KING FAHD UNIVERSITY OF PETROLEUM AND MINERALS

DEPARTMENT OF COMPUTER ENGINEERING

COE-485

Senior Design Project

Investigation of WLAN-Cellular Integration Architecture for Cellular Operators and Deployment Issues at King Fahd International Airport

Prepared for
Dr. Ashraf H. Mahmoud

Ejaz Rahman
ID 215485

Junaid Jaffar
ID 215637

January 24, 2006

Acknowledgements

We would like to express our sincerest gratitude to Dr. Ashraf H. Mahmoud for his supervision and assistance throughout our senior design project and for providing invaluable insights and directions whenever needed. His guidance in this exciting field helped us a lot and we are indebted to him in this regard as well as in bringing this project to completion.

We also thank Dr. Abdelhafid Bouhraoua, for his administrative support and assistance in carrying out and realizing this project. We would also like to thank our family and friends for their support throughout the semester.

Table of Contents

List of Figures

List of Tables

Acronyms

3G	3rd Generations
3G-SGSN	3rd Generation - Serving GPRS Support Node
3G-GGSN	3rd Generation - Gateway GPRS Support Node
3GPP	3rd Generation Partnership Project
AAA	Authentication, Authorization and Accounting
AP	Access Point
APE	Access Point Estimator
AR	Access Router
ASCONF	Address Configuration
AUC	Authentication Center
AZR	Access Zone Router
BSSID	Basic Service Set Identifier
CC&BS	Customer Care and Billing Services
CG	Charging Gateway
CN	Correspondent Node
CoA	Care of Address
DAR	Dynamic Address Reconfiguration
DHCP	Dynamic Host Configuration Protocol
DNS	Domain Name Server
EAP	Extensible Authentication Protocol
EAP-AKA	EAP for UMTS Authentication and Key Agreement
EAPoL	EAP over LAN
EAPoW	EAP over Wireless
EAP-SIM	EAP Method for GSM Subscriber Identity
EAP-TTLS	EAP-Tunneled Transport Layer Security
EDGE	Enhanced Data Rates for GSM Evolution
EIR	Equipment Identification Register
ETSI	European Telecommunications Standards Institute
FA	Foreign Agent

GCoA	Global Care-of Address
GERAN	GSM/EDGE Radio Access Network
GGSN	Gateway GPRS Support Node
GMSC	Gateway MSC
GPRS	General Packet Radio Service
GSM	Global System for Mobile Communications
HA	Home Agent
HAWAII	Handoff Aware Wireless Access Internet Infrastructure
HCRAS	Hybrid Coupling with Radio Access System
HLR	Home Location Register
IAPP	Inter-Access Point Protocol
IDMP	Intra-Domain Mobility Management Protocol
IEEE	Institute of Electrical and Electronic Engineers
I-GGSN	Intelligent Gateway GPRS Serving Node
IMS	IP Multimedia Subsystem
IMSI	International Mobile Subscriber Identity
ITP-MAP	IP Transfer Point-Mobile Application Part
ITU	International Telecommunications Union
IWU	InterWorking Unit
KFIA	King Fahd International Airport
KWAA	KFIA WLAN-Cellular Access Architecture
KWIF	KFIA WLAN- Cellular Interworking Framework
LCoA	Local Care-of Address
LDAP	Lightweight Directory Access Protocol
LEAP	Light Extensible Authentication
MA	Mobility Agent
MDS	Multi-access Data server
MIP	Mobile IP
MMS	Multimedia Message Service
MN	Mobile Node
MSC	Mobile Switching Center
mSCTP	Mobile Stream Control Transmission Protocol
OCS	Online Charging System

PDG	Packet Data Gateway
PLMN	Public Land Mobile Network
PPP	Point-to-Point Protocol
PWLAN	Public Wireless Local Area Network
QoS	Quality of Service
RADIUS	Remote Authentication Dial-In User Service
RNC	Radio Network Controller
RNS	Radio Network Sub-systems
SA	Subnet Agent
SCTP	Stream Control Transmission Protocol
SESM	Subscriber Edge and Services Manager
SGSN	Serving GPRS Support Node
SIM	Subscriber Identification Module
SIP	Session Initiation Protocol
SS7	Signaling System 7
SSG	Service Selective Gateway
STC	Saudi Telecommunications Company
TACACS+	Terminal Access Controller Access Control System
TCWA	Tight Coupling with Wireless Access
UCN	UMTS Core Network
UE	User Equipment
UMTS	Universal Mobile Telecommunications System
UTRAN	UMTS Terrestrial Radio Access Network
VLR	Visitor Location Register
VPLMN	Visited PLMN
W-APN	WLAN Access Point Name
WARG	WLAN Access Router Gateway
WCDMA	Wide-band Code Division Multiple Access
WLAN	Wireless Local Area Network

1 Introduction

It is increasingly evident that the growth of wireless local access networks (WLANs) based on 802.11x standards like Wi-Fi will soon be massive and widespread. Enterprises and end users are enthusiastically committing resources into WLAN deployment to benefit from the technology. Thus the natural trend today is to utilize the high-bandwidth WLANs in hot spots and switch to UMTS networks when the coverage of WLAN is not available or the network condition is not good enough. Thus the complementary characteristics of Universal Mobile Telecommunications System (UMTS) and Wireless Local Area Network (WLAN) make them ideal for interworking. UMTS provides a low-bandwidth circuit- and packet-switched service to users with relatively high mobility in large areas whereas WLAN provides a high-bandwidth packet-switched service to users with low mobility in smaller areas. The WLAN therefore complements UMTS on the packet-switched services. Hence, this senior project undertaking endeavored to develop and realize a WLAN-cellular interworking framework for operators with an initial deployment of a PWLAN at King Fahd International airport as its starting point in its series of PWLAN deployments aimed at targeting this vast and potentially rewarding market.

This report begins with section 2 wherein background information related to developing the interworking architecture is discussed. This includes an overview of WLAN and UMTS systems as well as a discussion of the various interworking scenarios and handovers. A brief discussion of AAA protocols is also included. This is followed by section 3 which provides a comprehensive survey of various WLAN-cellular interworking solutions. This section is divided into vendor solutions which focus on interworking solutions available in the market and academic solutions, which are solutions being proposed and usually still at the research stage but have promise in the future. Section 4 provides an all inclusive coverage of mobility management protocols that are essential for developing any interworking architecture that can provide continuity of service in case of handovers. It basically focuses on MIP, mSCTP and SIP.

