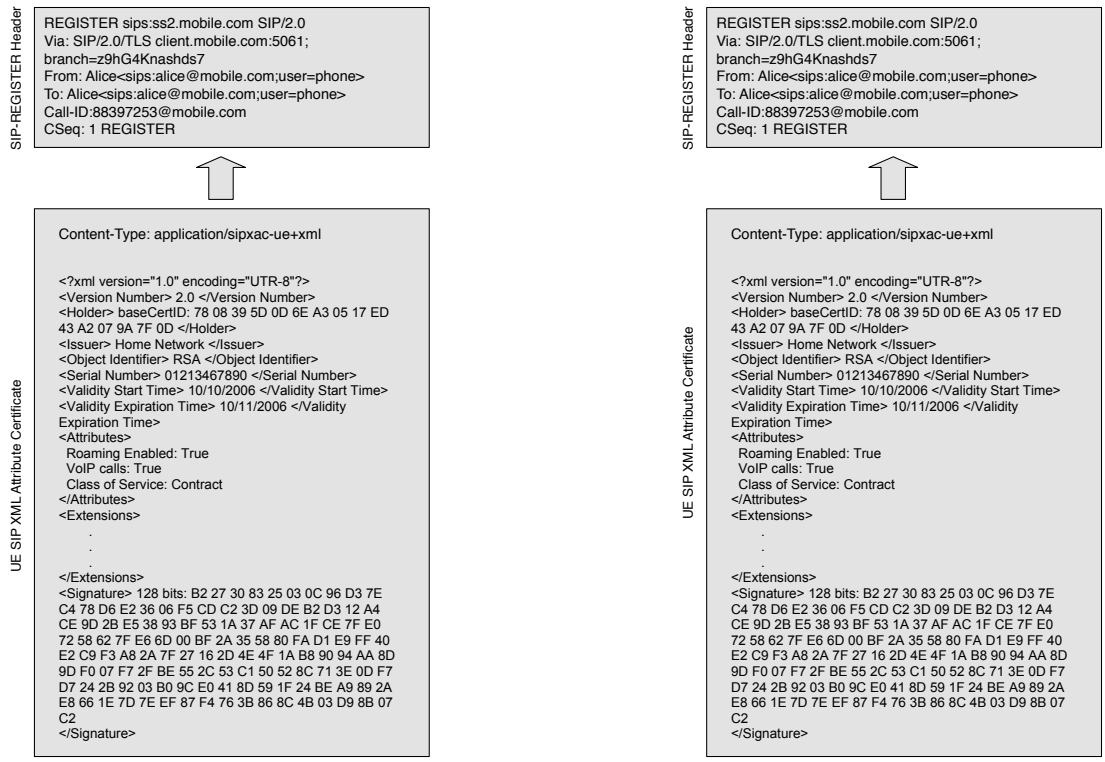


Figure 5.10: SIP-XML Attribute Certificates

5.5.1 User Equipment XAC

The UE XACs provide a mean to validate the user roaming profile hence determining the actions or services the UE is allowed to perform once in the VN. The UE SIPXAC is composed by the following information.

- **Content-Type**: this SDP content allows us to identify the nature of the certificate. It can take two values: UE SIPXAC and VN SIPXAC depending on the entity holding the certificate.
- **Holder**: in our architecture this field is used to bind the UE SIPXAC to its X.509 PKC (*issued by the HN*).
- **Issuer**: the information in this field identifies the authority that issued this certificate. In the case of UE SIPXAC the issuer is always the HN. The issuer can be identified by the MCC (*Mobile Country Code*) and MNC (*Mobile Network Code*) as specified in the ITU-T Rec E.212 (E.212, 2004-2005).
- **Object identifier**: This field identifies the algorithm used to sign the AC.
- **Serial number**: This field contains a unique AC serial number. That is, a unique integer assigned by the SOA (*HN*) to identify the holder in case of name duplication.
- **Validity Start/Expiration Time**: these values are chosen based on the type of contract the mobile user has with the network operator. For example, if the mobile user has a monthly contract, the SIPXAC can be valid only for one month and renovated upon the payment of the next month. Thus, the mobile user will be authorized, only if the monthly fees have been already paid.
- **Attributes**: The information contained under this field defines the user roaming profile. Here the HN specifies what the UE is allowed to do once in the VN. For example, if it is able to roam into VNs, if it is able to perform VoIP calls or initiate streaming video sessions. Additionally, it also contains the class of service that specifies the type of contract the user has with their network operator i.e *Monthly-contract or pre-paid service*. This attribute is useful for network operators to identify the user's roaming profile when attempting to roam into VNs.
- **Extensions**: the extensions field is used to provide extra information about user's privileges such as time control. The time extension constrains the time when the privilege can be exercised e.g. *certain days of the week and certain times of the day*. This can be useful for network operators if they want to constrain services to be provided only during certain period of the day. We can also use this field to provide additional information such as UE's IP Address, MAC Address, IMEI (*International Mobile Equipment Identity*) or IMSI (*International Mobile Subscriber Identity*)

5.5.2 Visiting Network XAC

The VN SIPXACs enables the VN to interact directly with the HN without the need of proxies (*the RB*). When the RB issues a XAC to a VN, it implicitly authorizes this to query the HN as it was the Roaming Broker itself. The VN SIPXAC contains the same fields than the UE SIPXAC however there are certain differences in the information contained, as described below.

- **Content-Type**: in the case of an attribute certificate on the VN it takes this value: VN SIPXAC.
- **Holder** : as in the UE SIPXAC, this field is used to bind the VN SIPXAC to its X.509 PKC (*issued by the RB*).
- **Issuer**: in the case of VN SIPXAC the issuer is the RB federating such VN.
- **Validity Start/Expiration Time**: the validity period of this certificate is set by the RB. Once the expiration time is reached and after a VN's reputation-based evaluation the RB decide whether or not to renew the SIPXAC.
- **Attributes**: the RB specifies the interaction level between the VN and the HN. It also defines the query types the VN can perform to the HN i.e. *location database update*. In this field, the RB also defines the class of service based on the wireless access network *i.e. WiFi, WiMAX*. This attribute is useful for network operators to know more details about the wireless network their mobile users are attempting to roam into.
- **Extensions**: the extensions field are used to provide extra information about the VN's privileges. The information in this field is defined by the RB.

5.6 PMI-Enhanced Roaming Architecture

Attribute certificates determine the roles and attributes of the key elements of our architecture. In this section, we present a PMI-based integration architecture that binds trust relation between heterogeneous operators. To achieve this, we rely on PKI & PMI mechanisms to guarantee efficient network security and SIP as roaming signaling protocol. It is worth-mentioning that this trust infrastructure is built on top of the SIP-based Roaming architecture presented in chapter 4, see Figure 5.11. In this perspective, we introduce in this section the necessary enhancement to the prior architecture to enable fast & secure service mobility between the HN and independent UWNs.

The aim of PMI and SIPXAC in the roaming architecture is to enhance the network registration and session initiation process. We aims at reducing the signaling message exchange while enhancing the robustness of the network security.

The role of PMI and SIPXAC in each element of the SIP-based roaming architecture, as well, as the integration of such technologies in the architecture is explained in details in the remaining of this section.

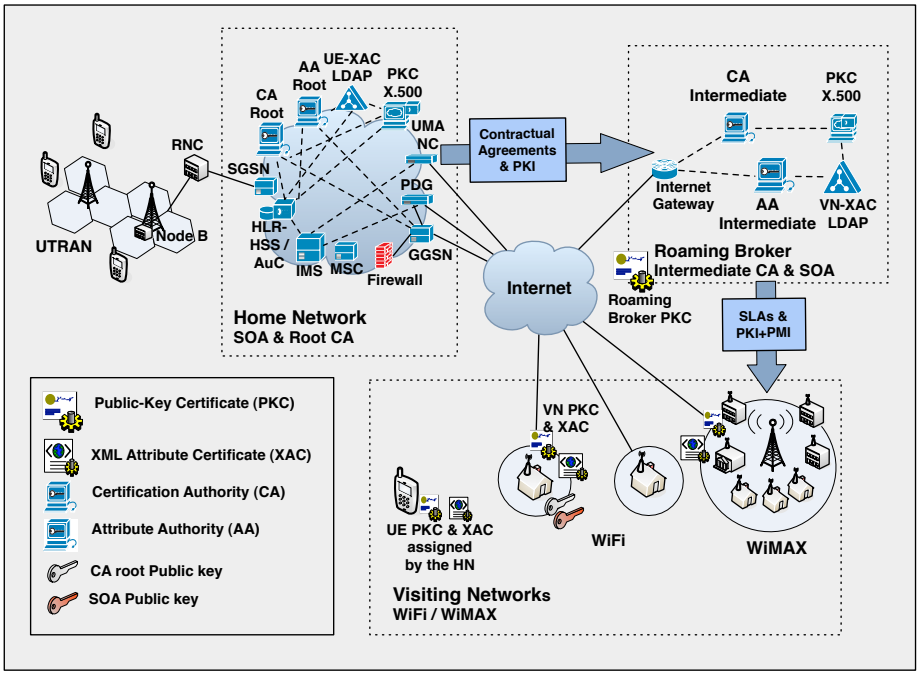


Figure 5.11: SIP & PMI-based Roaming Architecture

5.6.1 PMI & SIPXAC integration

The certification path in our proposed roaming architecture comprises the HN, the RB, the VN, and the UE, as illustrated in Figure 5.12. In the roaming architecture, the SOA (*the HN*) delegates privileges to the RB (*flow 1*) and assigns the UE's privileges, based on the user service profile (*flow 2*). Additionally ,The RB plays the role of intermediate AA hence assigning privileges to the VN. Among these privileges is the capability of offering roaming services (*flow 3*). The UE is the privilege holder, this element carries the attribute certificate that grants him access to the VN and specific services (*flow 4*). Finally, The VN acts as a privilege verifier, hence it decides whether or not the UE is authorized to roam or initiate services in the VN (*flow 5*).

The elements participating in the PMI-based Roaming Architecture are explained in details in the remaining of this section.

5.6.1.1 The Home Network

The HN is considered the SOA in the PMI and the root CA in the PKI. It is the trusted entity with ultimate responsibility for assignment of a set of privileges. The HN acts as an Attribute Authority (*AA*) that issues ACs to the UEs based on their user profile. Thus, in the HN find:

- The CA root
- The AA root *(SOA)*

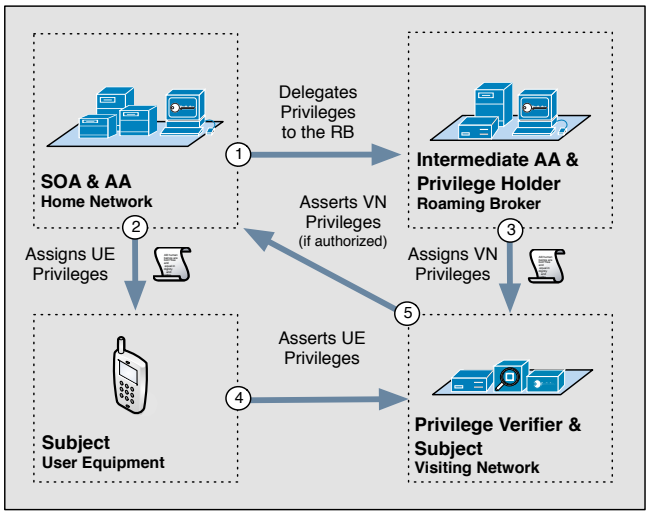


Figure 5.12: Attribute Certification Path

- A PKC Repository (*X.500 Directory*) containing the PKC of each UE (*User Equipment*) in the network and the PKC of the RBs.
- An LDAP (*Lightweight Directory Access Protocol*) repository with the UE's AC.
- An up-to-date revocation database for both PKCs and ACs.
- The privilege policies for the UEs (*roaming profile*).

Figure 5.12 depicts how the attribute certification path used in our architecture supports privilege delegation. Moreover, the SOA assign the privilege to the RB to act as an intermediate AA with the capability of issuing ACs to the VNs. The privilege delegation path in our architecture is defined in the following manner: $HN \rightarrow RB \rightarrow VN$

5.6.1.2 The Roaming Broker

As stated previously, the RB is an intermediate AA however it is also a privilege holder as it is authorized by the HN to perform roaming brokering services and assign the VN's privileges. To do this, the RB relies on the following elements:

- A PKC that proves the RB identity.
- An intermediate CA that allows the RB to issue PKC's to the VNs
- An intermediate AA that allows to issue ACs to the VNs.
- A PKC X.500 repository containing the PKCS of all the VN's under the RB's domain.
- A LDAP repository containing the VN's ACs.

- An up-to-date revocation database for both VN's PKCs and ACs. This database is accessed by the HN to validate the VN's certificate.
- The privilege policies for the VNs. Here, the RB defines the rights and role of the VN under her domain.

Th RB in the role of intermediate root CA and intermediate AA issue PKCs and delegates privileges to the VNs. Thus,the RB authorizes the VN to play the role of privilege verifier (*assigned by the RB*), see Figure 5.12 and at the same time provide roaming services.

5.6.1.3 The Visiting Network

The RB through an AC grants the VN the role of privilege verifier, this allows the VN to make an admit/reject decision based on the privilege policy, the privileges of the UEs and the revocation list provided by the holder (*UE*). To perform these tasks efficiently the VN relies on the following resources:

- A PKC that identifies the VN as part of the RB domain.
- An AC that authorizes the UWN to become a VN.
- The public key of the trusted CA(s) (*the HN and the RB*) so it can verify the signatures on the ACs and the PCKs that it will evaluate.
- The name and public key of a trusted SOA (*the RB*) that can be validated against the RB's PKC, so the VN can validate that the ACs are issued directly or indirectly by this SOA.
- The privilege policy that defines how the VN can determine if the credentials presented by the UE are sufficient to access the resource. This policy is related to the SLA established between the RB and the UWNs.
- The ACs of the holder and a valid chain back to the SOA.
- An up-to-date revocation database. In our architecture this information is provided by the holder (*UE*) and confirmed by the OCSP.

Since the data is digitally signed it cannot be altered without detection hence it does not need to be configured by trusted means.

5.6.1.4 The User Equipment

Within our architecture the UE plays the role of the privilege asserter (*AC holder*) invoking some action/service request on a specific resource. In order to roam into a heterogeneous network or initiate a service from within, the UE relies on:

- A valid PKC issued by the root CA that proves its identity (*authentication*).
- A valid AC issued by the SOA describing the user profile (*authorization*).

5.7 Fast & Secure Roaming

In most of the cases, before accessing a resource we are required to carry out two mechanisms: authentication and authorization. In chapter 4, the Roaming-SIP protocol provides the user's credentials within the SIP-REGISTER message upon the reception of a SIP-407: Proxy Authentication Request packet as illustrated in Figure 5.13 (*flow 2*). The limitation of this approach lies on the the additional signaling message exchange introduced by the triangulation between the VN, the RB and the HN for network registration and session initiation (*flows 4 to 9*). Nevertheless, this configuration is necessary as the VNs is not allowed to communicate directly to the HN. As stated in previous chapter the RB acts as an authorized proxy to establish SIP dialogs with the HN on behalf the VNs, as illustrated in Figure 5.13. The signaling message exchange starts when the mobile user initiates the registration process in the VN or when the UE attempts to initiate a VoIP call. The other inconvenient of this approach is the Internet delay introduced by the *remote* authentication and authorization.

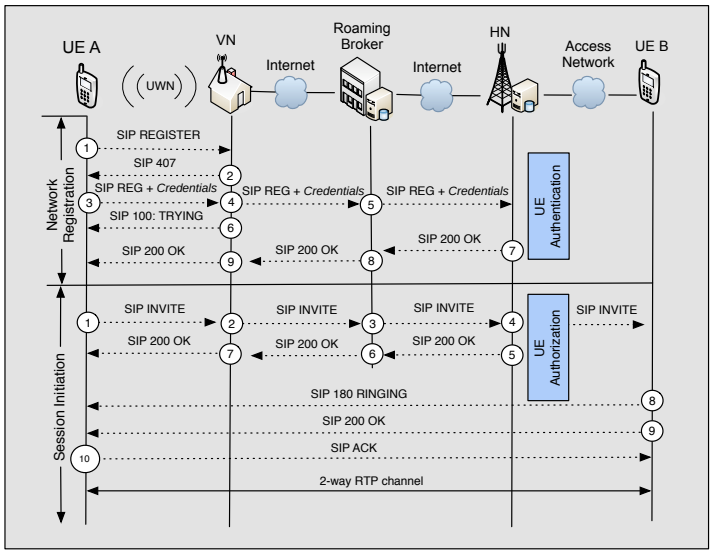


Figure 5.13: Roaming-SIP Signaling Exchange

In the SIPXAC approach the network registration and session initiation signaling message exchange is reduced considerably. The process, as illustrated by Figure 5.14 is the following:

1. Upon entering into the VN coverage area the UE attempts to register. To do this, the UE transmits a SIP-REGISTER. The UE's registration message contains the SIPXAC of the UE providing the right to roam into an heterogeneous wireless network.

2. As the XAC are linked to the PKC, the VN can authenticate and authorize the mobile user *locally*, without requesting authorization from the RB or the

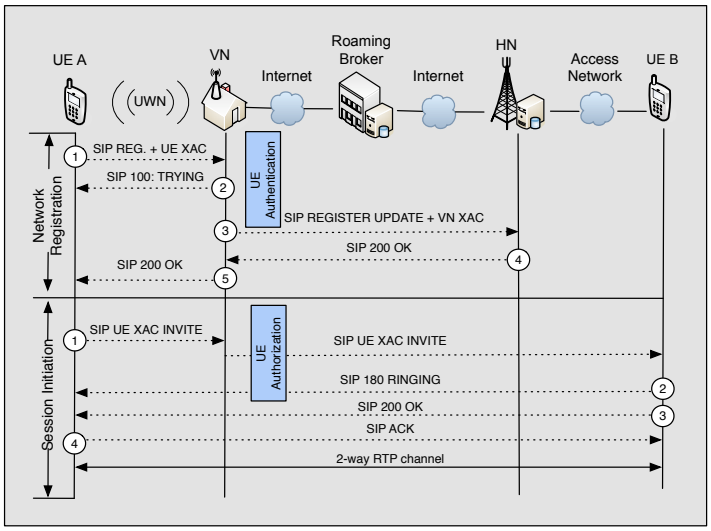


Figure 5.14: SIPXACs Signaling Exchange

HN. Thus, based on the privilege policy and the SLA information, the privileges in the UE SIPXAC, and the revocation list, the VN determines whether or not let the user register in the network. If the UE's privileges are sufficient to let the user in, the next step of the VN is to update the user location in the VN.

3. To do this, the VN uses its own VN SIPXAC to obtain direct access to the HN's location databases. Thus, the VN can update the user location database in the HN without passing through the RB.

4. Once the HN has an up-to-date information about the UE's location it transmits a SIP-200 OK message to the UE.

5. The reception of the SIP-200 OK message by the UE indicates a successful network registration.

On the other hand, as depicted in the Session Initiation section in Figure 5.14 session initiation signaling starts when the UE transmits an invitation message when attempting to initiate a multimedia service such as a VoIP call.

1. The UE includes in the SIP-INVITE message the UE SIPXAC. If the UE SIPXAC authorizes the UE to perform VoIP calls from VNs, then the VN forwards the invitation message to the other peer. In this way, only services defined in the roaming profile are authorized, hence the HN has more control of the mobile user when roaming into VNs.

2. The other peer starts the required signaling to accept the session invitation and start the VoIP conference. Here, the peer transmits a SIP-180 RINGING message.

3. Followed by a SIP-200 OK message indicating the signaling exchange was successful.

4. Finally, the originating peer acknowledges the reception of the 200 OK messages. At this point both peers are ready to initiate the multimedia session, in this case a VoIP call.

5.8 Roaming Signaling Delay Analysis

The enhanced roaming architecture was evaluated through computer simulation. In this section, we present some results that demonstrate the improvements of this new contributions compared to the Roaming-SIP approach proposed in previous chapter.

5.8.1 Simulation Objective

Since SIP is an application layer protocol, the processing of SIP messages in the intermediate and destination servers/entities may experience certain delay due to queueing of messages that need to be served. Nevertheless, we consider that due to the increasing computing power in network devices the processing delay introduced by the processing of SIP messages and the cryptographic algorithms (*used for digital signature validation*) does not a have considerable impact on our architecture. Due to the geographical distribution of entities in the architecture, we consider that major delays are introduced by the Internet delay, as consequence of the signaling message exchange between the VN, the RB and the HN. Thus, for analysis we consider two important mechanisms in the roaming process: network registration and session initiation. The parameter to evaluate is network registration and session initiation delay. The rationale is that delay impacts directly on the roaming process, the fastest the UE updates its location the fastest it can get service.

5.8.2 Simulation Model

Figure 5.15 illustrates the network model used in the simulations of the roaming architecture. As we can observe, for this analysis we relied on the same network model used in the previous chapter, however we made certain modifications. The first modification implied the interconnection between the RB and the HN. As we rely on a trust infrastructure supported by PKI and PMI, the roaming signaling messages do not have to pass through the RB anymore. The required message exchange is performed directly between the UE, the VN and the HN.

In addition, as illustrated in Figure 5.16, the wireless node were also modified to support the new parameters introduced by the XML ACs such as packet size (*increased due to the embedded certificate*) and SIPXAC headers. Moreover, the delay introduced by the cryptographic algorithms when signing and verifying the validity of the certificate was also taken into account. We obtained some of these

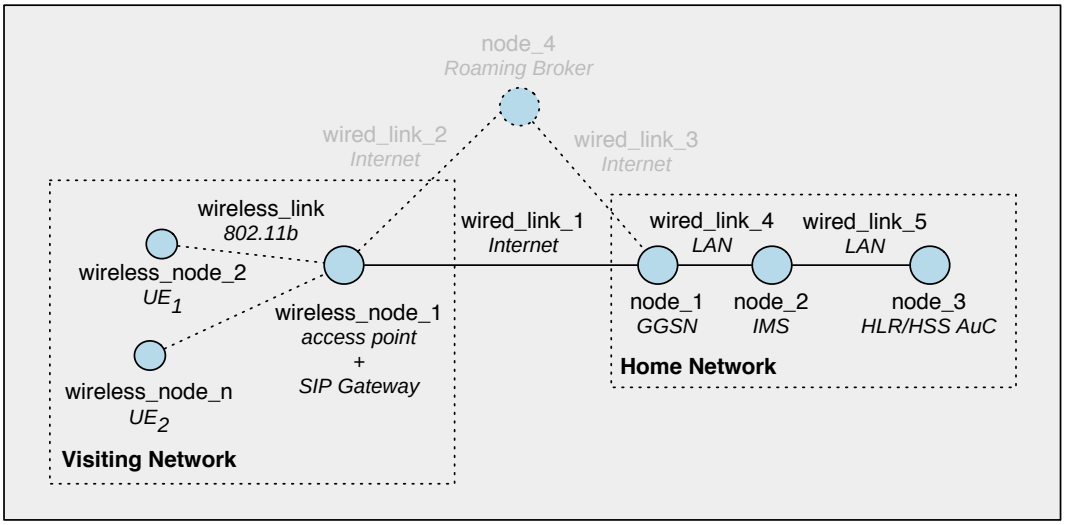


Figure 5.15: Simulation Model

values from Shafer's work (Hess & Schafer, 2002) and through the Openssl command in the access point. From the command line of the WRT54G access point, we typed:

```
wrtg54g-osalazar:root$ openssl speed
```

This command triggers the algorithm speed measurement and provides the following outcome:

```
                  sign    verify    sign/s verify/s
rsa  512 bits 0.001895s 0.000135s    527.8   7423.7
rsa 1024 bits 0.008975s 0.000389s    111.4   2568.2
rsa 2048 bits 0.051038s 0.001278s     19.6    782.2
rsa 4096 bits 0.315193s 0.004456s      3.2    224.4
                  sign    verify    sign/s verify/s
dsa  512 bits 0.001182s 0.001424s    845.8    702.2
dsa 1024 bits 0.003654s 0.004446s    273.7    224.9
dsa 2048 bits 0.012350s 0.015120s     81.0     66.1
```

Then, we used these values as input in our simulation model to recreate the processing delay introduced by cryptographic algorithms.

5.8.3 Simulation Parameters

Similar to previous chapter, the simulation was performed on *ns2* (NS-2, 1995) using an enhanced version of Rui Prior's SIP module for *ns2* (Prior, 2006). Our contributions to this module include: the packet length when using Roaming-SIP and SIPXACs, and the processing delay introduced by the cryptographic algorithms.

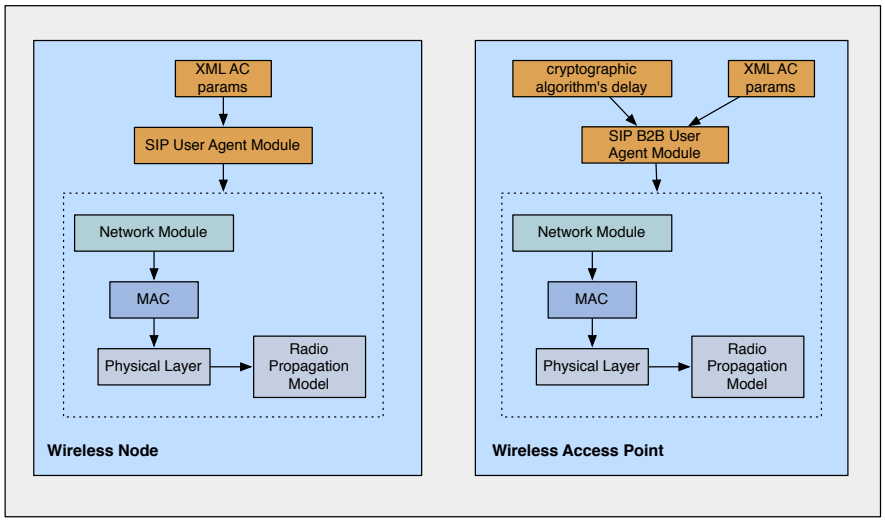


Figure 5.16: Wireless Node Model

Most of the parameters used in the simulation, as depicted by Table 5.3, are the same than the parameters used in the simulation in previous chapter. The rationale of the assumed parameters value is, for comparison purposes, to be consistent with the simulation environment used in our previous simulations (*Chapter 4*). Thus, we assumed for the VN, 802.11 MAC layer providing a theoretical data rate of 11 Mbps. In terms of transmission delay, we assumed $10ms$ in the links connecting the GGSN, the IMS and the HLR-HSS/AuC in the HN (Δ_{HN}), as we consider that it is an acceptable delay for local area networks. Similar to previous chapter we assumed an Internet delay of $150ms$ (Δ_i) and ADSL links connecting the VNs with a data rate of $2Mbps$. The Internet link connecting the RB to the HN was assumed to be $20Mbps$.

On the other hand, we used different SIP packet size due to the utilization of the XML ACs embedded in the SIP messages. In this context, for our simulation we used two different packet sizes: 587 bytes for Roaming-SIP and 1300 bytes for SIPXAC. The processing delay in both cases was defined by the simulator accordingly to the packet size and the signature validation algorithm, as described in section 5.8.2. The GSM vocoder was chosen as it provided acceptable performance in VoIP communications and it is supported by a considerable number of cellulars in the market.

5.8.4 Simulation Outputs

As the parameter to evaluate is network registration and session initiation delay, we were interested on studying the different components of the delay in the network. Thus, we considered that the delay components in our architecture in a single message exchange are: the wireless delay, the processing delay (*time that the VN spends processing Roaming-SIP and SIPXAC messages*) in the VN, the Internet delay, and in the HN processing delay, see Equation 5.1.

Table 5.3: Simulation Parameters

Parameters	Values
Wireless MAC	802.11
Δ_{HN}	10 ms
Δ_i	150 ms
VN Internet bandwidth	$2Mbps$
RB Internet bandwidth	$20Mbps$
SIP_{AAA} Pkt Length	587 bytes
SIP_{XAC} Pkt Length	1300 bytes
Number of VoIP sources	variable 1,5,10,15
VoIP codec	GSM - 13.3 kbps
Simulation time	600 seconds

$$D_{total} = D_{Wtx} + D_{proc} + \Delta_i + D_{HN} \tag{5.1}$$

Where, D_{total} expresses the overall delay for a successful network registration or session initiation attempt, D_{Wtx} describes the average wireless transmission delay in the VN, D_{proc} refers to the average processing delay in the VN, Δ_i is the Internet delay, and finally D_{HN} which represents the processing delay in the HN.

The Internet delay is a parameter that neither the VN nor the HN can control, in our simulation environment this is fixed to $150ms$. On the other hand, to evaluate the impact of congestion in the wireless LAN, we decided to recreate four congestion scenarios. To do this, we set a variable number (*1, 5, 10 and 15*) of wireless nodes continuously transmitting VoIP traffic to the wireless network. We defined a maximum number of 15 sources because it is proved that 802.11b cannot support more than 15 simultaneous VoIP sources (Hole & Tobagi, 2004) (Coupechoux et al., 2004).

We consider that the HN has all the computing and network capabilities to support multiple users therefore hence for evaluation purposes, we decided to observe the behavior of a 802.11b VN under the presence of multiple UEs as the access points (*VNs*) are devices with limited resources. To support this assumption, we observed that the processing delay in the HN was so slow that it could be easily neglected. Notwithstanding, the processing and transmission delay in the VN exhibited different performance.

In this context, we present in this section, the results for the network registration and session initiation delay for both Roaming-SIP and SIPXAC using the simulation environment described in section 5.8.3. Similar to chapter 4, the results presented are the average of twenty independent replications of the terminating simulation of duration 600 seconds. The independence of replication was accomplished by using different random number seeds for each simulation. The validation and verification of the simulation outputs were performed following the same methodology as in chapter 4. The confidence interval of the simulation outputs presented in this paper are 95%.

5.9 Simulation Results

As mentioned previously, due to the varying nature of the wireless channel, it was interesting to evaluate the delay components introduced by the VN's wireless environment. In this perspective, we present the results for the average wireless transmission delay, the average processing delay and the overall delay in both network registration and session initiation process.

5.9.1 Average Wireless Transmission Delay

The average wireless transmission delay is defined as the time required to transmit a SIP message through the wireless network. For instance, with an average wireless delay of 1.5 seconds, the transmission and/or reception of three SIP messages through the VN's wireless interface could add up to 4.5 seconds to the overall network registration process.

5.9.1.1 Network Registration

Delay in 802.11 networks under high load conditions is generally caused due to the initial size of the congestion window and the number of back-off stages. Consequently, as depicted by Figure 5.17, the Roaming-SIP and SIPXAC exhibited similar average wireless transmission delay in most of the congestion levels, despite of the difference in SIP packet size.

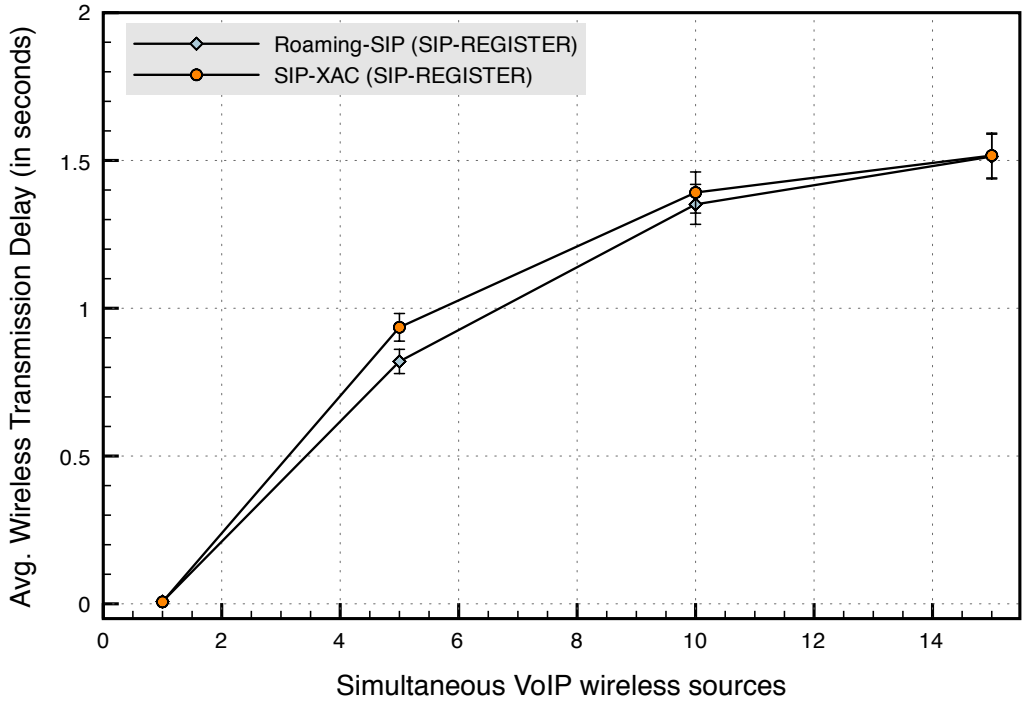


Figure 5.17: Avg. Wireless Transmission Delay for Network Registration

In the case of 5 simultaneous VoIP wireless sources SIPXAC (*SIP-REGISTER message*) displays slightly higher delay compared to Roaming SIP. Nevertheless, the difference is minimal as being in the order of $\approx 3ms$. Under high load congestion levels (*15 sources*) both approaches exhibited the same wireless transmission delay. Based on these results, we can state that the difference in SIP packet size introduced by the XAC does not impact considerably the wireless transmission delay in the network registration process.

5.9.1.2 Session Initiation

Accordingly to the network registration process, we were interested on evaluating if the use of SIPXAC produces a negative impact in terms of wireless transmission under load conditions. Not surprisingly, the results exhibited similar performance than in the case of network registration, see Figure 5.18. Again, the simulation outputs demonstrate that the increment in the SIP-INVITE packet size due to the XAC doest not have a great impact on wireless transmission delay.

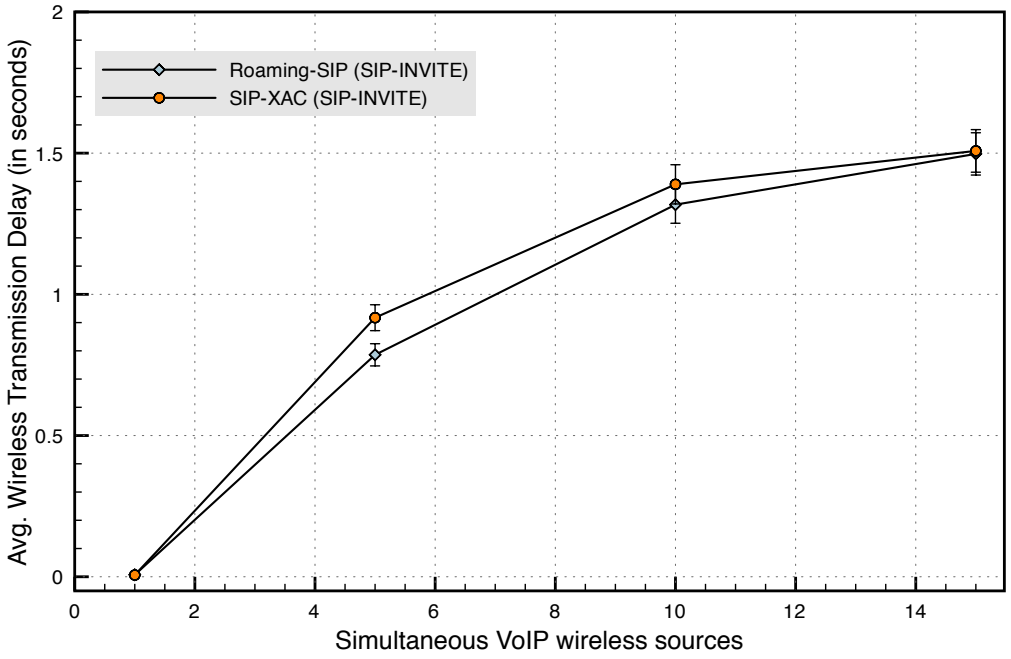


Figure 5.18: Avg. Wireless Transmission Delay for Session Initiation

Consequently, from both network registration and session initiation results we can imply that under high congestion levels and regardless of the packet size of the SIP message, the average transmission delay tends to increase. Thus, efficient admission control mechanisms are required to maintain acceptable load conditions and lower average wireless transmission delay.