The entire crux of the project is established in section 5. It begins with an overview of the architecture being proposed for establishing a WLAN-cellular interworking architecture for cellular operators. It covers a number of aspects such as the various mobility management protocols that can be incorporated. A complete access architecture that can provide a robust and secure mechanism to access the WLAN service is discussed. Various deployment issues relate PWLAN are also addressed. Finally this report is concluded with a look at some promising future directions this research is likely to take.

2 Background

In this section, a brief overview of the interworking technologies i.e. WLAN and UMTS will be presented. A discussion of some common terms used in this interworking field such as interworking scenarios and handover types is also provided. Moreover, to provide a detailed discussion of the AAA features later in the report, some background information of AAA protocols such as RADIUS, DIAMETER and so on is also included.

2.1 Wireless Local Area Network (WLAN)

A WLAN is a type of local area network that uses electromagnetic waves (radio and infrared) instead of wired connections to send and receive information between the communicating entities [34]. These entities or mobile hosts communicate with the wired backbone network via one or more access points as shown in the following figure. The mobile hosts contain WLAN adapters which allow them to access the WLAN.

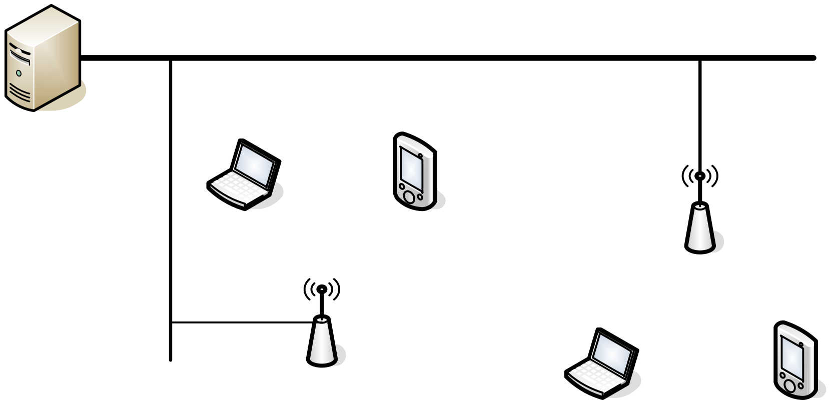


Fig. 1. Basic WLAN architecture

There are two main types of WLAN configurations: Peer-to-peer or ad-hoc mode and infrastructure mode. In ad-hoc mode, the mobile hosts communicate with each other directly without involving access points. When two or more hosts are within range of each other they can interconnect and share information. The benefit of this approach is that it requires no administration.

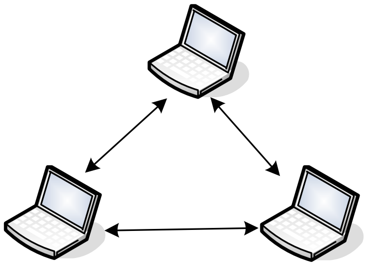


Fig. 2. Ad-hoc mode configuration

The infrastructure mode consists of multiple access points connected to the wired backbone network. This allows the WLAN users to share network resources efficiently. The access points act as central communicating stations and balance the network traffic.

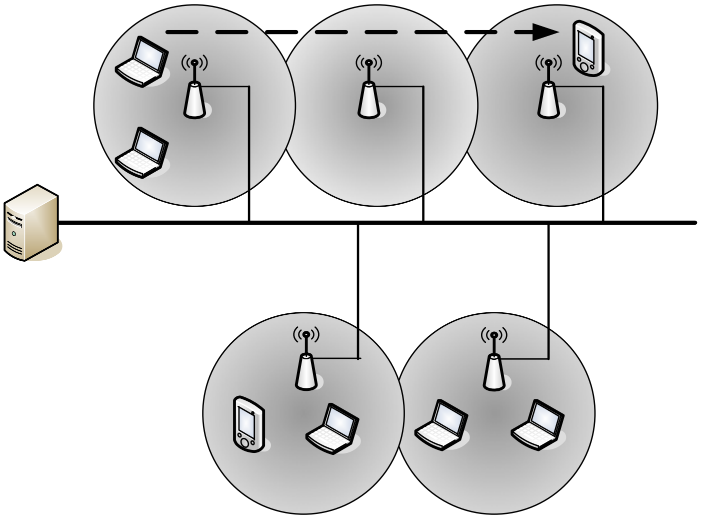


Fig. 3. Infrastructure mode configuration

2.1.1 Inter-Access Point Protocol (IAPP)

When a mobile host moves out the coverage area of one access point (AP) and into that of another, it has to associate with the new access point. After this, the wired network has to be informed of the new association. This inter-AP communication was not standardized and different vendors had their own way of implementing a solution [35]. However, the IAPP (IEEE 802.11f) has now been developed and allows the transfer of information from one AP to another to carry out fast handovers between AP's of different vendors.

The IAPP message exchange during handover is initiated by the re-associate message, which carries the Basic Service Set Identifier (BSSID) of the old AP (message 1 in figure 4). In order to communicate with the old AP, the new AP needs its IP address. The mappings between BSSID and IP addresses can either be stored inside all AP's or obtained from a Remote Authentication Dial-In User Service (RADIUS) server (messages 2 and 3). If there is a need for encryption, security blocks are exchanged before the exchange of data between AP's (messages 4 and 5). The two most significant messages of the IAPP are the MOVE-notify (message 6), issued by the new AP to inform the old AP that the mobile host has moved to its own area, and the MOVE-response (message 7), issued by the old AP, containing the context block which contains the information to be exchanged [36]. It can consist of a variable number of Information Elements (IE's) of the form (Element ID, Length, and Information). In this way, every IE can contain variable length information, whose type is specified by the Element ID. Processing of the information transferred inside the IE's depends on the functionality of the AP's [36].