5.9.2 Average Processing Delay

As defined previously, processing delay is the time the VN or the HN spend processing the SIP messages. We considered that the HN has the computing resources to maintain processing delay in considerable low levels. However, due to the limited computing resources in current wireless LAN access points, we decided to analyze the impact of SIPXAC on processing in the VN.

5.9.2.1 Network Registration

The results expressed in Figure 5.19 indicate that SIPXAC contributes with higher processing delay. This is due to the increased size of the SIP message and the processing of cryptographic algorithms when verifying the validity of the digital certificate.

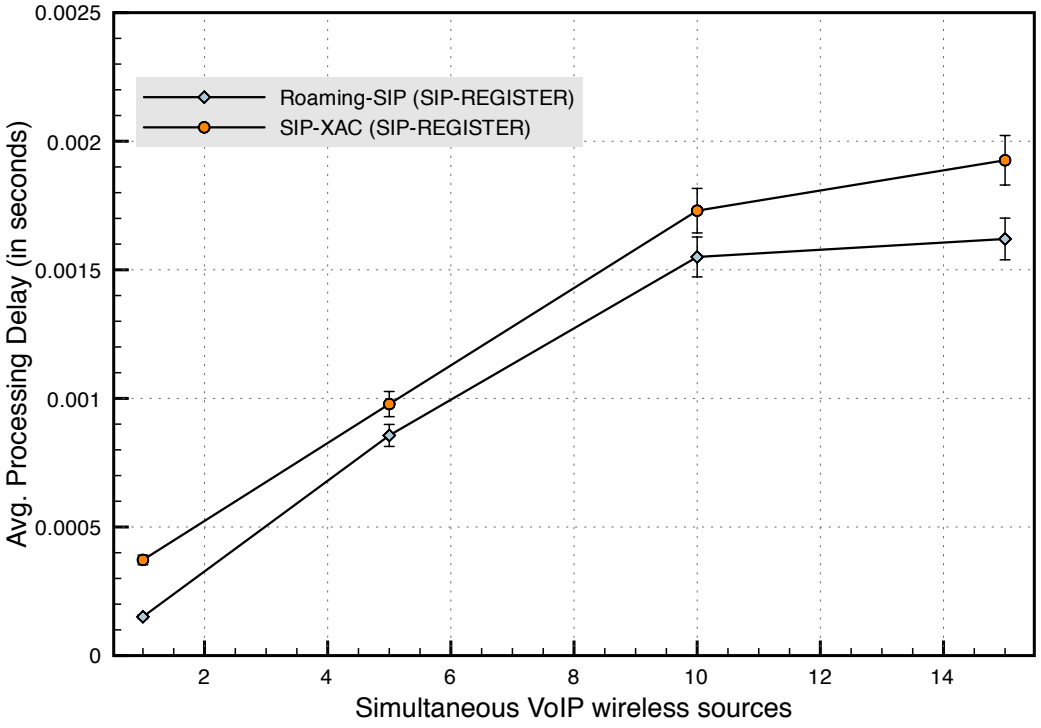


Figure 5.19: Avg. Processing Delay for Network Registration

Furthermore, the processing delay introduced by the SIPXAC approach is less than $3ms$ under high condition levels and $0.5ms$ under low condition levels. From this, we can state that the processing delay introduced by SIPXAC definitely does not represent any harm to the overall delay performance.

5.9.2.2 Session Initiation

From the results in section 5.9.2.1 we can see that the processing delay is definitely impacted by the use of SIPXAC and the cryptographic mechanisms. Hence, resulting in a slight increment of the processing delay. Nevertheless, as depicted by Figure 5.20, the difference is not considerable.

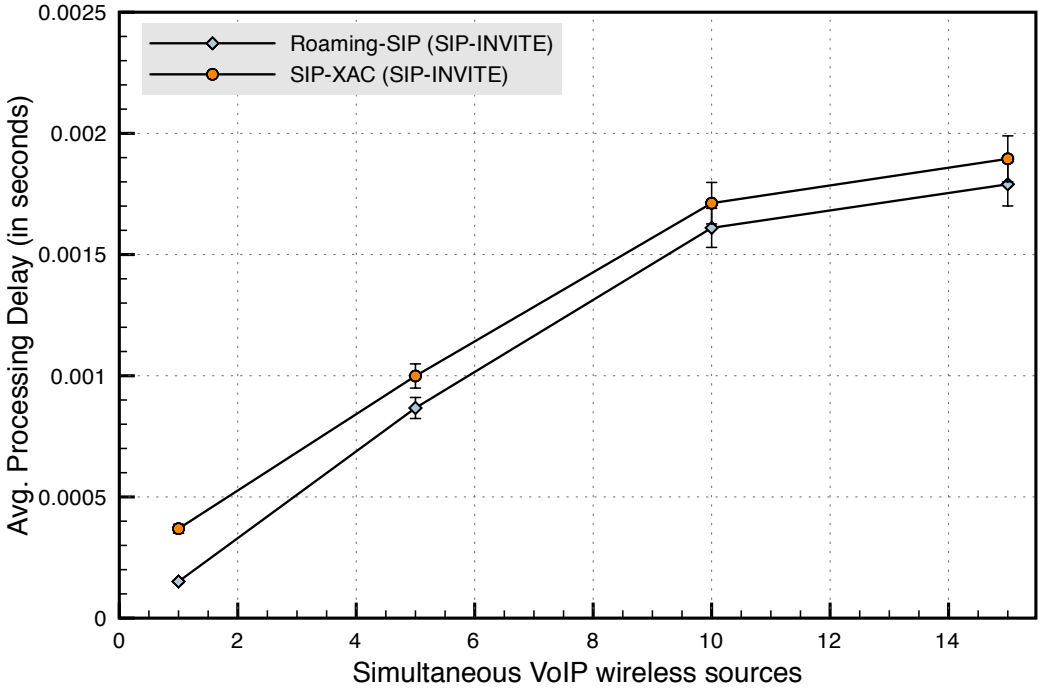


Figure 5.20: Avg. Processing Delay for Session Initiation

As exhibited in Figure 5.20, under all the congestion scenarios, the SIPXAC approach display higher delay. Nevertheless, as this difference is less than $1ms$ in some cases, it can be negligible . Based on the fact that processing capability of APs are improving on daily basis, we can state that even though the use of SIPXAC contributes in increasing the processing delay this increment does not have a negative impact on the overall session initiation delay.

5.9.3 Overall Delay

Overall delay expresses the time experimented by the UE since the transmission of a SIP-REGISTER or SIP-INVITE message until the UE_1 receives the SIP-200 OK message from the other peer (*the HN or the* UE_2). Accordingly, overall delay takes into account all the delay components previously evaluated: average wireless transmission delay, VN processing delay, Internet delay, and HN processing delay.

5.9.4 Overall Network Registration

The overall network registration delay is the time a UE spends in network registration when entering into the coverage of a VN. This parameter is important as low network registration delay might improve certain functionalities in the vertical handover process.

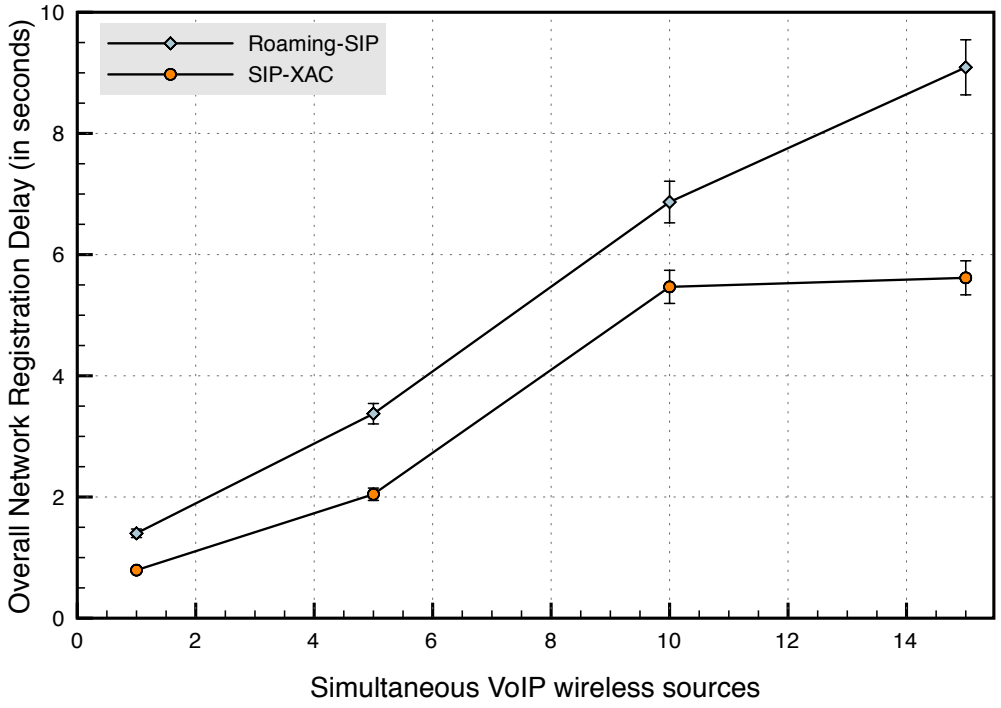


Figure 5.21: Overall Network Registration Delay

The results illustrated in Figure 5.21 show that in the case of low congestion levels (*1 wireless*) VoIP source in the network an UE using SIPXAC required 0.79 seconds to successfully register into the VN, whereas an UE using Roaming-SIP required 1.13 seconds. On the other hand, under higher congestion levels (*15 wireless VoIP sources*) the UE using SIPXAC could register after 5.61 seconds whereas the UE using Roaming-SIP needed 9.98 seconds. From this, we can state that again congestion in the VN definitely has a negative impact on the overall signaling performance.

5.9.5 Overall Session Initiation

Likewise to network registration, for evaluation purposes, we also considered the overall session initiation delay as the sum of all the delays components required to establish a session between two peers.

The simulation results from overall session initiation delay are depicted in Figure 5.22. In the case of session initiation, the use SIPXAC also contributed in reducing the overall session initiation delay. Under a low congestion scenario (*1*

wireless source), the UE using SIPXAC was able to initiate a call in 1.07 seconds whereas the UE using Roaming-SIP required 1.53 seconds.

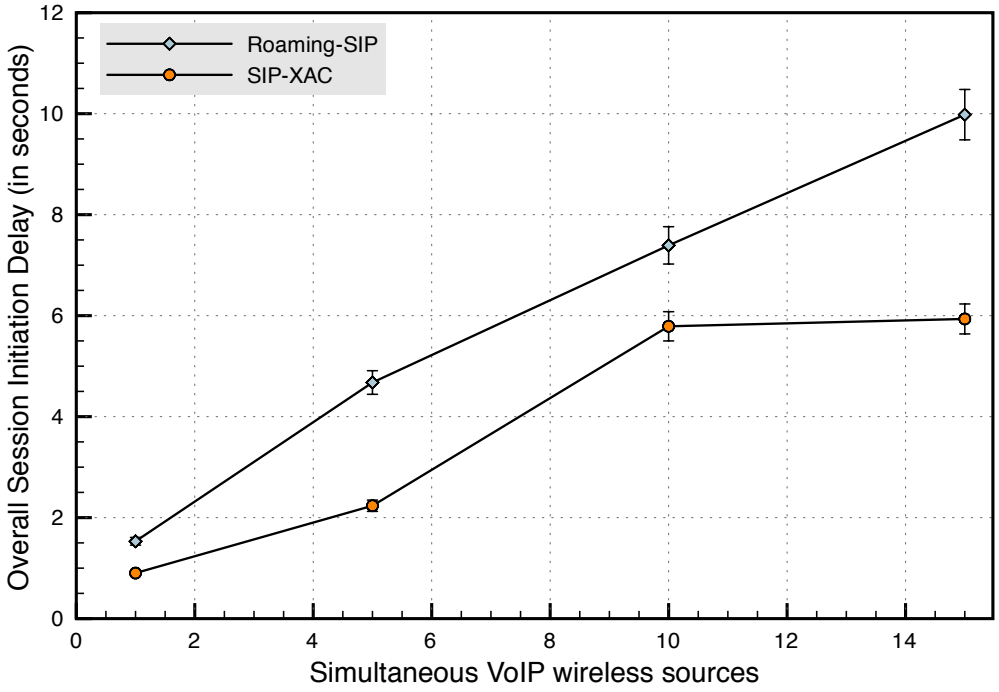


Figure 5.22: Overall Session Initiation Delay

The evaluation of SIPXAC-based session initiation under high congestion level, 15 wireless VoIP sources, also showed acceptable results. The UE using SIPXAC started a VoIP call in 5.95 seconds when the UE using Roaming-SIP required 10.21 seconds. Once again, the simulation results demonstrated that by reducing the number of SIP messages exchanged, see Figure 5.14, we obtained significant improvement in the session initiation process.

From Figures 5.21 and 5.22, we can observe that the session initiation process exhibited a slightly greater delay compared to the network registration process. This is due an extra message required in the session initiation process and the size of the attribute certificate in the SIP-INVITE message, as it describes the user roaming profile.

5.10 Conclusions

In previous chapters we have discussed the issues and challenges of roaming in heterogeneous multi-operator wireless environments. We also proposed in chapter 4 a SIP-based Roaming Architecture to overcome the majority of such issues and provide an efficient and robust roaming platform to enable heterogeneous service migration.

In engineering every proposal has a trade-off, in our case the proposed SIP-based roaming architecture, while performing adequately, introduces certain delay due to the SIP-based signaling messages exchange among the key-players in the architecture: the HN, the UE, the RB, and the VN.

From computer simulation analysis, we noticed that such delay was mainly introduced by the authentication and authorization processes when the UE attempted to register into the network or initiate a session from a VN. Thus, in this chapter we proposed an enhancement of such architecture by reducing the signaling delay in the key-processes: network registration and session initiation, all this while providing a robust authentication and authorization platform. To achieved this, we relied on PKI for authentication and PMI for authorization. Additionally, the SIP-embedded XML Attribute Certificates contributed to the support of the user roaming profile when roaming into independent UWNs. Our enhancements to the SIP-based roaming architecture are the following:

1. the implementation of PKC and PMI-based mechanism to improve the efficiency and security of our proposed SIP-based Roaming Architecture.
2. the de-localization of authentication and authorization processes and,
3. the use of SIP-embedded XML Attribute Certificates (*SIPXAC*) for service authorization in heterogeneous networks.

The results obtained through computer simulation indicated that the use of SIPXAC reduces significantly the network registration and session initiation delay, hence outperforming the traditional Roaming-SIP method. We also confirmed that the wireless delay introduced by the VN increases considerably when increasing the traffic congestion level in the UWN. Thus, from the simulations results we can state that by reducing the signaling message exchange, and maintaining acceptable congestion levels in the wireless network, hence reducing the wireless transmission delay, we can improved the overall delay in both network registration and session initiaition process. Consequently, as future work we are interested on the design of efficient admission control mechanisms to maintain the congestion in the VNs under acceptable levels.

Although, we have improved the key-processes in our architecture and provide and efficient and robust roaming platform for heterogeneous multi-operator environments, we consider there are still certain areas that could benefit from an enhancement. Such areas will be described in the next chapter as possible perspectives for our current research.

Chapter 6

Conclusion and Perspectives

6.1 Thesis Conclusions

This thesis addresses the issue of heterogeneous roaming in multi-operator wireless environments by proposing a SIP-based roaming architecture. The objective of such architecture is to provide a seamless and efficient roaming platform to enable interoperability between cellular networks and independent UWNs. Our main contributions to this domain as presented in chapters 4 and 5 respectively are:

- A SIP-based roaming architecture for heterogeneous multi-operator wireless networks.
- The necessary enhancements to improve our proposed SIP-based roaming signaling protocol while maintaining robust network security.
- The implementation of SIP-embedded XML Attribute Certificate as service authorization mechanism.
- The evaluation of our proposed architecture and enhancements through computer simulation and testbed.

In chapter 4, our roaming architecture enables seamless roaming through the addition of a new architectural element called the Roaming Broker. Thus, through this element, trust relations are built between the cellular and the UWNs. Trust between heterogeneous wireless networks is established in the following manner: trust between the RB and the cellular network is endorsed by contractual roaming agreements whereas the RB and the UWNs rely on SLAs placed on the wireless access points.

The roaming signaling message exchange is performed through an enhanced version of SIP. Consequently, SIP is modified based upon the recommendations of the architecture to convey critical information for the roaming process. Particularly, two SIP messages are modified: the SIP-REGISTER and the SIP-INVITE messages. After our modifications, both messages convey information regarding the mobile user, the UWN, and the SLAs. Thus, the HN (*cellular network*) can determine whether or not to let the mobile user roam in to the VN (*UWN*).

The evaluation of our architecture was performed through computer simulation, in addition a feasibility study was also performed to demonstrate the real applicability of our proposal. In this perspective, computer simulation results displayed that the deployment of our SIP-based roaming architecture slightly increases the delay in the SIP messages processing. Furthermore, due to the roaming-related information conveyed in the SIP messages (*increasing SIP message size*), the signaling overhead also experiences a slight increment. Nevertheless, as indicated by the simulation results, these changes does not represent a considerable impact to SIP. Moreover, the feasibility study demonstrated that based on the current technical characteristics of wireless access points i.e. *computing processing capacity and storage memory*, the implementation of such platform is possible.

In engineering every enhancement comes with a trade-off hence due to the addition and functionalities of the RB, our roaming architecture suffers the effect of roaming signaling triangulation. This introduces an increment in the overall signaling delay in both network registration and session initiation process. An important assumption in our architecture is that the cellular network is always the roaming decision maker hence it must be queried continuously with roaming signaling messages. Thus, we decided to reduce the effect of triangulation while providing robust network security and without compromising the user's service profile. To achieve this, our architecture incorporates PKI & PMI, as described in chapter 5. The implementation of such security mechanisms also relies on our SIP-based roaming protocol. In this perspective,we rely on a trust infrastructure endorsed by PKI & PMI elements distributed along the roaming architecture and the use of SIP-embedded XML Attribute Certificates to prove the rights of the mobile user when roaming or initiating services from a VN.

For evaluation of the impact of security elements and the SIP-embedded XML Attribute Certificates on our architecture, we relied on computer simulation. The performance of our roaming architecture with such additions displayed an enhancement in terms of overall signaling delay, specifically in both network registration and session initiation delay. The fact that authentication and authorization process (*roaming decision making*) are now perform locally (*under the consent of the cellular network*) definitely reduces the signaling message exchange and the delay introduced by the Internet. All this without compromising the network security.

Although we propose a roaming architecture that integrates heterogeneous wireless networks without major changes in current wireless architectures, we still consider there are still open issues for further development. Some of them are the following:

- When using plain SIP for roaming signaling exchange, an intruder can eavesdrop the wireless network and read the information in the SIP messages as they travel in plain text. Although, we suggest secure SIP (*SIP+TLS*) to protect the integrity of such messages, we still have to evaluate the impact of such implementation.
- Our architecture does not relies on an specific security mechanism to guarantee the integrity of the databases in the access points in the VNs.
- As our architecture relies on EAP-based mechanisms for authentication i.e.

TLS, it is also exposed to the security weaknesses of such protocol.

Our research focuses on the technical-related issues that could prevent heterogeneous roaming, however the implementation of such type of roaming also rises new challenges that have nothing to do with technology. As consequence, for an efficient deployment of our architecture it is necessary to provide solutions that address issues different than technological. In this connection, we present in the next section some perspectives and the future work that can improve our proposed roaming architecture and provide additional efficiency and robustness to our research work from technical and business-related perspectives.

6.2 Engineering Significance of Research Findings

The impact of our research findings on the industry and wireless vendors relates to the ability of providing an open platform for the inter-operability of heterogeneous wireless networks within a multi-operator environment. Currently, roaming in heterogeneous wireless networks is only possible between networks that are controlled or managed by the same operator such as the case of UMA technology. We based our roaming architecture on the web 2.0 philosophy, where user provided content contributes to the grow and expansion of the world wide web network. Thus, through the proposed roaming architecture, residential or enterprise wireless local area network can become extensions of cellular network and support cellular services and/or provide new and enhanced services. In consequence, we provide in this section, what we consider our most significant research findings from an engineering point of view:

- User-provided network access: the access point owners will be able to provide open access to cellular users while maintaining the quality of service and security standards provided by the cellular network. Cellular operators will see their coverage area increased by independent wireless operators. With our platform, one wireless access point can offer network access to mobile users of different cellular operators.

- Broker-based access network management: WiFi aggregators can become roaming brokers and establish contracts with cellular operators to manage VNs. Thus, cellular operators will be leveraged of the burden of managing and controlling hundreds or thousands wireless networks deployed throughout the cities.

- Enhanced cellular services: while in the 802.11 or 802.16 wireless network, the mobile user can benefit from higher data rates and enhanced mobile multimedia services.

- A new approach for service authorization: the mobile user service profile can be validated locally by the VN without the need of querying the HN's databases. Cellular operators can create a service profile that will be respected in whatever network the mobile user is. Additionally, they control the kind of services the user is allowed to execute once in the VN. Thus, in

spite of the open access network issue the HN keeps control of the cellular services of the roaming user. Cellular operator can keep track of mobile user services for accounting and billing purposes. Moreover, mobile users cannot access services that they have not paid for.

- Service authentication and authorization: Open network access does not necessarily mean more charge to cellular networks. In this regard, we separate the user authentication and authorization from service authentication and authorization. Thus, the access to the user service profile it is not compromised and the signaling burden in the HN is reduced considerably. Cellular network do not have to respond to AAA queries coming from the VNs.
- Ubiquitous roaming: the mobile users will have access to their cellular services regardless of their location. A user traveling in a country in which their cellular operator have no cellular roaming agreements will be able to access his services through a WiFi network.

The results from our research show that the deployment of such roaming architecture is feasible and can be done with the current wireless technologies in the market. Nowadays, Internet service providers and wireless aggregators such as FON (ref) are establishing agreements in an attempt to form a global community to provide Internet and wireless telephony services. Nevertheless, there are still thousands of wireless hotspots or wireless networks of different kinds that are potential cellular extensions. Our research provides the platform for such networks to join the community and become extensions of cellular networks.

6.3 Perspectives and Future Work

We consider that heterogeneous roaming in multi-operator wireless networks is not a trivial task hence to provide an efficient roaming environment we must rely on: robust vertical handover techniques, efficient mobility management protocols, and optimal roaming decision making.

In this thesis, our SIP-based Roaming Architecture was evaluated using delay as a primary metric. Future work is to evaluate combined QoS metrics such as delay, jitter or packet loss to analyze and extend the capabilities of vertical handover in our architecture. Furthermore, expansion to other network environments such as wireless distribution systems or wireless mesh-networks is also recommended. Additionally, although our research does not focus on business-related issues in the heterogeneous roaming process, we still consider that this subject should be addressed to improve the transparency and efficiency of this process.

In this section, we present what we consider an interesting future research topics in this subject matter.

6.3.1 Future Work

Along our research work and during the conception phase of our roaming architecture, we identified certain issues that drew our attention and which, we are convinced would be interesting research topics.

- Integration of RTCP-based QoS measurements in the VN:
  The RB in our roaming architecture relies on QoS statistics to evaluate the SLA conditions in the VN. With this information, it creates and updates the VN's reputation list. Moreover, relying on real time QoS statistics the RB provides the necessary information to the HN to determine whether or not to trigger the roaming process. The feasibility of this technique was confirmed during the implementation of the testbed. We developed a software agent that continuously monitored QoS statistics from VoIP calls using the RTCP packets in the network, for more details see the Appendix B. Unfortunately, we could not evaluate the performance of such mechanism. The objective of RTCP-based QoS measurements is to provide additional intelligence to the wireless access points, specifically 802.11, to be aware of the QoS statistics in the VN.

- Voice Quality-based Call Admission Control:
  Traditionally, 802.11-based CACs take into account the physical limitations of the wireless access network to decide to admit or reject a call into the network. Current CAC mechanisms limit the calls by number of stations in the network. The main assumption is that all the stations are using the same VoIP applications, same vocoders and protocols.

  In the literature, we can find certain analysis that evaluate the maximal capacity of a 802.11 wireless network based subjective voice quality measurements, more specifically, by using the ITU E-Model (Coupechoux et al., 2004). We consider that such metrics can contribute in the development of a voice quality-based CAC. Such mechanism would admit or reject VoIP calls based on how the call can be perceived by the mobile user upon the current conditions of the wireless network rather than by the number of stations currently in the network.

- Optimal-Network Roaming Decision Making:
  The adoption and deployment of unlicensed wireless technologies contribute to the rapid increment in the density of UWNs in some cities. In this respect, when attempting to roam into a VN, it will be common to find more than one UWN willing to offer roaming services. Here, the cellular network (*the roaming decision maker*) must select the optimal network for the mobile user. The combination of decision theory, MADM (*Multiple Attribute Decision Making*) approaches, and optimization techniques can provide important insight to improve such process.

Our research focused primarily on the technical aspects of heterogeneous roaming however we are aware that heterogeneous networks does not differ only in technology but also in business models and management strategies. In this perspective, we have identified three research areas that could improve heterogeneous multi-operator roaming: efficient accounting mechanisms, new business models, and incentives for roaming provisioning.

- Efficient Accounting Mechanisms:
  The billing issue is paramount when offering roaming services, cellular op-

erators will not deploy any roaming architecture if there is no way to bill mobile users. Thus, we consider that new accounting mechanisms to provide efficient distributed billing platform would motivate the network operators or service providers to participate in the heterogeneous roaming.

- New Business Models:
  New business models can provide solutions about how to commercialize this kind of roaming and maybe to define a future killer application. Remember that under our architecture, there are three main elements participating in this subject: the cellular network, the roaming broker, and the visiting network. Therefore, it is necessary to come up with new models that provide a clear business strategy to allow all the participants of our architecture to be economically rewarded.

- Incentives for Roaming Provisioning:
  The independent UWN are important elements under our architecture as they are willing to share their resources to offer roaming services to external mobile users. Thus, the importance of intelligent and efficient roaming provisioning incentives that motivate the UWNs to become VNs.

We are certain that by addressing the aforementioned technical and business-related issues, the performance of our roaming platform will improve considerably and could be considered for future deployment under real scenarios.

Appendix A

SIP server on a WRT54GL access point

OpenSER is an open source SIP server implementation. Currently, OpenSER supports a set of features including: UDP/TCP/TLS, ENUM, and AAA with RADIUS and database (*MySQL and Postgres*), load balancing, call processing language (*CPL*), and NAT traversal. The project is managed by a board of seven people from different countries, ensuring the project's independence and a fair environment for all contributors.

For the testbed we used an OpenSER-based SIP server for embedded linux-based wireless routers called *milkfish*. The idea behind the *milkfish* project is simple:

> *"to make your internet router able to handle not only your internet traffic - surfing, email, online banking, etc. - but also these inexpensive internet-based phone calls"*. (Milikfish, 2006)

The objective of this study is to analyze the feasibility of implementing an open source SIP server on a wireless router for the further development of a basic B2BUA in a wireless access point. This would add the necessary intelligence to traditional access points to interact with other networks by providing real time network statistics.

To achieve this, we performed the following steps:

1. We installed a GNU/Linux based firmware for embedded devices such as residential wireless access point. For this, we chose OpenWRT as operating system and the Linksys WRT54GL wireless router as device.

2. We configured all the services, NAT and Firewall rules to provide connectivity to wireless devices.

3. We installed an OpenSER-based SIP server on the wireless access point with the following purposes:

   (a) Communication between local users stay local

   (b) Communication between local users and remote internet peers stay between the two milkfished internet routers

(c) Fee-based communication between local and, say, mobile users is done by sharing provider accounts in the same way as sharing trunks of telephone companies in the old telephone systems.

The installation and set-up procedures are explained in detail in the remaining of this appendix.
For further information please refer to http://www.openser.org and http://www.milkfish.org.

A Installing OpenWrt on a WRT54GL

A.1 Select the Firmware

The next step is to locate the distribution of OpenWrt you want to install. It comes in three variations located in three differen directories. The only difference between them is that each has a slightly different set of packages installed.

1. micro/ - contains the least amount of packages required for a functional system. This one is only recommended for those experience with OpenWrt and Linux as it does not contain a web interface to the OpenWrt firmware.
2. bin/ - contains a decent amount of packages (*referred to as a Standard image*) needed for a functional system plus it contains a web interface and pppoe. This one is recommended and is considered the default.
3. pptp/ - contains almost the exact set of packages as the previous Standard image except it includes pptp and not pppoe. It also has the web interface.

A.2 Selects the ./bin/.trx file

In that directory you will notice two different types of files. ”trx” and ”bin” files, the bin files simply repackage the trx in the vendor’s default firmware format and are only used when the trx files can’t be used directly. That may be a little confusing for some people however, since we are using the WRT54GL you should use:

openwrt-wrt54g-<type>.bin

where <type> refers to ”SquashFS” and ”JFFS2”.

- quashFS - SquashFS files include a small compressed filesystem within the firmware itself. The disadvantage is that Squashfs is a readonly filesystem, to save changes and make the filesystem appear writable a separate JFFS2 partition has to be used; the advantage is that Squashfs takes up slightly more space than JFFS2, and you’ll always have the original files on the readonly filesystem which can be used as a boot device for recovery.
- JFFS2 - JFFS2 make the entire filesystem JFFS2. The disadvantage is that this takes slightly more space; the advantage is that changes to included files nolonger leaves behind an old copy on the readonly filesystem.

A.3 Download the Firmware Image

At this point you have read the sections above, and since we are using the LinkSys WRT54GL you now know to choose the openwrt-wrt54g-squashfs.bin , although you can also choose openwrt-wrt54g-jffs2.bin if you don't mind rebooting after the install and the fact that is uses a little more space.

You can choose to download the OpenWrt firmware image onto a Windows machine or Linux. It doesn't matter.

Now, download the latest OpenWrt file found at:

http://downloads.openwrt.org/whiterussian/newest/

After downloading the OpenWrt firmware image it is good practice to make sure that the file is not corrupt. This can be verified by comparing the md5sum from your downloaded image with the md5sum listed in the md5sums file found in the download directory. For win32 platforms use md5sums.exe for GNU/Linux systems use the md5sum command. After running the command with the OpenWrt firmware image you just downloaded, compare the results to the corresponding file found on this page md5sums

A.4 How to connect the WRT54G

Be sure to have your WRT54GL router powered up. Plug the power adaptor into a standard outlet and plug the other end into your router. You should see lights on the front of the router. Now you need to your computer to your router. Your connection may vary based on your network configuration. To connect your computer directly to the router run one end of an ethernet cable into your computer's network card and the other end directly into one of the four numbered ports on the back of your router (not the internet/LAN port). If you are on a network and are plugged into a hub you can plug one end of the ethernet cable into a numbered port on the hub and the other end into a numbered port on the router (you must unplug the hub's uplink button, which will disconnect it from the internet). In either case, if you have a dedicated IP you need to change your IP address so that you can access the router on the same subnet. For example, if your IP address is 192.168.170.55 change it to 192.168.1.55 and you should be able to access your router.

To confirm your connection you can try pinging your router (*ping 192.168.1.1*) or using either Windows or Linux, simply open up your favorite Web Browser and access page 192.168.1.1. If your LinkSys WRT54GL router web interface appears you are connected. If not, you need to solve this problem before moving forward.

A.5 Installing OpenWrt

So you have the OpenWrt firmware image downloaded and have confirmed it is not corrupt. As well, you have confirmed that your computer can communicate with the WRT54GL router.

You have two choices for installation, both methods work on both Linux and Windows. You can can use the LinkSys Web Interface for installing OpenWrt or you can use tftp to install OpenWrt.

If you are not compiling your own OpenWrt source code, simply install OpenWrt using the Routers Web Interface (this is simple... brief instructions below). In fact, even if you plan on installing your own self compiled OpenWrt source code it is easier to install an official version of OpenWrt first using the Router's Web Interface. Why you ask? It's all about boot_wait.

By default, boot_wait is "off" on the WRT54GL routers, in fact, most routers have boot_wait "off" by default. Turning boot_wait "on" simply increases the time it takes to boot. Why is this relevant you may ask? Well, try using tftp to install OpenWrt and you'll find out. Some people claim that atftp for linux is quicker than tftp for windows, this is not necessarily true although the 'a' in atftp does stand for "advanced". If you try to install OpenWrt using tftp on windows, using tftp on Linux or atftp on Linux you get the same error: *error received from server <Invalid Password>*

This happens because tftp client cannot access the router's server. Why? Because there is a small window of opportunity to connect and install OpenWrt using tftp and as fast as you try to be it is near impossible to catch without knowing the trick. Turning boot_wait "on" increases that window of opportunity and makes it easier to catch and install OpenWrt using ftp, however, turning boot_wait "on" is a task in itself and varies across routers. Apparently, on some routers you can install older firmware that enables a ping exploit where commands can be typed directly into the routers Web Interface to turn boot_wait "on" or "off". This does not work with WRT54GL so do not try to roll back the firmware, it is not recommended. In fact, turning boot_wait "on" for installing OpenWrt on the WRT54GL using tftp is not necessary (but it helps).

So, it is recommended that you install an official version of OpenWrt first using the Router's Web Interface before using tftp to install your own self compiled version. This is because once you have installed OpenWrt you can access its Web Interface and turn boot_wait "on". Then you can use tftp and have a much easier time installing OpenWrt. This is one of the easiest ways to turn on the boot_wait option. Hopefully, it saves you hours of searching the forums.

Just so that you know. You do not need the boot_wait option on to successfully install OpenWrt using tftp. It does make it easier but it's not necessary. The trick is to press the reset button and timing. Run the ftp client, power up the WRT54GL router and hit the reset button. If that doesn't work the first time, try again, but it does work.

A.5.1 Router Web Interface Installation

We installed OpenWrt using the the router's web interface due to the simple access to the web interface. To do this, open the web browser and go to the page at 192.168.1.1 then go to system → administration → firmware upgrade. Locate the OpenWrt Firmware Image file and that's it. Be sure that the power supply is stable and not disconnected during transfer. After your installation depending on the type of OpenWrt you chose you may need to restart you router.