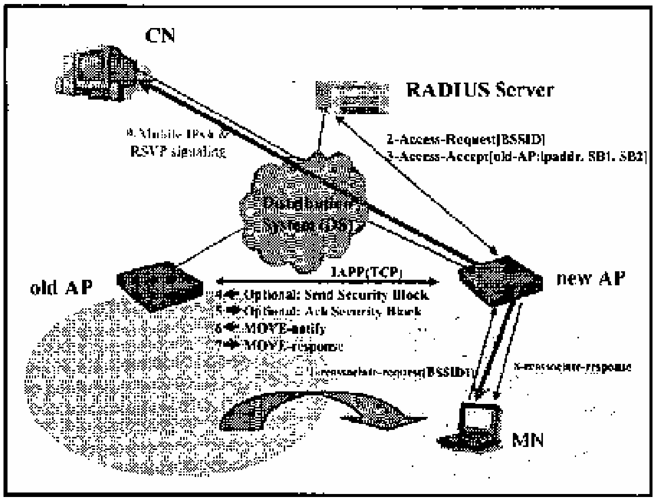


Fig. 4. IAPP signaling [36]

2.2 Universal Mobile Telecommunication System (UMTS)

UMTS represents an evolution in terms of services and data speeds from today's "second generation" mobile networks. As a key member of the "global family" of third generation (3G) mobile technologies identified by the ITU, UMTS is the natural evolutionary choice for operators of GSM networks, currently representing a large customer base mostly in Asia and Europe.

UMTS is a third generation (3G) mobile communications system that provides a range of broadband services to the world of wireless and mobile communications. UMTS delivers low-cost, mobile communications at data rates of up to 2 Mbps. It preserves the global roaming capability of second generation GSM/GPRS networks and provides new enhanced capabilities.

It is designed to deliver pictures, graphics, video communications, and other multimedia information, as well as voice and data, to mobile wireless subscribers. The UMTS takes a phased approach toward an all-IP network by extending second generation (2G) GSM/GPRS networks and using Wide-band Code Division Multiple Access (WCDMA) technology. Handover capability between the UMTS and GSM is supported. The GPRS is the convergence point between the 2G technologies and the packet-switched domain of the 3G UMTS.

2.2.1 Network architecture

The UMTS network architecture (Release 99) consists of three domains: The User Equipment (UE) domain, the UMTS Terrestrial Radio Access Network (UTRAN) domain and the Core Network (CN) domain as shown in the figure below. The UE domain represents the equipment used by the user to access UMTS services while the UTRAN domain and the CN domain, together known as the infrastructure domain, consist of the physical nodes which perform the various functions required to terminate the radio interface and to support the telecommunication services requirements of the user.

The three domains are further described in the following sections.

User Equipment (UE) DOMAIN:

The UE domain encompasses a variety of equipment types with different levels of functionality such as cellular phones, PDAs, laptops etc. These equipment types are typically referred to as user equipment. The UE domain consists of two parts: The UMTS Subscriber Identity Module (USIM) and the Mobile Equipment (ME). The USIM is a smartcard that contains user-specific information and the authentication keys that authenticates a user’s access to a network. The USIM is physically incorporated into a SIM card and linked to the

ME over an electrical interface at reference point Cu. The ME is a radio terminal used for radio communication with the UTRAN domain over the Uu radio interface. [23][24][25].

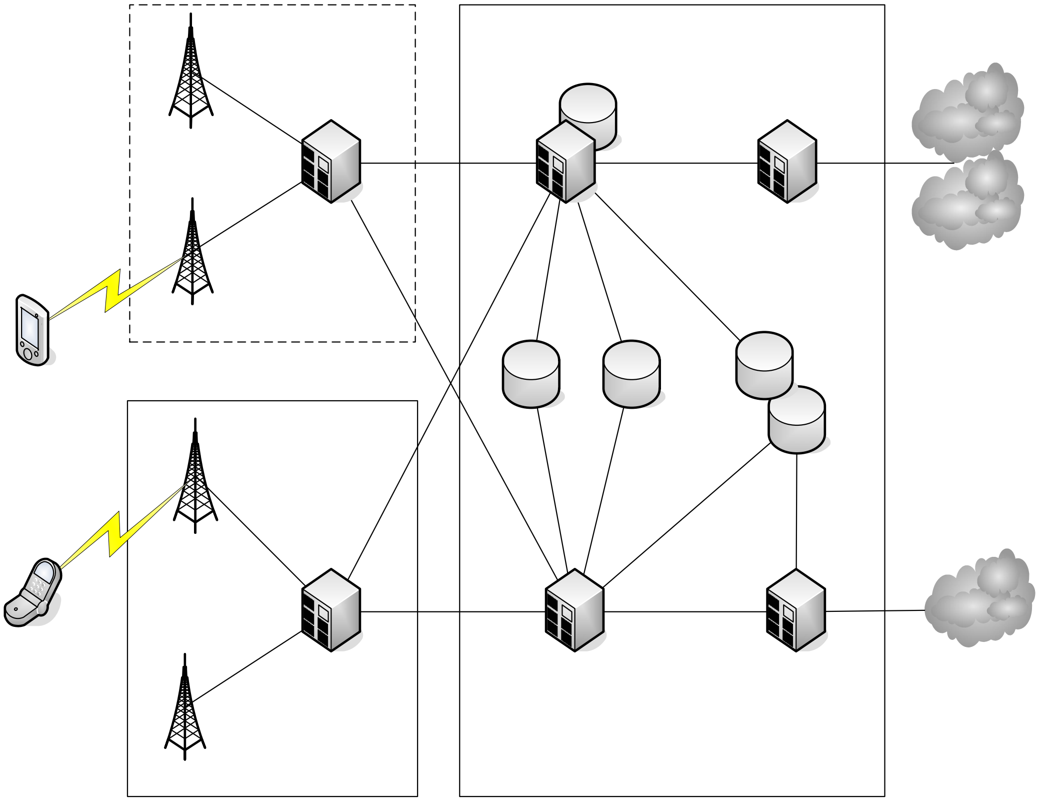


Fig. 5. UMTS Architecture

UMTS Terrestrial Radio Access Network (UTRAN) DOMAIN:

The UTRAN domain handles all radio-related functionality. It consists of one or more Radio Network Sub-systems (RNS) where each RNS consists of one or more Node Bs and one Radio Network Controller (RNC). The Node B, also known as a Base Station and equivalent to the Base Transceiver Station (BTS) from GSM, converts the signals of the radio interface into a data stream and forwards it to the RNC over the Iub interface. In the opposite direction, it prepares incoming data from the RNC for transport over the radio interface. The area covered by a Node B is called a cell.