A Installing OpenSER (milkfish) in OpenWrt

1. Go to Category "System" - "Installed Software" and click "Update package lists" (*internet connection needs to be up from now on*). See, Figure A.1

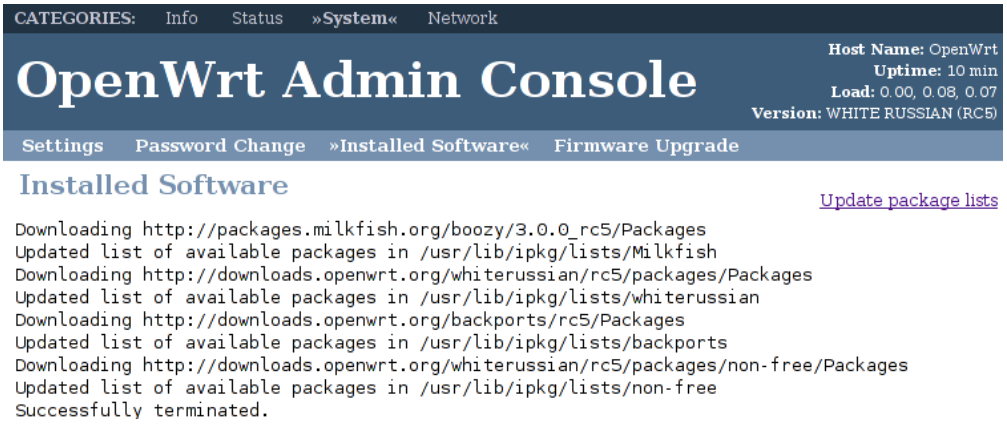


Figure A.1: Step 1: Installing the milkfish software, (Milikfish, 2006)

2. Install the milkfish-sip-with-webif package, as illustrated in Figure A.2.

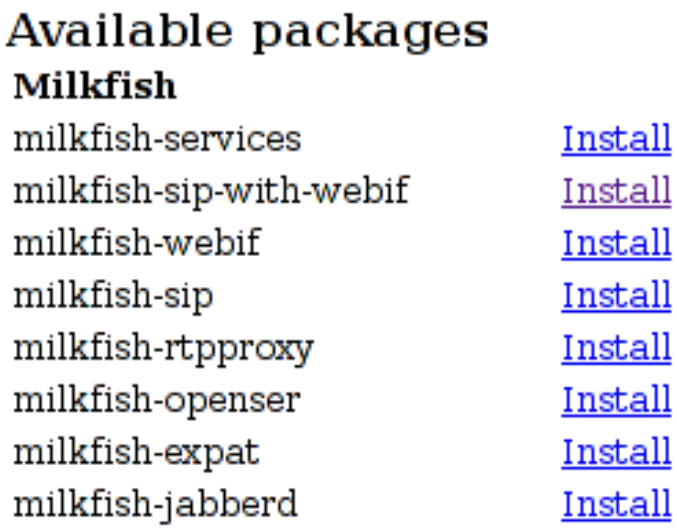


Figure A.2: Step 2: Installing the milkfish web interface, (Milikfish, 2006)

3. This should lead to this screen. If you get errors, remove all milkfish packages and repeat this step once - if the same errors still appear, try with the firmware we provide and/or take the error to the forum. (*all the necessary packages for the main SIP router scenario are downloaded and installed with milkfish-sip-with-webif due to dependencies*). See Figure A.3
4. Reboot the router.

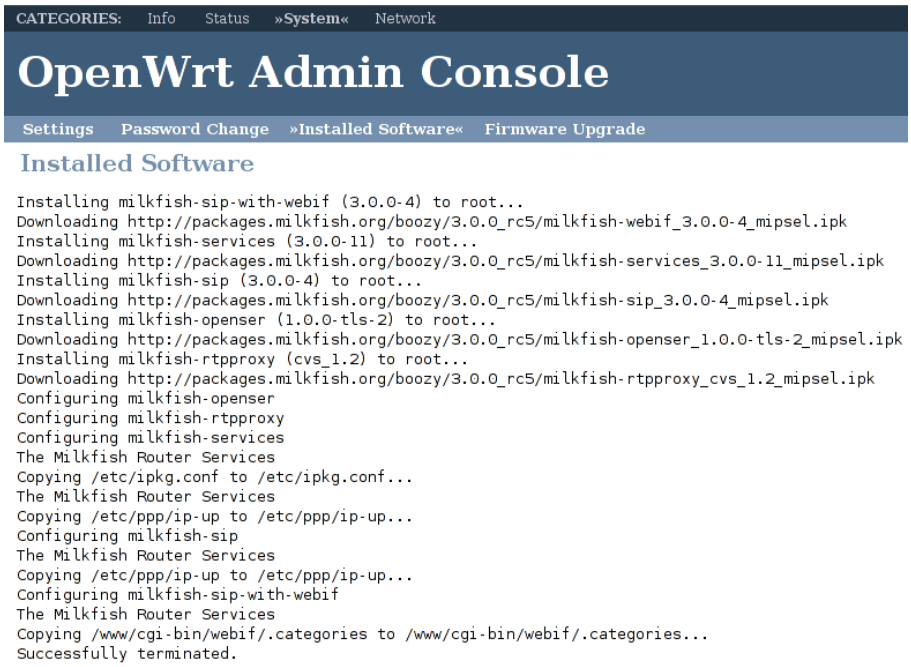


Figure A.3: Step 3: Completing milkfish installation, (Milikfish, 2006)

5. Set the outbound proxy in your phones with provider accounts to the milkfish router ip and disable STUN. Phones which are only used internally need to be set to use the milkfish router as their registrar..

6. and an account needs to be created on the milkfish web interface database section (to use the database frontend on the web interface hit the ”Update Database” button first). See Figure A.4.

A Conclusion

OpenWRT provides a robust platform for the development of new set of tools to enhance wireless communications. A linux-based firmware for wireless access points allows the installation for new services and applications that would increase the *intelligence* of such devices. The support of a RADIUS server and MySQL database enables the wireless router to perform authentication and authorization locally as indicated in this thesis for the case of SIP-based XML Attribute Certificates. In addition, the open *milkfish* project also provides a platform for the further development of new SIP-based services. As a future work, we aim at the development of a simplified B2BUA operating on a wireless access point.

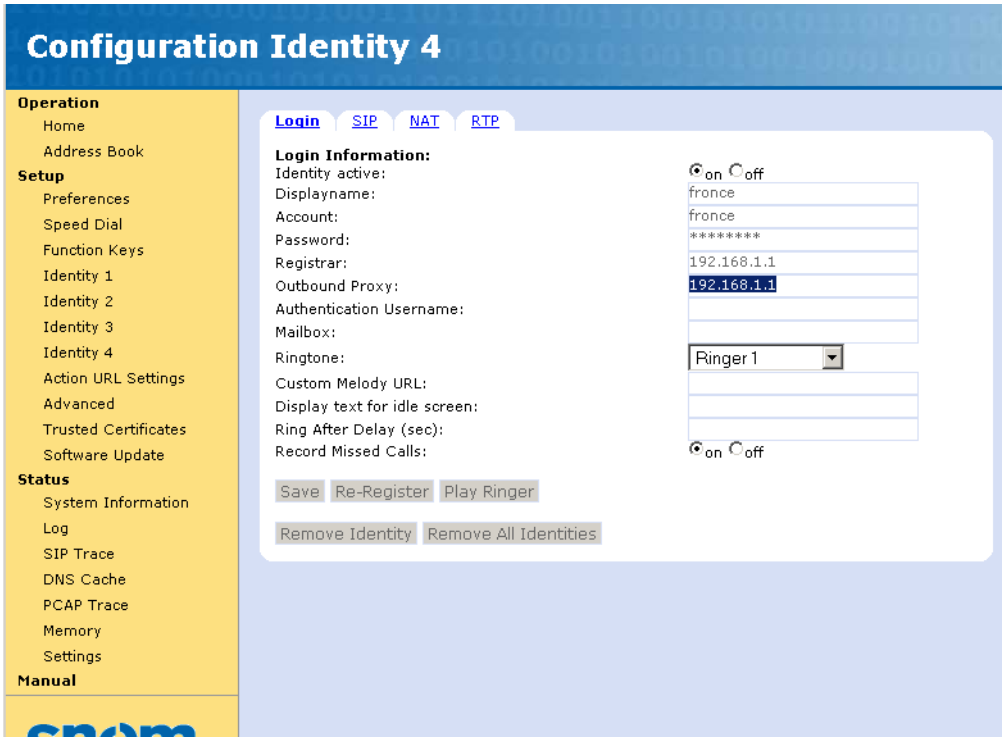


Figure A.4: Step 4: Configuring milkfish, (Milikfish, 2006)

Appendix B

RTCP-based QoS measurements

The objective of this testbed is to analyze the feasibility of obtaining real time QoS statistics from VoIP calls in the WLAN. For this, we developed and install a RTCP analyzer on the wireless access point. This software was able to obtain statistics from ongoing VoIP calls using the information provided by the protocol RTCP. Among such information we were able to calculate the packet loss rate, the RTT delay, and the jitter.

For the successfully accomplishment of this testbed we relied on two SIP User Agents, an OpenWRT-based wireless access point, a RTCP analyzer that we developed, and a SIP outbound server on the access point.

B RTCP Packet Format

RTCP is the Real-time Transport Control Protocol, which may be used as a lightweight companion to RTP to convey a number of statistics and other information about an RTP flow between recipients and senders (Schulzrinne et al., 2003). Thus, all RTCP packets MUST be sent in a compound packet of at least two individual packets, with the following format:

- **SR or RR:**
  The first RTCP packet in the compound packet must always be a report packet to facilitate header validation. This is true even if no data has been sent nor received, in which case an empty RR is sent, and even if the only other RTCP packet in the compound packet is a BYE.

- **Additional RRs:**
  If the number of sources for which reception statistics are being reported exceeds 31, the number that will fit into one SR or RR packet, then additional RR packets should follow the initial report packet.

- **SDES:**
  An SDES packet containing a CNAME item must be included in each compound RTCP packet. Other source description items may optionally be included if required by a particular application, subject to bandwidth constraints.

- **BYE or APP:**
  Other RTCP packet types, including those yet to be defined, may follow in any order, except that BYE should be the last packet sent with a given SSRC/CSRC. Packet types may appear more than once.

B.1 Sender and Receiver Reports

RTP receivers provide reception quality feedback using RTCP report packets which may take one of two forms depending upon whether or not the receiver is also a sender. The only difference between the sender report (*SR*) and receiver report (*RR*) forms, besides the packet type code, is that the sender report includes a 20-byte sender information section for use by active senders. The SR is issued if a site has sent any data packets during the interval since issuing the last report or the previous one, otherwise the RR is issued.

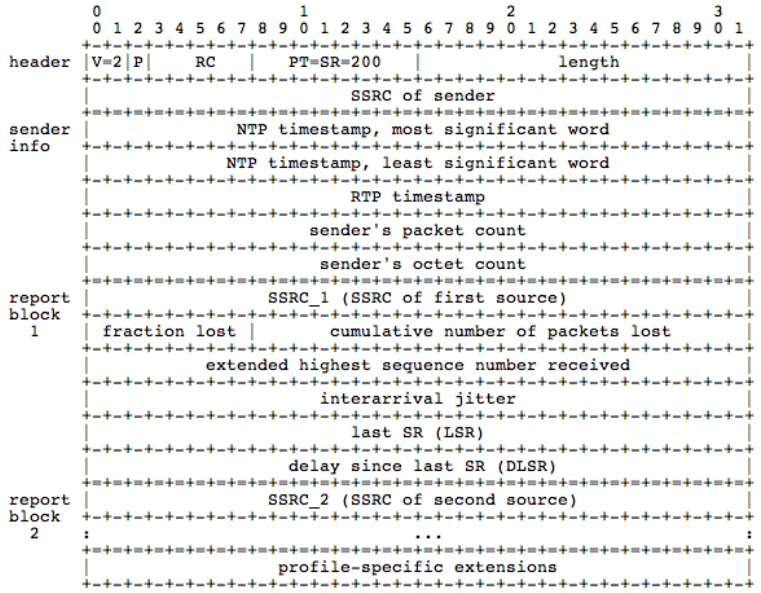


Figure B.1: Sender Report RTCP Packet Format, (Schulzrinne et al., 2003)

As illustrated in Figure B.1, the SR consist of three sections, the fist section (*the header*) is 8 octets long. The fields in the first section have the following meaning:

- **version (V):** 2 bits
  Identifies the version of RTP, which is the same in RTCP packets as in RTP data packets. The version defined by this specification is two (2).

- **padding (P):** 1 bit
  If the padding bit is set, this individual RTCP packet contains some additional padding octets at the end which are not part of the control information but are included in the length field. The last octet of the padding is a count of how many padding octets should be ignored, including itself (*it*

will be a multiple of four). Padding may be needed by some encryption algorithms with fixed block sizes. In a compound RTCP packet, padding is only required on one individual packet because the compound packet is encrypted as a whole for the method in Section 9.1. Thus, padding MUST only be added to the last individual packet, and if padding is added to that packet, the padding bit MUST be set only on that packet. This convention aids the header validity checks and allows detection of packets from some early implementations that incorrectly set the padding bit on the first individual packet and add padding to the last individual packet.

- **reception report count (RC):** 5 bits
  The number of reception report blocks contained in this packet. A value of zero is valid.

- **packet type (PT):** 8 bits
  Contains the constant 200 to identify this as an RTCP SR packet.

- **legth:** 16 bits
  The length of this RTCP packet in 32-bit words minus one, including the header and any padding. (*the offset of one makes zero a valid length and avoids a possible infinite loop in scanning a compound RTCP packet, while counting 32-bit words avoids a validity check for a multiple of 4*).

- **SSRC:** 32 bits
  The synchronization source identifier for the originator of this SR packet.

The second section(*the sender information*), is 20 octets long and is present in every SR packet. Thus, this section summarizes the data transmissions from this sender. The fields have the following meaning:

- **NTP timestamp:** 64 bits
  Indicates the wallclock time when this report was sent so that it may be used in combination with timestamps returned in reception reports from other receivers to measure round-trip propagation to those receivers. Receivers should expect that the measurement accuracy of the timestamp may be limited to far less than the resolution of the NTP timestamp. The measurement uncertainty of the timestamp is not indicated as it may not be known. On a system that has no notion of wallclock time but does have some system-specific clock such as *system uptime*, a sender MAY use that clock as a reference to calculate relative NTP timestamps. It is important to choose a commonly used clock so that if separate implementations are used to produce the individual streams of a multimedia session, all implementations will use the same clock. Until the year 2036, relative and absolute timestamps will differ in the high bit so (*invalid*) comparisons will show a large difference; by then one hopes relative timestamps will no longer be needed. A sender that has no notion of wallclock or elapsed time MAY set the NTP timestamp to zero.

- **RTP timestamp:** 32 bits
  Corresponds to the same time as the NTP timestamp (*above*), but in the same units and with the same random offset as the RTP timestamps in data packets. This correspondence may be used for intra- and inter-media synchronization for sources whose NTP timestamps are synchronized, and may be used by media-independent receivers to estimate the nominal RTP clock frequency. Note that in most cases this timestamp will not be equal to the RTP timestamp in any adjacent data packet. Rather, it MUST be calculated from the corresponding NTP timestamp using the relationship between the RTP timestamp counter and real time as maintained by periodically checking the wallclock time at a sampling instant.

- **sender's packet count:** 32 bits
  The total number of RTP data packets transmitted by the sender since starting transmission up until the time this SR packet was generated. The count SHOULD be reset if the sender changes its SSRC identifier.

- **sender's octet count:** 32 bits
  The total number of payload octets *i.e., not including header or padding*, transmitted in RTP data packets by the sender since starting transmission up until the time this SR packet was generated. The count SHOULD be reset if the sender changes its SSRC identifier. This field can be used to estimate the average payload data rate.

Finally, the third section contains zero or more reception report blocks depending on the number of other sources heard by this sender since the last report. Each reception report block conveys statistics on the reception of RTP packets from a single synchronization source. Receivers SHOULD NOT carry over statistics when a source changes its SSRC identifier due to a collision. These statistics are:

- **SSRC_n (source identifier)**: 32 bits
  The SSRC identifier of the source to which the information in this reception report block pertains.

- **fraction lost:** 8 bits
  The fraction of RTP data packets from source SSRC_n lost since the previous SR or RR packet was sent, expressed as a fixed point number with the binary point at the left edge of the field. (*That is equivalent to taking the integer part after multiplying the loss fraction by 256.*) This fraction is defined to be the number of packets lost divided by the number of packets expected, as defined in the next paragraph. If the loss is negative due to duplicates, the fraction lost is set to zero. Note that a receiver cannot tell whether any packets were lost after the last one received, and that there will be no reception report block issued for a source if all packets from that source sent during the last reporting interval have been lost.

- **cumulative number of packets lost:** 24 bits
  The total number of RTP data packets from source SSRC_n that have been

lost since the beginning of reception. This number is defined to be the number of packets expected less the number of packets actually received, where the number of packets received includes any which are late or duplicates. Thus, packets that arrive late are not counted as lost, and the loss may be negative if there are duplicates. The number of packets expected is defined to be the extended last sequence number received, as defined next, less the initial sequence number received.

- **extended highest sequence number received:** 32 bits
  The low 16 bits contain the highest sequence number received in an RTP data packet from source SSRC_n, and the most significant 16 bits extend that sequence number with the corresponding count of sequence number cycles.Note that different receivers within the same session will generate different extensions to the sequence number if their start times differ significantly.

- **interarrival jitter:** 32 bits
  An estimate of the statistical variance of the RTP data packet interarrival time, measured in timestamp units and expressed as an unsigned integer. The interarrival jitter J is defined to be the mean deviation (*smoothed absolute value*) of the difference D in packet spacing at the receiver compared to the sender for a pair of packets. As shown in the equation below, this is equivalent to the difference in the *relative transit time* for the two packets. The relative transit time is the difference between a packet's RTP timestamp and the receiver's clock at the time of arrival, measured in the same units. If Si is the RTP timestamp from packet i, and Ri is the time of arrival in RTP timestamp units for packet i, then for two packets i and j, D may be expressed accordingly, B.1

$$D(i,j) = (Rj - Ri) - (Sj - Si) = (Rj - Sj) - (Ri - Si) \tag{B.1}$$

  The interarrival jitter SHOULD be calculated continuously as each data packet i is received from source SSRC_n, using this difference D for that packet and the previous packet i-1 in order of arrival (not necessarily in sequence), according to the equation B.2

$$J(i) = J(i-1) + (|D(i-1,i)| - J(i-1))/16 \tag{B.2}$$

  Whenever a reception report is issued, the current value of J is sampled.

  The jitter calculation MUST conform to the formula specified here in order to allow profile-independent monitors to make valid interpretations of reports coming from different implementations. This algorithm is the optimal first-order estimator and the gain parameter 1/16 gives a good noise reduction ratio while maintaining a reasonable rate of convergence.

- **last SR timestamp (LSR):** 32 bits
  The middle 32 bits out of 64 in the NTP timestamp received as part of the

most recent RTCP sender report (*SR*) packet from source SSRC_n. If no SR has been received yet, the field is set to zero.

- **delay since last SR (DLSR):** 32 bits
  The delay, expressed in units of 1/65536 seconds, between receiving the last SR packet from source SSRC_n and sending this reception report block. If no SR packet has been received yet from SSRC_n, the DLSR field is set to zero.

  Let SSRC_r denote the receiver issuing this receiver report. Source SSRC_n can compute the round-trip propagation delay to SSRC_r by recording the time A when this reception report block is received. It calculates the total round-trip time A-LSR using the last SR timestamp (*LSR*) field, and then subtracting this field to leave the round-trip propagation delay as: *A - LSR - DLSR*, as illustrated by Figure B.2.

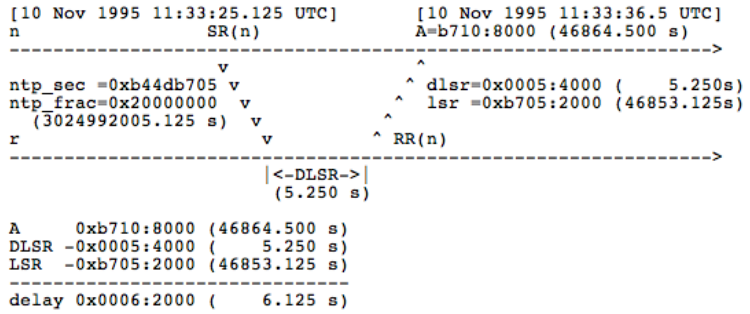


Figure B.2: Round-trip time computation, (Schulzrinne et al., 2003)

Times are shown in both a hexadecimal representation of the 32-bit fields and the equivalent floating- point decimal representation. Colons indicate a 32-bit field divided into a 16-bit integer part and 16-bit fraction part.

This may be used as an approximate measure of distance to cluster receivers, although some links have very asymmetric delays.

B.1.1 Receiver Report RTCP Packet

As depicted by Figure B.3, the format of the receiver report (RR) packet is the same as that of the SR packet except that the packet type field contains the constant 201 and the five words of sender information are omitted (*these are the NTP and RTP timestamps and sender's packet and octet counts*). The remaining fields have the same meaning as for the SR packet.

An empty RR packet (*RC = 0*) MUST be put at the head of a compound RTCP packet when there is no data transmission or reception to report.

B RTCP statistics with the WRT54GL

As stated previously, we enabled the wireless access point to extract RTCP statistics from ongoing VoIP calls. The main reason of this was to prove that the extrac-

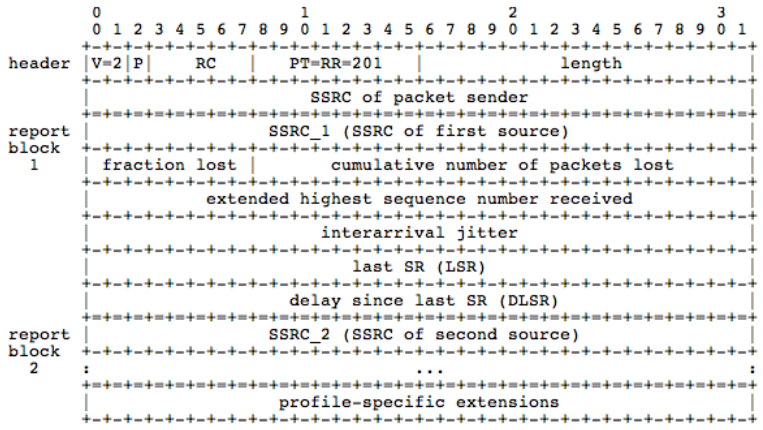


Figure B.3: Receiver Report RTCP Packet Format, (Schulzrinne et al., 2003)

tion of real-time QoS measurements is feasible. Then, with this QoS information the wireless access point will be able to create or update the reputation list. As explained in this thesis, the average QoS statistics of a specific VoIP call can be conveyed within a SIP BYE packet once the conversation is ended. In this section, we explain the testbed configuration and the evaluation process.

B.1 Testbed Configuration

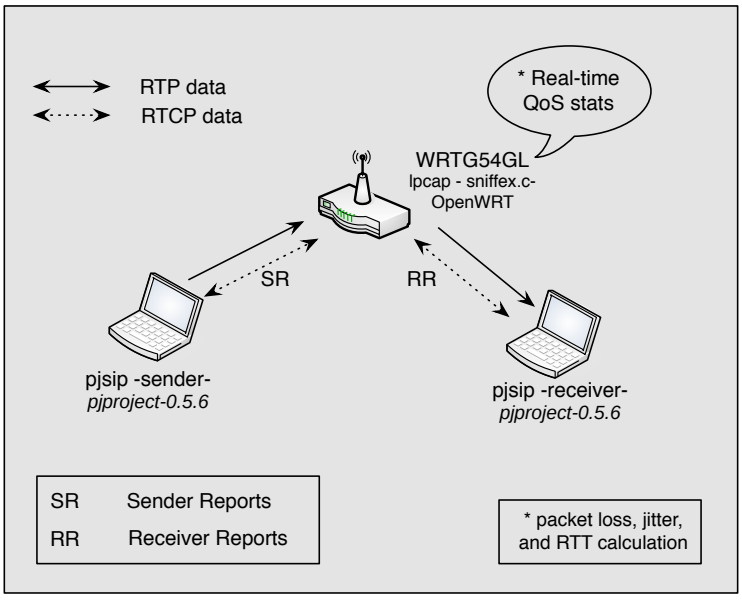


Figure B.4: RTCP statistics with the WRT54GL

As illustrated in Figure B.4, the testbed included two SIP User Agents (*UA*) installed on two laptops with a IEEE 802.11g wireless interface. For the SIP UA

we used the PJSIP Open Source SIP stack (*PJSUA*) due to its robustness and the support of RTCP statistics (Prijono, 2005). In our configuration one computer establishes a VoIP session (*UDP (User Datagram Protocol)/RTP (Real Time Protocol)/RTCP*) and transmitted audio packets to the other peer. The receiver was also running a PJSIP-based SIP UA. A screenshot of the PJSUA application is illustrated in Figure B.5.

Figure B.5: PJSUA Main Interface

One of the clients behave as the caller. Thus, as illustrated in Figure B.6, it transmitted a SIP INVITE message to the other peer to initiate a VoIP session. In the SIP INVITE message (*in the SDP context*) the caller provided all the necessary information to initiate the call e.g. *IP address, softphone, voice codecs supported, etc*. The SIP INVITE message was transmitted to the callee (*IP: 192.168.1.2*).

Upon the reception of the SIP INVITE message, the callee answered with a SIP 200 OK to the caller. This action indicated the that the callee is willing to accept the call. Before the RTP data transmission, the caller acknowledges the reception of the SIP 200 OK message. This is illustrated in Figure B.7. This is the last SIP message between both peers, from now both caller and callee will exchange VoIP data through the RTP protocol.

All the data and control messages (*RTC and RTCP*) exchanged between both peers will pass through the wireless access point. Thus, to enable the extraction of QoS measurements for each VoIP session, we installed our RTCP analyzer on the access point.

B.2 RTCP statistics with lpcap

The RTCP analyzer is a software daemon that runs on the wireless access points. It has a behavior similar to a network sniffer as it is listening continuously the

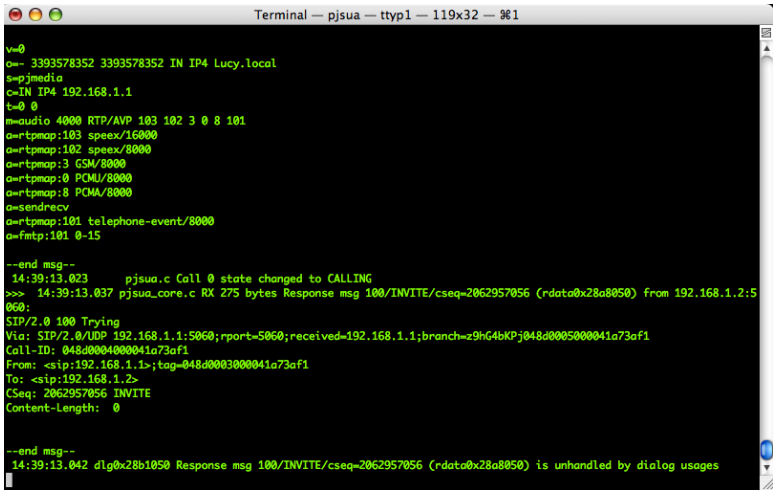


Figure B.6: PJSUA Caller

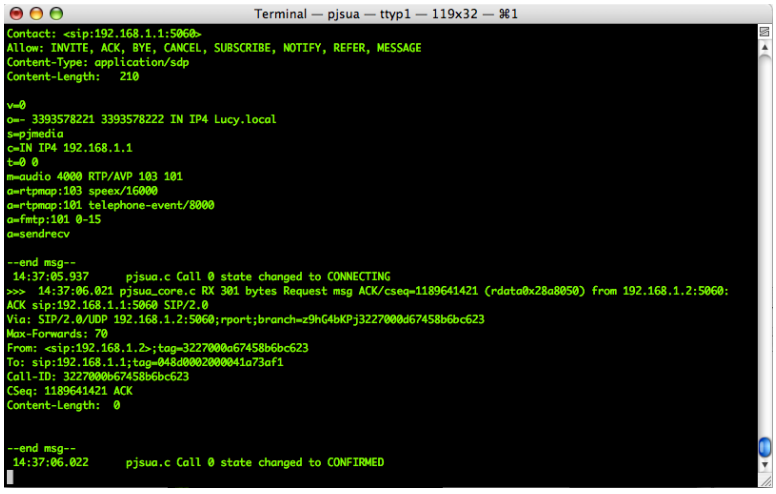


Figure B.7: PJSUA Callee

RTCP port (*in our case the 4001*). Upon the detection of a RTCP packet, it analyzes its content and calculates the QoS statistics. With the information provided by transport and network layer protocol, it is capable of identify the source and the destination and store the real-time QoS information for a specific VoIP call.

To do this we used the Linux pcap library (*lpcap*) for packet capturing and the specifications and equations from (Schulzrinne et al., 2003) for QoS statistics calculation.

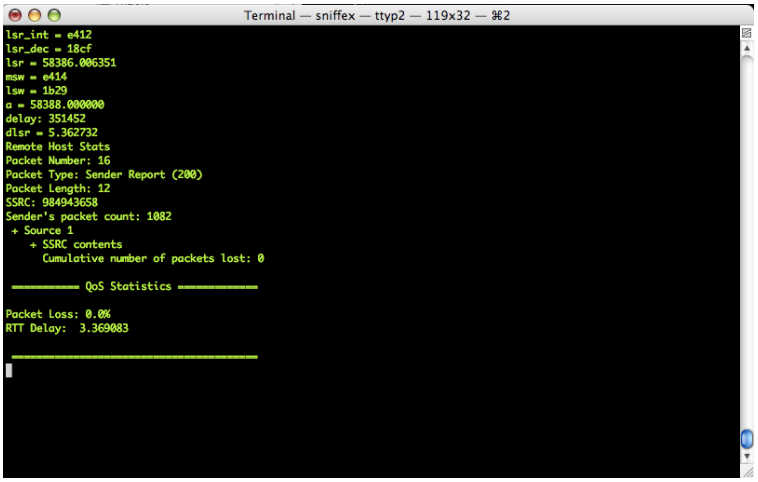


Figure B.8: RTCP statistics with lpcap

Figure B.8 displays an example of the RTCP measurements and QoS statistic calculation for a specific VoIP call. In this example, we see a screenshot of a terminal connected to a wireless access point displaying real time QoS statistics. At this point, the VoIP session had no packet loss and an average delay of 3.36 ms.