BTS

The RNC is the central node in the UTRAN and equivalent to the Base Station Controller (BSC) from GSM. It controls one or more Node Bs over the Iub interface and is responsible for the management of all the radio resources in the UTRAN. The RNC interfaces the CN

BTS

Um

MN

domain over the Iu interface. If there are more than one RNC, they can be interconnected via an Iur interface. [23][24][25].

Core Network (CN) DOMAIN:

The CN domain is responsible for switching and routing calls and data connections between the UTRAN domain and external packet and circuit switched networks. It is divided into a Packet Switched network (PS), a Circuit Switched network (CS) and a Home Location Register (HLR).

The PS network consists of a Serving GPRS Support Node (SGSN) and a Gateway GPRS Support Node (GGSN). The SGSN is responsible for routing packets inside the PS as well as handling authentication and encryption for the users. The GGSN serves as the gateway towards external packet switched networks like the Internet, Local Area Networks (LANs), Wide Area Networks (WANs), General Packet Radio Service (GPRS) networks, Asynchronous Transfer Mode (ATM) networks, Frame Relay networks, X.25 networks etc., and thus completes the routing function of the SGSN.

The CS network consists of a Mobile Services Switching Centre (MSC)/Visitor Location Register (VLR) and a Gateway MSC (GMSC). The MSC/VLR serves as a switch and database. The MSC part is responsible for all signaling required for setting up, terminating, and maintaining connections, and mobile radio functions such as call rerouting, as well as the allocation/de-allocation of radio channels, i.e. the switching function. The VLR part is controlled by the MSC part and is used to manage users that are roaming in the area of the associated MSC. It stores information transmitted by the responsible HLR for mobile users operating in the area under its control, i.e. the database function.

The GMSC serves similar to the GGSN as the gateway towards external circuit switched networks like other Public Land Mobile Networks (PLMNs), Public Switched Telephone Networks (PSTNs) and Integrated Service Digital Networks (ISDNs) etc.

The HLR is a database located in the user’s home system that stores all important information relevant to the user, e.g. telephone number, subscription basis, authentication key, forbidden roaming areas, supplementary service information etc. The HLR also stores the UE location for the purpose of routing incoming transactions to the UE. [23][24][25].

2.3 Coupling

The architecture for interworking between WLANs and cellular networks is classified into three different couplings: tight, loose and open. In tight coupling the Iu interface can be modified to provide the connectivity between the UMTS core and the WLAN via its radio controller. This requires the standardization of the interface and requires the use of SIM cards in the mobile terminals. In loose coupling the WLAN connects to the HLR entity in the UMTS core network via its AAA server. This solution requires less standardization and does not require specific network access equipment. With open coupling, there is no standardization of the interface and the WLAN is linked to the UMTS core network only at the customer care and billing system.

2.4 Interworking Scenarios

The interworking requirements are classified into the following six scenarios [37]

Scenario 1: Common billing and customer care

This form of interworking provides only a common bill and customer care to the subscriber and has no other interworking between the WLAN and GPRS network. Hence, it does not need any other standardization features.

Scenario 2: 3GPP system based access control and charging

This requires AAA for the subscribers in the WLAN to be based on the same AAA procedures utilized in the GPRS system. So users in a hot spot can use the same SIM card for authentication that they would use for GPRS. The authorization is taken care of by GPRS on the basis of subscription information. IP connectivity is maintained for GPRS users via the WLAN.

Scenario 3: Access to 3GPP GPRS based services

This scenario allows users in a WLAN area to access GPRS based services i.e. services provided over the GPRS network. These include IP multimedia based services, location and presence based services, instant messaging. For example, if a cellular operator provides WAP service to the subscribers, then subscribers in a WLAN area should also be able to access this WAP service.

Scenario 4: Service continuity

This scenario builds on scenario 3 by providing service continuity over the GPRS and WLAN environments in addition to providing access to 3GPP GPRS based services. Considering the same example above, when a subscriber using WAP in a GPRS area moves to a WLAN area, he/she should be able to continue using WAP. However, the service continuity requirements are not very strict in this scenario. So it is possible that GPRS services requiring tight delay performance would be terminated when a user moves to a WLAN area.

Scenario 5: Seamless service

This scenario enhances scenario 4 by providing seamless service continuity between GPRS and WLAN systems. Therefore, subscribers will be able to use GPRS based services over the WLAN without any noticeable changes in performance.

Scenario 6: Access to 3GPP circuit switched services

This scenario enables the provision of circuit switched services (e.g. normal voice calls) from the WLAN environment. These services require seamless mobility.

Table 1 summarizes the various scenarios with regard to interworking WLAN and cellular technologies focusing on the features available in each scenario.

Scenarios	**Features**
1	Common billing and customer care
2	Scenario 1 + 3G-based access control and charging
3	Scenario 2 + access to 3G packet-switched based services
4	Scenario 3 + service continuity
5	Scenario 4 + seamless service continuity
6	Scenario 5 + access to 3G circuit-switched based services with seamless mobility

Table 1. Interworking scenarios and their features

2.5 Handovers

A handover (also known as handoff) is the process by which a mobile terminal changes its point of attachment to the network. There are two main types of handoffs: vertical and horizontal. When a handoff occurs between two different types of networks it is referred to as a vertical handoff. For example, switching from a GPRS/UMTS network to a WLAN or vice versa is done via a vertical handoff. When a handoff occurs within the same network it is called a horizontal handoff. For example, in the GSM network, a mobile terminal switches between two base stations through a horizontal handoff.

2.6 Authentication Authorization and Accounting

An AAA server is a server program that handles user requests for access to computer resources and, for an enterprise, provides authentication, authorization, and accounting (AAA) services. The AAA server typically interacts with network access and gateway servers and with databases and directories containing user information. The current standard by which devices or applications communicate with an AAA server is the Remote Authentication Dial-In User Service (RADIUS) protocol, although there are many other AAA protocols that have been in use. The DIAMETER protocol is a recent AAA protocol that ahs been designed to replace RADIUS. In this section a brief overview of these protocols will be provided.

2.6.1 RADIUS (Remote Authentication Dial In User Service)

RADIUS Evolved from work by Steve Willens and Carl Rigney at Livingston Enterprises in 1992 and rapidly became the industry standard with broad support across nearly all vendors of networking equipment with the support of the IETF. RADIUS is a UDP based protocol suitable for high volume service control applications such as regulation of dial in or VPN services.

With the emergence of 802.1x port security for wired and wireless LANs, RADIUS has recently seen even greater usage. A RADIUS client (typically an access server such as a dial-up server, VPN server, or wireless access point) sends user credentials and connection parameter information in the form of a RADIUS message to a RADIUS server.

The RADIUS server authenticates and authorizes the RADIUS client request, and sends back a RADIUS message response. RADIUS clients also send RADIUS accounting messages to RADIUS servers. Additionally, the RADIUS standards support the use of RADIUS proxies. A RADIUS proxy is a computer that forwards RADIUS messages between RADIUS clients, RADIUS servers, and other RADIUS proxies. RADIUS messages are never sent between the access client and the access server. [17]

RADIUS messages are sent as User Datagram Protocol (UDP) messages. UDP port 1812 is used for RADIUS authentication messages and UDP port 1813 is used for RADIUS accounting messages. Some access servers might use UDP port 1645 for RADIUS authentication messages and UDP port 1646 for RADIUS accounting messages. Only one RADIUS message is included in the UDP payload of a RADIUS packet. [17]

RFC 2865 define the following RADIUS message types:

- Access-Request. Sent by a RADIUS client to request authentication and authorization for a network access connection attempt.
- Access-Accept. Sent by a RADIUS server in response to an Access-Request message. This message informs the RADIUS client that the connection attempt is authenticated and authorized.

- Access-Reject. Sent by a RADIUS server in response to an Access-Request message. This message informs the RADIUS client that the connection attempt is rejected. A RADIUS server sends this message if either the credentials are not authentic or the connection attempt is not authorized.

- Access-Challenge. Sent by a RADIUS server in response to an Access-Request message. This message is a challenge to the RADIUS client that requires a response.

- Accounting-Request. Sent by a RADIUS client to specify accounting information for a connection that was accepted.

- Accounting-Response. Sent by the RADIUS server in response to the Accounting-Request message. This message acknowledges the successful receipt and processing of the Accounting-Request message.

A RADIUS message consists of a RADIUS header and RADIUS attributes. Each RADIUS attribute specifies a piece of information about the connection attempt. For example, there are RADIUS attributes for the user name, the user password, the type of service requested by the user, and the IP address of the access server. RADIUS attributes are used to convey information between RADIUS clients, RADIUS proxies, and RADIUS servers. For example, the list of attributes in the Access-Request message includes information about the user credentials and the parameters of the connection attempt. In contrast, the list of attributes in the Access-Accept message includes information about the type of connection that can be made, connection constraints, and any vendor-specific attributes (VSAs).[17]

To provide security for RADIUS messages, the RADIUS client and the RADIUS server are configured with a common shared secret. The shared secret is used to secure RADIUS traffic and is commonly configured as a text string on both the RADIUS client and server. [17]

2.6.2 TACACS+ (Terminal Access Controller Access Control System)

Invented at around the same time as RADIUS, the protocol was developed by Cisco to service similar needs. TACACS was proposed to the IETF as a standard but remains only in draft RFC form to this day. Whilst supported by a number of device and server vendors, TACACS+ is most pervasively used by Cisco Systems. TACACS+ is a TCP based protocol that separates all three stages of the AAA process and provides an excellent foundation for the construction of sophisticated device administration regimes. In addition to the normal "session" accounting, the TACACS+ protocol is often used to log the commands entered by administrator users when configuring a network device which can be invaluable for isolating mis-configuration problems. [18]

2.6.3 DIAMETER

Diameter is an AAA (authentication, authorization and accounting) protocol for applications such as network access or IP mobility. The basic concept is to provide a base protocol that can be extended in order to provide AAA services to new access technologies [19].

Attribute-Value Pairs: AVPs are a method of encapsulating information relevant to the Diameter message. The Diameter protocol defines things in terms of attributes. Each attribute may take on one of a set of values. When a Diameter packet is exchanged among clients and servers, one or more attributes and values are sent pair wise from the client to the server [20].

Protocols: A Diameter client generates Diameter messages to request authentication, authorization, and accounting services for the user. A Diameter agent is a node that does not authenticate and/or authorize messages locally (proxies). A Diameter server performs authentication and/or authorization of the user. A Diameter node may act as an agent for certain requests while acting as a server for others. Any node can initiate a request, in that sense Diameter is a peer to peer protocol.

Diameter uses both TCP and SCTP transport protocols. Clients must support either TCP or SCTP, while servers must support both.

Diameter clients, such as Network Access Servers (NAS) and Mobility Agents must support IP Security, and may support TLS. Diameter servers must support TLS and IPsec. The Diameter protocol must not be used without any security mechanism [20].

DIAMETER message: A Diameter message consists of a fixed-length header followed by a variable number of AVPs. There are two types of messages, Requests and Answers. There are few circumstances where a request is silently discarded, and therefore the originator of a request will receive an answer. Every answer message carries a Result-Code AVP indicating whether a particular request was completed successfully or whether an error occurred [21].

Below is a drawing of a sequence diagram when a user accesses the network through the Network Access Server and disconnects itself.