The code source of our RTCP monitor is described bellow:

```
/*
 * sniffex.c
 *
 * Sniffer example of TCP/IP packet capture using libpcap.
 *
 * Version 0.1.1 (2005-07-05)
 * Copyright (c) 2005 The Tcpdump Group
 *
 * This software is intended to be used as a practical example and
 * demonstration of the libpcap library; available at:
 * http://www.tcpdump.org/
 *
****************************************************************************
 *
 * This software is a modification of Tim Carstens' "sniffer.c"
 * demonstration source code, released as follows:
```

```
*
* sniffer.c
* Copyright (c) 2002 Tim Carstens
* 2002-01-07
* Demonstration of using libpcap
* timcarst -at- yahoo -dot- com
*
* "sniffer.c" is distributed under these terms:
*
* Redistribution and use in source and binary forms, with or without
* modification, are permitted provided that the following conditions
* are met:
* 1. Redistributions of source code must retain the above copyright
*    notice, this list of conditions and the following disclaimer.
* 2. Redistributions in binary form must reproduce the above copyright
*    notice, this list of conditions and the following disclaimer in the
*    documentation and/or other materials provided with the distribution.
* 4. The name "Tim Carstens" may not be used to endorse or promote
*    products derived from this software without prior written permission
*
* THIS SOFTWARE IS PROVIDED BY THE REGENTS AND CONTRIBUTORS ''AS IS'' AND
* ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE
* IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE
* ARE DISCLAIMED.  IN NO EVENT SHALL THE REGENTS OR CONTRIBUTORS BE LIABLE
* FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL
* DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS
* OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION)
* HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT
* LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY
* OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF
* SUCH DAMAGE.
* <end of "sniffer.c" terms>
*
* This software, "sniffex.c", is a derivative work of "sniffer.c" and is
* covered by the following terms:
*
* Redistribution and use in source and binary forms, with or without
* modification, are permitted provided that the following conditions
* are met:
* 1. Because this is a derivative work, you must comply with the "sniffer.c"
*    terms reproduced above.
* 2. Redistributions of source code must retain the Tcpdump Group copyright
*    notice at the top of this source file, this list of conditions and the
*    following disclaimer.
* 3. Redistributions in binary form must reproduce the above copyright
*    notice, this list of conditions and the following disclaimer in the
*    documentation and/or other materials provided with the distribution.
* 4. The names "tcpdump" or "libpcap" may not be used to endorse or promote
*    products derived from this software without prior written permission.
*
* THERE IS ABSOLUTELY NO WARRANTY FOR THIS PROGRAM.
* BECAUSE THE PROGRAM IS LICENSED FREE OF CHARGE, THERE IS NO WARRANTY
* FOR THE PROGRAM, TO THE EXTENT PERMITTED BY APPLICABLE LAW.  EXCEPT WHEN
* OTHERWISE STATED IN WRITING THE COPYRIGHT HOLDERS AND/OR OTHER PARTIES
* PROVIDE THE PROGRAM "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED
* OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF
* MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE.  THE ENTIRE RISK AS
* TO THE QUALITY AND PERFORMANCE OF THE PROGRAM IS WITH YOU.  SHOULD THE
* PROGRAM PROVE DEFECTIVE, YOU ASSUME THE COST OF ALL NECESSARY SERVICING,
```

```
 * REPAIR OR CORRECTION.
 *
 * IN NO EVENT UNLESS REQUIRED BY APPLICABLE LAW OR AGREED TO IN WRITING
 * WILL ANY COPYRIGHT HOLDER, OR ANY OTHER PARTY WHO MAY MODIFY AND/OR
 * REDISTRIBUTE THE PROGRAM AS PERMITTED ABOVE, BE LIABLE TO YOU FOR DAMAGES,
 * INCLUDING ANY GENERAL, SPECIAL, INCIDENTAL OR CONSEQUENTIAL DAMAGES ARISING
 * OUT OF THE USE OR INABILITY TO USE THE PROGRAM (INCLUDING BUT NOT LIMITED
 * TO LOSS OF DATA OR DATA BEING RENDERED INACCURATE OR LOSSES SUSTAINED BY
 * YOU OR THIRD PARTIES OR A FAILURE OF THE PROGRAM TO OPERATE WITH ANY OTHER
 * PROGRAMS), EVEN IF SUCH HOLDER OR OTHER PARTY HAS BEEN ADVISED OF THE
 * POSSIBILITY OF SUCH DAMAGES.
 * <end of "sniffex.c" terms>
 *
 ********************************************************************************
 *
 * Below is an excerpt from an email from Guy Harris on the tcpdump-workers
 * mail list when someone asked, "How do I get the length of the TCP
 * payload?" Guy Harris' slightly snipped response (edited by him to
 * speak of the IPv4 header length and TCP data offset without referring
 * to bitfield structure members) is reproduced below:
 *
 * The Ethernet size is always 14 bytes.
 *
 * <snip>...</snip>
 *
 * In fact, you *MUST* assume the Ethernet header is 14 bytes, *and*, if
 * you're using structures, you must use structures where the members
 * always have the same size on all platforms, because the sizes of the
 * fields in Ethernet - and IP, and TCP, and... - headers are defined by
 * the protocol specification, not by the way a particular platform's C
 * compiler works.)
 *
 * The IP header size, in bytes, is the value of the IP header length,
 * as extracted from the "ip_vhl" field of "struct sniff_ip" with
 * the "IP_HL()" macro, times 4 ("times 4" because it's in units of
 * 4-byte words).  If that value is less than 20 - i.e., if the value
 * extracted with "IP_HL()" is less than 5 - you have a malformed
 * IP datagram.
 *
 * The TCP header size, in bytes, is the value of the TCP data offset,
 * as extracted from the "th_offx2" field of "struct sniff_tcp" with
 * the "TH_OFF()" macro, times 4 (for the same reason - 4-byte words).
 * If that value is less than 20 - i.e., if the value extracted with
 * "TH_OFF()" is less than 5 - you have a malformed TCP segment.
 *
 * So, to find the IP header in an Ethernet packet, look 14 bytes after
 * the beginning of the packet data.  To find the TCP header, look
 * "IP_HL(ip)*4" bytes after the beginning of the IP header.  To find the
 * TCP payload, look "TH_OFF(tcp)*4" bytes after the beginning of the TCP
 * header.
 *
 * To find out how much payload there is:
 *
 * Take the IP *total* length field - "ip_len" in "struct sniff_ip"
 * - and, first, check whether it's less than "IP_HL(ip)*4" (after
 * you've checked whether "IP_HL(ip)" is >= 5).  If it is, you have
 * a malformed IP datagram.
 *
 * Otherwise, subtract "IP_HL(ip)*4" from it; that gives you the length
```

```
 * of the TCP segment, including the TCP header.  If that's less than
 * "TH_OFF(tcp)*4" (after you've checked whether "TH_OFF(tcp)" is >= 5),
 * you have a malformed TCP segment.
 *
 * Otherwise, subtract "TH_OFF(tcp)*4" from it; that gives you the
 * length of the TCP payload.
 *
 * Note that you also need to make sure that you don't go past the end
 * of the captured data in the packet - you might, for example, have a
 * 15-byte Ethernet packet that claims to contain an IP datagram, but if
 * it's 15 bytes, it has only one byte of Ethernet payload, which is too
 * small for an IP header.  The length of the captured data is given in
 * the "caplen" field in the "struct pcap_pkthdr"; it might be less than
 * the length of the packet, if you're capturing with a snapshot length
 * other than a value >= the maximum packet size.
 * <end of response>
 *
 ****************************************************************************
 *
 * Example compiler command-line for GCC:
 *   gcc -Wall -o sniffex sniffex.c -lpcap
 *
 ****************************************************************************
 *
 * Code Comments
 *
 * This section contains additional information and explanations regarding
 * comments in the source code. It serves as documentaion and rationale
 * for why the code is written as it is without hindering readability, as it
 * might if it were placed along with the actual code inline. References in
 * the code appear as footnote notation (e.g. [1]).
 *
 * 1. Ethernet headers are always exactly 14 bytes, so we define this
 * explicitly with "#define". Since some compilers might pad structures to a
 * multiple of 4 bytes - some versions of GCC for ARM may do this -
 * "sizeof (struct sniff_ethernet)" isn't used.
 *
 * 2. Check the link-layer type of the device that's being opened to make
 * sure it's Ethernet, since that's all we handle in this example. Other
 * link-layer types may have different length headers (see [1]).
 *
 * 3. This is the filter expression that tells libpcap which packets we're
 * interested in (i.e. which packets to capture). Since this source example
 * focuses on IP and TCP, we use the expression "ip", so we know we'll only
 * encounter IP packets. The capture filter syntax, along with some
 * examples, is documented in the tcpdump man page under "expression."
 * Below are a few simple examples:
 *
 * Expression                  Description
 * ----------                  -----------
 * ip                                Capture all IP packets.
 * tcp                               Capture only TCP packets.
 * tcp port 80                 Capture only TCP packets with a port equal to 80.
 * ip host 10.1.2.3            Capture all IP packets to or from host 10.1.2.3.
 *
 ****************************************************************************
 *
 */
```

```
#define APP_NAME                "sniffex"
#define APP_DESC                "Sniffer example using libpcap"
#define APP_COPYRIGHT   "Copyright (c) 2005 The Tcpdump Group"
#define APP_DISCLAIMER  "THERE IS ABSOLUTELY NO WARRANTY FOR THIS PROGRAM."

#include <pcap.h>
#include <stdio.h>
#include <string.h>
#include <stdlib.h>
#include <ctype.h>
#include <errno.h>
#include <sys/types.h>
#include <sys/socket.h>
#include <netinet/in.h>
#include <arpa/inet.h>

/* default snap length (maximum bytes per packet to capture) */
#define SNAP_LEN 1518

/* ethernet headers are always exactly 14 bytes [1] */
#define SIZE_ETHERNET 14

/* Ethernet addresses are 6 bytes */
#define ETHER_ADDR_LEN  6

/* Ethernet header */
struct sniff_ethernet {
        u_char  ether_dhost[ETHER_ADDR_LEN];    /* destination host address */
        u_char  ether_shost[ETHER_ADDR_LEN];    /* source host address */
        u_short ether_type;                     /* IP? ARP? RARP? etc */
};

/* IP header */
struct sniff_ip {
        u_char  ip_vhl;                 /* version << 4 | header length >> 2 */
        u_char  ip_tos;                 /* type of service */
        u_short ip_len;                 /* total length */
        u_short ip_id;                  /* identification */
        u_short ip_off;                 /* fragment offset field */
        #define IP_RF 0x8000            /* reserved fragment flag */
        #define IP_DF 0x4000            /* dont fragment flag */
        #define IP_MF 0x2000            /* more fragments flag */
        #define IP_OFFMASK 0x1fff       /* mask for fragmenting bits */
        u_char  ip_ttl;                 /* time to live */
        u_char  ip_p;                   /* protocol */
        u_short ip_sum;                 /* checksum */
        struct  in_addr ip_src,ip_dst;  /* source and dest address */
};
#define IP_HL(ip)               (((ip)->ip_vhl) & 0x0f)
#define IP_V(ip)                (((ip)->ip_vhl) >> 4)

struct sniff_udp {
        u_short th_sport;               /* source port */
        u_short th_dport;               /* destination port */
        u_short th_len;                 /* length */
        u_short th_sum;                 /* checksum */
};

void
```

```
got_packet(u_char *args, const struct pcap_pkthdr *header, const u_char *packet);

void
print_payload(const u_char *payload, int len);

unsigned int
shift(int array[100], int start, int end);

unsigned int shift(int array[100], int start, int end)
{
int i;
unsigned int t=0;
        for (i=start; i<end; i++) {
        t = t << 8;
        t = t+array[i];
        }

 return(t);
}

/*
 * print packet payload data (avoid printing binary data)
 */
void
print_payload(const u_char *payload, int len)
{
        int offset = 0;                                 /* zero-based offset counter */
        const u_char *ch = payload;

        static int count = 1;                   /* packet counter */

        int i, tempo_sndpkt, cum_pkt_lost, rtcp_pkt[100];
        unsigned int ssrc, lsr_msw, lsr_lsw=0, lsr_int=0, lsr_dec, snd_pktcnt, delay;
        double avg_pkt_loss, dlsr, lsr, a, rtt=0;


        if (len <= 0)
        return;

        printf("\n");

        /* Print Packet Stats */
        ch = payload;
        for (i = 0; i < 100; i++) {
                rtcp_pkt[i] = *ch;
                ch++;
        }

        /* Preparing SSRC */

        ssrc = shift(rtcp_pkt, 4, 8);

        system("clear");

        /* Preparing LSR  */

        lsr_int = shift(rtcp_pkt, 10, 12);
        lsr_dec = shift(rtcp_pkt, 12, 14);
```

```
printf("lsr_int = %x\n", lsr_int);
printf("lsr_dec = %x\n", lsr_dec);

lsr = (double)(lsr_int) + (double)(double)((lsr_dec)/1000000.);

printf("lsr = %f\n",lsr);

/* Preparing packet count */
snd_pktcnt = shift(rtcp_pkt, 20, 24);

if (snd_pktcnt > 0)
tempo_sndpkt = snd_pktcnt;

/* Preparing Last SR timestamp */

lsr_msw = shift(rtcp_pkt, 44, 46);
lsr_lsw = shift(rtcp_pkt, 46, 48);

printf("msw = %x\n", lsr_msw);
printf("lsw = %x\n", lsr_lsw);
a = (double)(lsr_msw) + (double)(double)(lsr_lsw/1000000.);
a = (double)(lsr_msw);

printf("a = %f\n", a);

/* Preparing Delay */

delay = shift(rtcp_pkt, 48, 52);
printf("delay: %u\n", delay);
dlsr = (double)(delay/65536.);


printf("dlsr = %f\n", dlsr);

/* Calculating Packet Loss in % */

count++;
tempo_sndpkt++;

cum_pkt_lost = shift(rtcp_pkt, 33, 36);
avg_pkt_loss = (cum_pkt_lost*100.)/(tempo_sndpkt);

/* Calculating RTT delay */

rtt = a-dlsr-lsr;

/* printing info .... */

printf("Remote Host Stats \n");
printf("Packet Number: %d\n", count);

 /* Packet Type */
if (rtcp_pkt[1] == 200)
printf("Packet Type: Sender Report (%d)\n",rtcp_pkt[1]);
if (rtcp_pkt[1] == 201)
printf("Packet Type: Receiver Report (%d)\n",rtcp_pkt[1]);
if (rtcp_pkt[1] == 202)
printf("Packet Type: Source Description (%d)\n", rtcp_pkt[1]);
if (rtcp_pkt[1] == 203)
```

```
        printf("Packet Type: Goodbye (%d)\n", rtcp_pkt[1]);

        printf("Packet Length: %d\n",rtcp_pkt[3]);
        printf("SSRC: %u\n", ssrc);
        printf("Sender's packet count: %u\n", snd_pktcnt);
        printf(" + Source 1\n");
        printf("    + SSRC contents\n");
        printf("      Cumulative number of packets lost: %d\n", cum_pkt_lost);

        printf("\n");
        printf(" ========== QoS Statistics ============\n");
        printf("\n");
        printf("Packet Loss: %2.1f%%\n", avg_pkt_loss);
        printf("RTT Delay: %7.3f\n", rtt);
        printf("\n");
        printf(" =======================================\n");

return;
}

/*
 * dissect/print packet
 */
void
got_packet(u_char *args, const struct pcap_pkthdr *header, const u_char *packet)
{

        static int count = 1;                   /* packet counter */

        /* declare pointers to packet headers */
        const struct sniff_ethernet *ethernet;  /* The ethernet header [1] */
        const struct sniff_ip *ip;              /* The IP header */
        const struct sniff_udp *udp;            /* The UDP header */
        const char *payload;                    /* Packet payload */

        int size_ip;
        int size_udp;
        int size_payload;

        size_udp = sizeof(struct sniff_udp);

        /*printf("\nPacket number %d:\n", count);*/
        count++;

        /* define ethernet header */
        ethernet = (struct sniff_ethernet*)(packet);

        /* define/compute ip header offset */
        ip = (struct sniff_ip*)(packet + SIZE_ETHERNET);
        size_ip = IP_HL(ip)*4;
        if (size_ip < 20) {
                printf("   * Invalid IP header length: %u bytes\n", size_ip);
                return;
        }

        /* print source and destination IP addresses */
/*      printf("\n"); */
/*      printf("   From: %s\n", inet_ntoa(ip->ip_src)); */
/*      printf("   To: %s\n", inet_ntoa(ip->ip_dst));  */
```

```
        /* determine protocol */
        switch(ip->ip_p) {
                case IPPROTO_TCP:
                        printf("   Protocol: TCP\n");
                        return;
                case IPPROTO_UDP:
                        /*printf("   Protocol: UDP\n");*/
                        break;
                case IPPROTO_ICMP:
                        printf("   Protocol: ICMP\n");
                        return;
                case IPPROTO_IP:
                        printf("   Protocol: IP\n");
                        return;
                default:
                        printf("   Protocol: unknown\n");
                        return;
        }

        /*
         *  OK, this packet is UDP.
         */

        /* define/compute udp header offset */
        udp = (struct sniff_udp*)(packet + SIZE_ETHERNET + size_ip);

        /*printf("   Src port: %d\n", ntohs(udp->th_sport));*/
        /*printf("   Dst port: %d\n", ntohs(udp->th_dport));*/

        /* define/compute udp payload (segment) offset */
        payload = (u_char *)(packet + SIZE_ETHERNET + size_ip + size_udp);

        /* compute udp payload (segment) size */
        size_payload = ntohs(ip->ip_len) - (size_ip + size_udp);

        /*
         * Print payload data; it might be binary, so don't just
         * treat it as a string.
         */
        if (size_payload > 0) {
                /*printf("   Payload (%d bytes):\n", size_payload);*/
                print_payload(payload, size_payload);
        }

return;
}

int main (int argc, char **argv)
{

        char *dev = NULL;                       /* capture device name */
        char errbuf[PCAP_ERRBUF_SIZE];          /* error buffer */
        pcap_t *handle;                         /* packet capture handle */

        char filter_exp[] = "udp port 4001";            /* filter expression [3] */
        struct bpf_program fp;                  /* compiled filter program (expression) */
        bpf_u_int32 mask;                       /* subnet mask */
```

```
        bpf_u_int32 net;                        /* ip */
        int num_packets = 1000;                 /* number of packets to capture */

        /* check for capture device name on command-line */
        if (argc == 2) {
                dev = argv[1];
        }
        else if (argc > 2) {
                fprintf(stderr, "error: unrecognized command-line options\n\n");
                exit(EXIT_FAILURE);
        }
        else {
                /* find a capture device if not specified on command-line */
                dev = pcap_lookupdev(errbuf);
                if (dev == NULL) {
                        fprintf(stderr, "Couldn't find default device: %s\n",
                            errbuf);
                        exit(EXIT_FAILURE);
                }
        }

        /* get network number and mask associated with capture device */
        if (pcap_lookupnet(dev, &net, &mask, errbuf) == -1) {
                fprintf(stderr, "Couldn't get netmask for device %s: %s\n",
                    dev, errbuf);
                net = 0;
                mask = 0;
        }

        /* print capture info */
        /*printf("Device: %s\n", dev);*/
/*      printf("Number of packets: %d\n", num_packets);*/
        /*printf("Filter expression: %s\n", filter_exp);*/

        /* open capture device */
        handle = pcap_open_live(dev, SNAP_LEN, 1, 1000, errbuf);
        if (handle == NULL) {
                fprintf(stderr, "Couldn't open device %s: %s\n", dev, errbuf);
                exit(EXIT_FAILURE);
        }

        /* make sure we're capturing on an Ethernet device [2] */
        if (pcap_datalink(handle) != DLT_EN10MB) {
                fprintf(stderr, "%s is not an Ethernet\n", dev);
                exit(EXIT_FAILURE);
        }

        /* compile the filter expression */
        if (pcap_compile(handle, &fp, filter_exp, 0, net) == -1) {
                fprintf(stderr, "Couldn't parse filter %s: %s\n",
                    filter_exp, pcap_geterr(handle));
                exit(EXIT_FAILURE);
        }

        /* apply the compiled filter */
        if (pcap_setfilter(handle, &fp) == -1) {
                fprintf(stderr, "Couldn't install filter %s: %s\n",
                    filter_exp, pcap_geterr(handle));
                exit(EXIT_FAILURE);
```

```
        }

        /* now we can set our callback function */
        pcap_loop(handle, num_packets, got_packet, NULL);

        /* cleanup */
        pcap_freecode(&fp);
        pcap_close(handle);

        printf("\nCapture complete.\n");

return 0;
}
```

B Conclusion

This testbed was useful to prove the feasibility of real-time RTCP data extraction and QoS calculations from a wireless access point. A future work would be to use to develop a SIP-based application that uses such QoS information to update the reputation list in the Roaming Broker. Another interesting approach would be the implementation of a RTCP-based call admission control to determine whether or not to accept a new visitor in the network.

Appendix C

Impact of Access Control on SIP Network Registration

The objective of this analysis was to evaluate the impact of access control mechanisms on the SIP network registration process. To achieve this, we recreated three possible scenarios: an open access WLAN, a WLAN protected by WPA-PSK, and a WLAN protected by RADIUS-based WPA-TTLS. The testbed, as illustrated in Figure **??**, included a dual-interface laptop (*WiFi and Ethernet*), a Fast-Ethernet Hub, a 802.11g wireless access point, a RADIUS & SIP server, and a network sniffer for delay statistics.

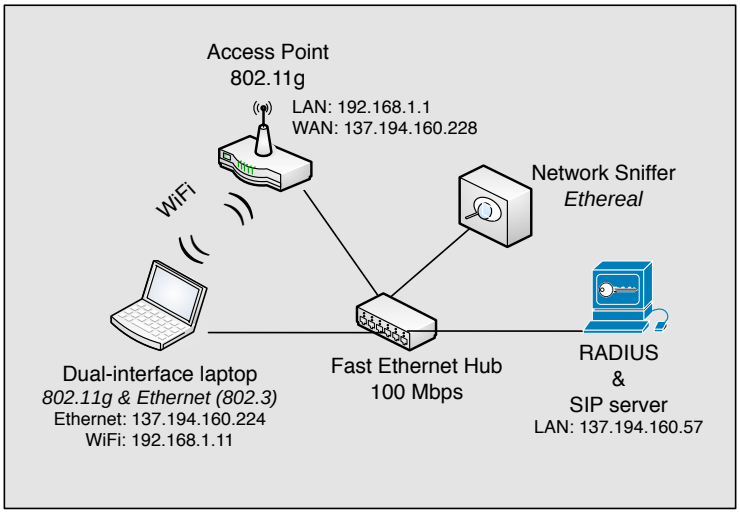


Figure C.1: Testbed for Access Control mechanism

The evaluation process was the following, as we are not interested on evaluating a vertical handover, but the delay introduced by wireless access point association (*DHCP configuration*) and the AAA mechanism: only one network interface in the laptop is active at the same time hence once the WiFi interface is enabled, the Ethernet interface is automatically disabled. This is performed through a shell script that invokes the *ifconfig* command. This action triggers the following pro-

cess:

1. Network interface switching (*Ethernet to WiFi*)
2. SIP de-registration
3. Access point association and AAA validation
4. SIP registration

C.1 Evaluation Technique

The delay evaluation is performed through a network sniffer based on the time stamp of the packets exchanged between the client and the servers (*AAA or SIP server*). The sniffer used for this evaluation was Ethereal and is located in the AAA/SIP server. This allows us to obtain the time- stamp of the de-registration SIP message and the time-stamp the of the new registration message. The time between both of them represents the network switching delay (*the time spent by the operating system disabling the old and enabling the new network interface*).