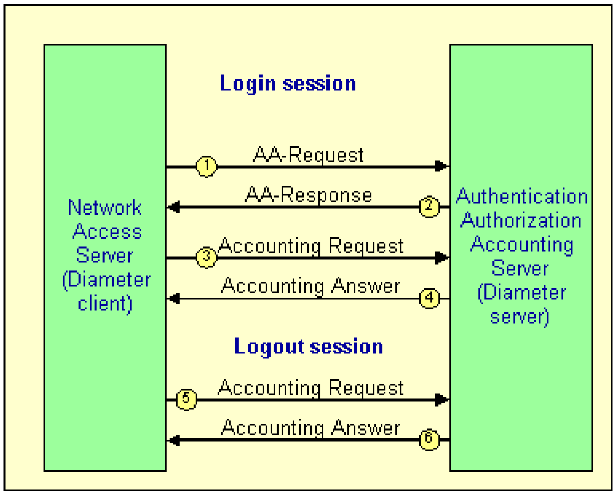


Fig. 6. DIAMETER's Message Flow [22]

2.6.4 Advantages of Diameter over RADIUS

Diameter provides better transport for packets as it runs over a reliable transport layer protocol (TCP or SCTP). These protocols allow lost packets to be retransmitted at each hop and also adapt to network congestion. In the event that a transport failure is detected with a peer, it is necessary for all pending request messages to be forwarded to an alternate agent, if possible. This is commonly referred to as failover. Diameter maintains a persistent connection with an application-level message to support failover. Diameter has proxy agents that route Diameter messages. Following a failover these proxies automatically retransmit pending request messages. Diameter also provides better session control as session management is independent of accounting. Diameter is also more secure as it provides hop-by-hop security using IPsec or TLS. Also, end-to-end security protects the integrity and the confidentiality of A-V pairs through intermediate proxies [21].

3 WLAN-Cellular Interworking Solutions Survey

In this section we will discuss in detail various PWLAN as well as interworking WLAN-UMTS solution available. These include solutions being offered in the market by vendors as well as academic solutions proposed in literature.

3.1 Vendor Solutions

In this section, a couple of the most recent vendor solutions available in the market for the PWLAN will be discussed.

3.1.1 Cisco PWLAN solution

Scenario: This solution can be most closely classified as a scenario 4 solution as service connectivity is provided over various networks.

Coupling: The Cisco PWLAN solution can be categorized as loose coupling based on the following reason -

- Cellular access gateway provides WLAN user with SIM based authentication. RADIUS based authentication is used
- Mobile IP handovers occur between GGSN and Access Zone router

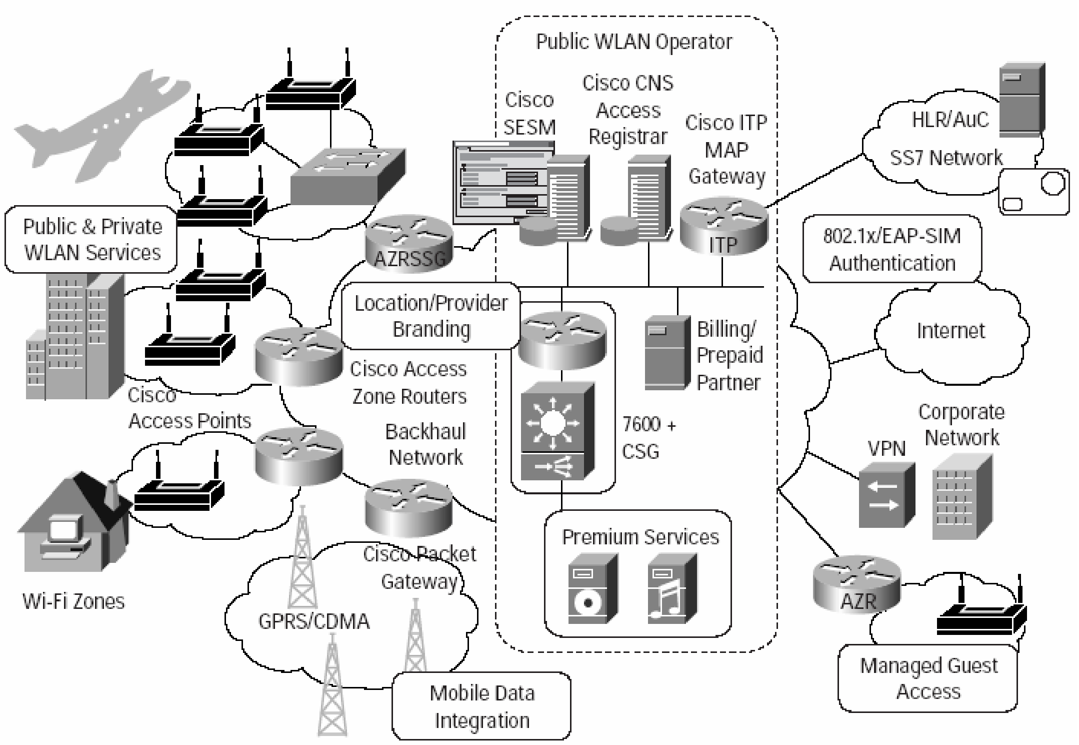


Fig. 7. Cisco PWLAN solution Architecture [1]

PWLAN solution architecture

The main components of the Cisco PWLAN solution contain the following:

1. Access Points
2. Access Zone Routers(AZR)
3. Service Selective Gateway technology
4. CNS Subscriber Edge and Services Manager (SESM)
5. CNS Access Registrar
6. IP Transfer Point-Mobile Application Part (ITP-MAP) gateway

Each of these components have detailed functionalities and are important parts that combine together to form the Cisco PWLAN Mobile Exchange.

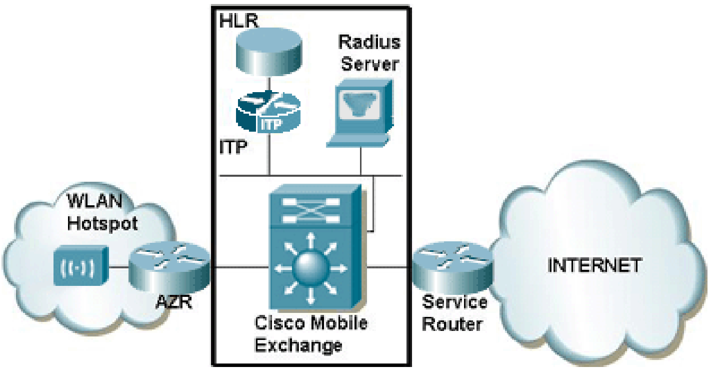


Fig. 8. Simplified PWLAN Mobile Exchange [5]

AAA Server: Cisco CNS Access Registrar

Cisco CNS Access Registrar supports deployment of access services by centralizing AAA information and simplifying provisioning and management with standards-based **RADIUS** protocol and proxy RADIUS server.

The Cisco AAA server provides a range of authentication types [1] including:

- LEAP- A Cisco propriety authentication protocol that stands for Light Extensible Authentication Protocol
- EAP-SIM- Extensible Authentication Protocol Method for GSM Subscriber Identity which is an Extensible Authentication Protocol (EAP) mechanism for authentication

and session key distribution using the Global System for Mobile Communications (GSM) Subscriber Identity Module (SIM).

Cisco CNS Access Registrar provides the following key benefits [2]:

• It supports multiple services (dial, DSL, voice, and wireless) with a single AAA platform.

• Its high-performance architecture reduces the number of AAA servers that must be deployed.

• It reduces operational costs and speeds service rollout by supporting integration with provisioning, billing, and other service management components via directory support and scriptable configuration interfaces and provides external interfacing with ODBC and LDAP protocols.

• Cisco CNS Access Registrar offers the extensibility and flexibility of non-supported freeware RADIUS software, while being supported by Cisco.

• Cisco CNS Access Registrar can be integrated with provisioning, billing, and other management and operational systems of the customer's choice.

Services and Features

Following are a list of the various features of the Cisco PWLAN solution and their discussion. [1]

• **Services enablement—**The Cisco PWLAN solution allows operators to enable the introduction of value-added services with incremental revenue streams. Examples of such services include anything from music, movies, sports, gaming, or ring tones for the consumer, to business-class, WLAN-optimized services for voice over IP (VoIP) for the enterprise or hosted VPN services that offer security for all users.

• **Comprehensive branding support—** The Cisco PWLAN solution enables operators to provide co-branded, customized portals for promoting their business in prime hot-spot locations such as hotel chains, convention centers, and airports. These portals can also feature localized information such as area guides and points of interest.

• **Flexible service billing options—**The Cisco PWLAN solution supports flexible billing models, including postpaid, prepaid (time or volume), tariff, and billing based on subscription content. Additional features enable the network to quickly detect when a user has left the service area without logging out and automatically close their session. In doing so, billing accuracy is preserved and network resources are freed.

• **Ready-to-use support for ease of use—**Key features developed for the Cisco PWLAN solution allow clients with static host client configurations to access the service without making changes to their laptop or mobile device. Cisco supports clients with static IP and Domain Name System (DNS) entries in addition to supporting clients with static HTTP proxy configurations.

• **Authentication transparency—**The Cisco SSG access control platform can proxy EAP authentication messages from hot-spot access points and automatically create user sessions upon successful EAP authentication, thereby eliminating the need for "double authentication," first at Layer 2 with 802.1x/EAP and then at Layer 3 through the Web portal. This feature allows an operator to take advantage of the Cisco SSG for centralized accounting record generation for both 802.1x/EAP and Web-authenticated users.

Integration with 3G systems:

The Cisco Mobile Exchange framework in conjunction with the Cisco Mobile Wireless Home Agent (based on the Internet Engineering Task Force (IETF) Mobile IP standard (RFC 3344)) identifies a host device by a single IP address even if the device moves its physical point of attachment from one network to another. [4]

Subscribers with mobile devices can roam to another network without restarting applications or terminating and reestablishing a connection. This allows seamless roaming over different access technologies and provider networks, including mobile data networks based on 3G. Thus, Cisco's mobile IP technology in conjunction with other technologies offers standards-based solutions for seamless roaming. [3]

3.1.2 Alcatel PWLAN solution

Scenario: The Alcatel PWLAN solution closely resembles scenario 4 because session continuity is provided by this solution using Mobile IP protocol. However, provision is provided to choose different scenario levels based on customer preference.

Coupling: The Alcatel PWLAN solution can be categorized as loose coupling due to the existence of a control connection with an AAA server and a direct connection with the GGSN.