C.2 Evaluation Results

From the results illustrated in Figure C.2, we observed that the network switching delay component is in the order of ≈ 5 seconds. This delay is product of the network interface switching as only one interface is active at the time and it is related to the way the operating system handles the activation of network interfaces. Although this process introduced considerable delay, under a vertical handover scenario we commonly assumed that both interfaces are enabled at the same time and the switching is performed from upper layers of the OSI model (*Data Link Layer and above*). On the other hand, the access point association, which comprises the DHCP and AAA validation exhibited also a considerable delay. In the open access WLAN (*no AAA*) the association delay was in the order of ≈ 2.8 seconds, mainly introduced by the DHCP process. In the protected networks WPA-PSK exhibited a slightly higher delay, presumably introduced by the AAA mechanism. Nevertheless, in this case the AAA only represented an increment of ≈ 350 ms. The WPA-TTLS scenario displayed the higher delay, as the AAA introduced a delay of $\approx 550 - 600$ ms compared to the open access WLAN. From this, we can see that there is a clear trade-off between delay and security robustness.

Additionally, from the three approaches, we see that the SIP registration and de-registration time remains similar $\approx 3ms$. From these results, we infer that AAA mechanisms introduce different delay levels based on their architecture. Although, the certificate-based AAA approach exhibited the higher delay, it remains acceptable as it was j ust $\approx 200ms$ higher than traditional approaches (*WPA-PSK*).

C.3 Conclusion

The network switching introduces an important delay component, however current physical-layer-based vertical handover techniques and the ability of dual-interface operation can enhance this process. Furthermore, the delay introduced

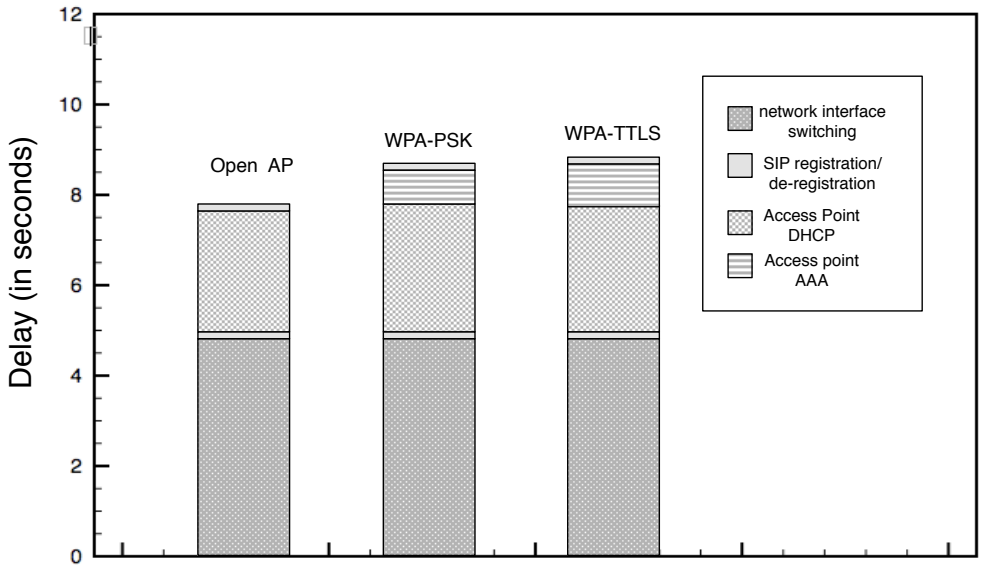


Figure C.2: Network Registration Delay

by wireless association can be improved through network pre-association techniques. On the other hand, we have also seen that when AAA required the delay figure goes higher. Specially, if the RADIUS server is accessed through the Internet, which in this case we have to add an extra Internet delay.

The data obtained through this testbed provided important information for the design of the roaming architecture. Concepts such as local authentication and authorization, and the separation of user AAA and service AAA were considered after the implementation of this testbed.

Publications

Peer reviewed

- O.Salazar, P. Martins, J. Demerjian, and S. Tohmé, *Enabling Roaming in Heterogeneous Multi-Operator Wireless Environments*, (*to appear*) in the Journal of Wireless Communications. Academy Publishers. September 2007.

Conference Proceedings

- O.Salazar, P. Martins, J. Demerjian, and S. Tohmé, *A SIP-based Roaming Architecture For Multi-Operator Wireless Networks*, in Proceedings of Innovations in Information Technologies, Dubai, United Arab Emirates, November 2006.
- O.Salazar, P. Martins, J. Demerjian, and S. Tohmé, *Fast & Secure Roaming For Heterogeneous Multi-Operator Wireless Networks*, in Proceedings of Wireless Telecommunications Symposium, April 2007. Pomona, California, USA.
- O.Salazar, P. Martins, J. Demerjian, and S. Tohmé, *SIP-based Roaming Signaling For Heterogeneous Multi-Operator Wireless Networks*, in Proceedings of IASTED Wireless and Optical Communications, May 2007. Montreal, Canada.
- O.Salazar, P. Martins, J. Demerjian, and S. Tohmé *SIP-embedded XML Attribute Certificates for Heterogeneous Roaming in Multi-Operator Wireless Environments*. (*to appear*) in Proceedings of Vehicular Technology Conference (*VTC*) Fall 2007.

Bibliography

3GPP Rel. 5 (2002). *3GPP Release 5, Technical Specifications and Technical Reports for a UTRAN-based 3GPP system*.

3GPP Rel. 6 (2004). *3GPP Release 6, Technical Specifications and Technical Reports for a UTRAN-based 3GPP system*.

3GPP Rel. 99 (2000). *3GPP Release 99, Technical Specifications and Technical Reports for a UTRAN-based 3GPP system*.

3GPP TR 22.980 (2006). *Network Composition*, 3GPP TR 22.980 version 8.1.0.

3GPP TS 23.234 (2006). *Technical Specification Group Services and System Aspects; 3GPP system to Wireless Local Area Network (WLAN) interworking; System description (Release 6)*, 3GPP TS 23.234 version 6.10.0.

3GPP TS 44.318 (2005). *Generic Access (GA) to the A/Gb interface*, Mobile GA interface layer 3 specification.

Aboba, B. & Simon, D. (1999). *PPP EAP TLS Authentication Protocol*, RFC 2716.

Ahlgren, B., Eggert, L., Ohlman, B. & Schieder, A. (2005). Ambient Networks: Bridging Heterogeneous Network Domains, *Proc. 16th Annual IEEE International Symposium on Personal Indoor and Proc. 16 th Annual IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), Berlin, Germany*.

Akyildiz, I. F., Xie, J. & Mohanty, S. (2004). A survey of mobility management in next-generation all-ip-based wireless systems, *Wireless Communications, IEEE [see also IEEE Personal Communications]* **11**(4): 16–28.

Balachandran, A., Voelker, G. M. & Bahl, P. (2005). Wireless hotspots: current challenges and future directions, *Mob. Netw. Appl.* **10**(3): 265–274.

Banerjee, N., Das, S. K. & Acharya, A. (2005). SIP-Based Mobility Architecture for Next Generation Wireless Networks, *PERCOM '05: Proceedings of the Third IEEE International Conference on Pervasive Computing and Communications*, IEEE Computer Society, pp. 181–190.

Calhoun, P., Loughney, J., Guttman, E., Zorn, G. & Arkko, J. (2003). *Diameter Base Protocol*, RFC 3588.

Cam-Winget, N., Housley, R., Wagner, D. & Walker, J. (2003). Security flaws in 802.11 data link protocols, *Commun. ACM* **46**(5): 35–39.

Chadwick, D. (2005). The X.509 Privilege Management Standard, *Upgrade - The European Journal for the Informatics Professional* **VI**(4): 41–46. Available from http://www.upgrade-cepis.org/issues/2005/4/up6-4Chadwick.pdf.

Chadwick, D. & Otenko, A. (2002). RBAC Policies in XML for X.509 Based Privilege Management, *in* M. A. Ghonaimy, M. T. El-Hadidi & H. Aslan (eds), *Security in the Information Society: Visions and Perspectives: IFIP TC11 17th Int. Conf. On Information Security (SEC2002), Cairo, Egypt*, Kluwer Academic Publishers, pp. 39–53.

Coupechoux, M., Kumar, V. & Brignol, L. (2004). Voice over IEEE 802.11b Capacity, *16th ITC Specialist Seminar on Performance Evaluation of Wireless and Mobile Networks*.

dd wrt (2006). http://www.dd-wrt.com, RFC 2560.

E.212, I.-T. R. (2004-2005). The international identification plan for mobile terminals and mobile users, *International Telecommunications Union*.

Edwards, N., el Khazen, K., Falk, R., Hjelm, J., Howard, P., Howker, K., Jefferies, N., Kowalski, S., Lucking, U., Melen, B., Nystrom, J., Roux, P., Selander, G. & Wang, H. (2004). WWI Whitepaper on Trust, *WWI Whitepaper*.

ETSI (June 1997). *ETSI TS 22.70. Universal Mobile Telecommunication System (UMTS): Virtual Home Environment*, *Draft Version*.

ETSI BRAN (2000a). *HIPERLAN Type 2; Data Link Control (DLC) Layer; Part2: Radio Link Control (RLC) Sublayer", ETSI Report TR 101 761-2, Ver. 1.1.1*, HIPERLAN Type 2; Data Link Control (DLC) Layer; Part2: Radio Link Control (RLC) Sublayer", ETSI Report TR 101 761-2, Ver. 1.1.1.

ETSI BRAN (2000b). *HIPERLAN Type 2; Physical (PHY) Layer", ETSI Report TR 101 475, Ver. 1.1.1*, HIPERLAN Type 2; Physical (PHY) Layer", ETSI Report TR 101 475, Ver. 1.1.1.

ETSI BRAN (2002). *P802.16 D5c-2002, HiperMAN Draft*, Tech. Rep. ETSI v.D5c.

Gao, D., Cai, J. & Ngan, K. N. (2005). Admission Control in IEEE 802.11e Wireless LANs, *IEEE Network Magazine* **19**(4): 6–13.

Garcia-Martin, M. (May 2005). *Input 3rd-Generation Partnership Project (3GPP) Release 5 Requirements on the Session Initiation Protocol (SIP)*, IETF RFC 4083.

Garg, S. & Kappes, M. (2003a). Admission control for VoIP traffic in IEEE 802.11 networks, *Global Telecommunications Conference, 2003. GLOBECOM '03. IEEE*, Vol. 6, pp. 3514–3518 vol.6.

Garg, S. & Kappes, M. (2003b). Can I add a VoIP Call?, *IEEE International Conference on Communications, ICC '03*, Vol. 2, pp. 779–783.

Grandison, T. & Sloman, M. (2000). A survey of trust in Internet application, *IEEE Communications Surveys, Fourth Quarter*. Available from: `citeseer.ist.psu.edu/grandison00survey.html`.

Handley, M. & Jacobson, V. (April 1998). *SDP: Session Description Protocol*, IETF RFC 2327.

Haverinen, H., Mikkonen, J. & Takamaki, T. (2002). Cellular access control and charging for mobile operator wireless local area networks, *Wireless Communications, IEEE* [*see also IEEE Personal Communications*] **9**(6): 52–60.

Hess, A. & Schafer, G. (2002). Performance Evaluation of AAA/Mobile IP Authentication, *Proc. of 2nd Polish-German Teletraffic Symposium (PGTS'02)*, Proc. of 2nd Polish-German Teletraffic Symposium (PGTS'02).

Hole, D. P. & Tobagi, F. A. (2004). Capacity of an IEEE 802.11b Wireless LAN supporting VoIP, *Proc. IEEE Int. Conference on Communications (ICC)*.

IEEE 802.11 Standard (1999). *EEE 802.11 Standard, Wireless LAN Medium Access Control (MAC) and Physical layer (PHY) Specifications*, IEEE Standard.

IEEE 802.16-2001 (2001). *IEEE 802.16 Standard, Air Interface for Fixed Broadband Wireless Access Systems*, IEEE Standard.

IEEE 802.16-2005 (2005). *IEEE 802.16 Standard, Air Interface for Fixed Broadband Wireless Access Systems*, IEEE Standard.

IEEE 802.16 Standard (2004). *IEEE 802.16 Standard, Air Interface for Fixed Broadband Wireless Access Systems*, IEEE 802.16 Standard, Air Interface for Fixed Broadband Wireless Access Systems.

ITU (2000a). *Technical Specifications for Third Generation Systems IMT-2000*, ITU Standard.

ITU (2005). ITU-T X.500 Directory Specifications, *ISO/IEC 9594-1*. Available from http://www.upgrade-cepis.org/issues/2005/4/up6-4Chadwick.pdf.

ITU (March 2000b). *The directory: Public-key and Attribute Certificate Frameworks*.

ITU-R WP8F (2006). *Additional Technical Details Supporting IP-OFDMA as an IMT-2000 Terrestrial Radio Interface*, Contribution ITU-R WP/1079.

Loughney, J. & Camarillo, G. (2004). *Authentication, Authorization, and Accounting Requirements for the Session Initiation Protocol (SIP)*, IETF RFC 3702.

Milikfish (2006). http://www.milkfish.org, Open SIP Express Router.

Myers, M., Ankney, R., Malpani, A., Galperin, S. & Adams, C. (1999). *X.509 Internet Public Key Infrastructure Online Certificate Status Protocol - OCSP*, RFC 2560.

NBS (1977). *Data Encryption Standard, FIPS-Pub.46. National Bureau of Standards, U.S. Department of Commerce, Washington D.C*, Data Encryption Standard, FIPS-Pub.46. National Bureau of Standards, U.S. Department of Commerce, Washington D.C.

NS-2 (1995). The Network Simulator - ns-2. Available from: `http://www.isi.edu/nsnam/ns/`.

OECD (October 1999). *OECD Communications Outlook 2007*, OECD Communications Outlook 2007.

OpenSER (2006). http://www.openser.org, Open SIP Express Router.

Paxson, V. (1997). End-to-end Internet packet dynamics, *SIGCOMM '97: Proceedings of the ACM SIGCOMM '97 conference on Applications, technologies, architectures, and protocols for computer communication*, ACM Press, New York, NY, USA, pp. 139–152.

Perkins, C. (August 2002). *IP Mobility Support for IPv4*, IETF RFC 3344.

Prijono, B. (2005). http://www.pjsip.org, PJSIP Project.

Prior, R. (2006). NS 2 SIP module. Available from: `http://www.ncc.up.pt/~rprior/ns/index-en.html`.

Project, A. (2005). *Ambient Network*, D 1.5 AN Framework Architecture.

Raman, B., Agarwal, S., Chen, Y., Caesar, M., Cui, W., Johansson, P., Lai, K., Lavian, T., Machiraju, S., Mao, Z. M., Porter, G., Roscoe, T., Seshadri, M., Shih, J. S., Sklower, K., Subramanian, L., Suzuki, T., Zhuang, S., Joseph, A. D., Katz, R. H. & Stoica, I. (2002). The SAHARA Model for Service Composition across Multiple Providers, *Pervasive '02: Proceedings of the First International Conference on Pervasive Computing*, Springer-Verlag, London, UK, pp. 1–14.

Rigney, C., Willens, S., Rubens, A. & Simpson, W. (2000). *Remote Authentication Dial In User Service (RADIUS)*, RFC 2865.

Rosenberg, J. & Schulzrinne, H. (2006). *Guidelines for Authors of Extensions of the Session Initiation Protocol (SIP)*, IETF RFC 4485.

Schulzrinne & Wedlund (2000). Application-Layer Mobility Using SIP, *SIGMOBILE Mob. Comput. Commun. Rev.* **4**(3): 47–57.

Schulzrinne, H. & Handley, M. (1999). *SIP: Session Initiation Protocol*, Request for Comments 2543, IETF.

Schulzrinne, H., Casner, S., Frederick, R. & Jacobson, V. (2003). *RTP: A Transport Protocol for Real-Time Applications*, IETF RFC 3550.

Sipsak (2002). http://sipsak.org, SipSak.

Soininen, J. (August 2003). *Transition Scenarios for 3GPP Networks*, IETF RFC 3574.

Stewart, R., Xie, Q., Morneault, K., Sharp, C., Schwarzbauer, H., Taylor, T., Rytina, I., Kalla, M., Zhang, L. & Paxson, V. (2000). *Stream Control Transmission Protocol*, IETF RFC 2960.

Wei, W., Banerjee, N., Basu, K. & Das, S. (2005). SIP-based vertical handoff between WWANs and WLANs, *Wireless Communications, IEEE [see also IEEE Personal Communications]* **12**(1): 66–72.

Wilhelm, U. G. & Butty, L. (1998). On the Problem of Trust in Mobile Agent Systems, *Symposium on Network and Distributed System Security*.

Wu, C. C. & Bertsekas, D. P. (n.d.). Admission control for wireless networks. Available from: `citeseer.ist.psu.edu/598372.html`.

Xiuchao, W. (2004). Simulate 802.11b channel within ns2, *Technical Report*.

Zeilenga, K. (2006). *Lighweight Directory Access Protocol (LDAP): Technical Specification Road Map*, RFC 4510.