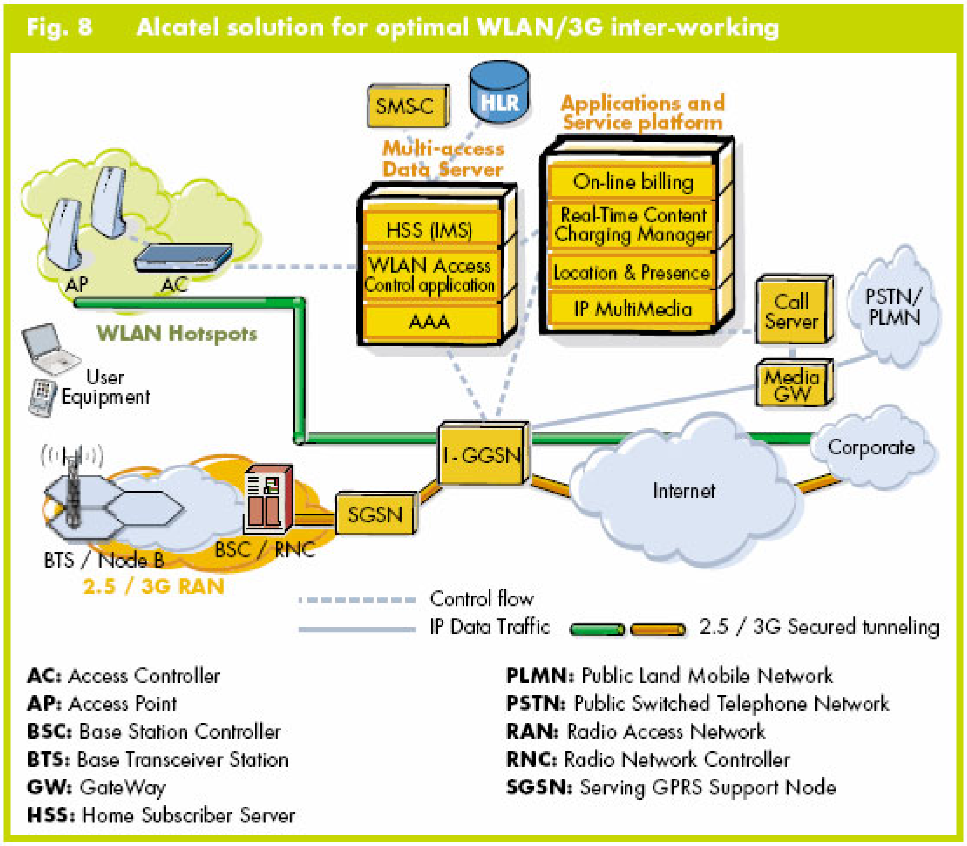


Fig. 8 Alcatel solution for optimal WLAN/3G inter-working
SMS-C
HLR
Applications and Service platform
Multi-access Data Server
On-line billing
Real-Time Content Charging Manager
Location & Presence
IP MultiMedia
HSS (IMS)
WLAN Access Control application
AAA
AP
AC
WLAN Hotspots
User Equipment
Call Server
PSTN/ PLMN
Media GW
I-GGSN
Corporate
Internet
SGSN
BTS / Node B
BSC / RNC
2.5 / 3G RAN
Control flow
IP Data Traffic
2.5 / 3G Secured tunneling
AC: Access Controller
AP: Access Point
BSC: Base Station Controller
BTS: Base Transceiver Station
GW: GateWay
HSS: Home Subscriber Server
PLMN: Public Land Mobile Network
PSTN: Public Switched Telephone Network
RAN: Radio Access Network
RNC: Radio Network Controller
SGSN: Serving GPRS Support Node

Fig. 9. Alcatel WLAN/3G interworking solution [7]

PWLAN solution architecture

The main components of the Alcatel PWLAN solution contain the following:

1. Access Points
2. Access Controllers
3. Mobile network equipment that includes
    a. Multi-access Data server(MDS)
    b. Intelligent Gateway GPRS Serving Node(I-GGSN)
    c. Real-time content charging suite

4. Revenue generating applications such as mobile office, video e.t.c.

This solution allows operators to both interface with existing third party hotspot equipment as well as deploy a complete end-to-end solution from the hotspots to the mobile network core.

AAA Server: Multi-access Data Server (MDS)

Roaming is supported by standard AAA mechanisms. The AAA server of the visited network forwards the authentication request to the home mobile operator AAA server (located in the multi access data server). This request can be conveyed through a chain of AAA proxies before it reaches the home network. The MDS authenticates the user and accesses the HLR for cellular-grade authentication. If authentication in the home network is successful, the AAA in the WLAN is instructed to enable the customer session. The AAA server in the visited network then forwards accounting information to the home network. [7]
By relying on IP roaming, as specified by the Third Generation Partnership Project (3GPP) standards for mobile networks, mobile operators can support roaming not only with other mobile operators, but also with Wireless ISP (WISP) and fixed operators.

Authentication types:

Alcatel provides secure WLAN access to the user. There are two possible authentication schemes possible based on the capabilities of the user device and profile.
IEEE 802.1x (Extensible Authentication Protocol; EAP) authentication: Reuses the existing mobile Subscriber Identity Module (SIM) card mechanisms. In this case, IEEE 802.1x supplies credentials to the Home Location Register (HLR) through the Authentication, Authorization and Accounting (AAA) server(s) and the Alcatel MDS in a fully transparent mode for the user. In this case, the user has subscribed to the WLAN access and his or her equipment includes a SIM card. It is also known as EAP-SIM authentication. [7]
Secure web-based authentication: Uses a user name/password pair submitted to the home AAA server. This could be a One Time Password (OTP) delivered on a scratch card or via the Short Message Service (SMS) to the user’s cell phone. [7]

Services and Features

The Alcatel PWLAN solution has a number of important features some which are enumerated below [7]:

- Means of associating the user's identity with his or her traffic activity over IP. This will allow services such as:
    - Enhanced real-time charging.
    - Optimal and innovative support for MMS over WLAN.
    - Full connectivity capability (e.g. QoS) so that packet-switched services (e.g. MMS and IMS) and applications can be optimally supported over heterogeneous access networks, such as WLAN, GPRS and UMTS.
- A Convenient connection kit, allowing:
    - seamless distribution and installation in most types of user equipment (e.g. PDAs, laptops);
    - Simple customization (e.g. user profile, operator preferred choices).
- Friendly graphical user interface:
    - simple to use thanks to an intelligent software client;
    - automatic identification of available networks (WLAN, GPRS) using continuous network scanning;
    - hotspot access decision logic aimed at selecting the preferred network in terms of throughput, billing scheme, etc;
    - Automatic triggering of the appropriate access authentication method, depending on the WLAN access technology.
- Support for macro-mobility to maintain an established application session while switching between access networks (e.g. WLAN <-> GPRS).
- Open software to simplify the integration of services and application clients, for instance, the Multimedia Messaging Service (MMS) or services on top of the IP Multimedia Subsystem (IMS).

Integration with 3G

The Alcatel PWLAN solution provides extensive support for integration with GSM/UMTS and WLAN architecture. To allow the operator to make optimal use of this integration technology, adequate interworking mechanisms have been provided between GPRS/UMTS and WLAN networks. Three levels of interworking targeted by Alcatel are enumerated below [8]:

- *Common access control, through the use of (U) SIM based authentication:* Users can take out a single subscription and receive a single bill. The preferred solution is Internet Protocol (IP) roaming, which means that Authentication, Authorization and Accounting (AAA) mechanisms are used between the WLAN and the home PLMN; within the home PLMN, the AAA server (Multi-access Data Server) communicates with the Home Location Register (HLR) for (U)SIM based authentication
- *Access to all 3GPP PS services from both 2G/3G and WLAN:* All flows are routed through the home PLMN by using tunneling mechanisms. This interworking scenario gives the home operator full control of the service offering (including billing, policy control), and provides the user with the same set of services that he or she is used to in the mobile network, but with higher throughput.
- *Service continuity across different access technologies (2G/3G, WLAN) through support for mobile IP*: Although initially users will simply want to access their services, they will rapidly also demand service continuity, despite the implications for throughput and quality.

3.1.3 Nortel PWLAN solution

Nortel's solution for a Public WLAN is called Wireless Mesh Network. It is an enhanced public WLAN architecture that extends the reach of WLAN technology [11].

Scenario: Scenario 4 as it provides seamless online connection to public hot spots/hot zones in different areas [9], [11].

Coupling: Identified as loose coupling as the AAA server is being used.

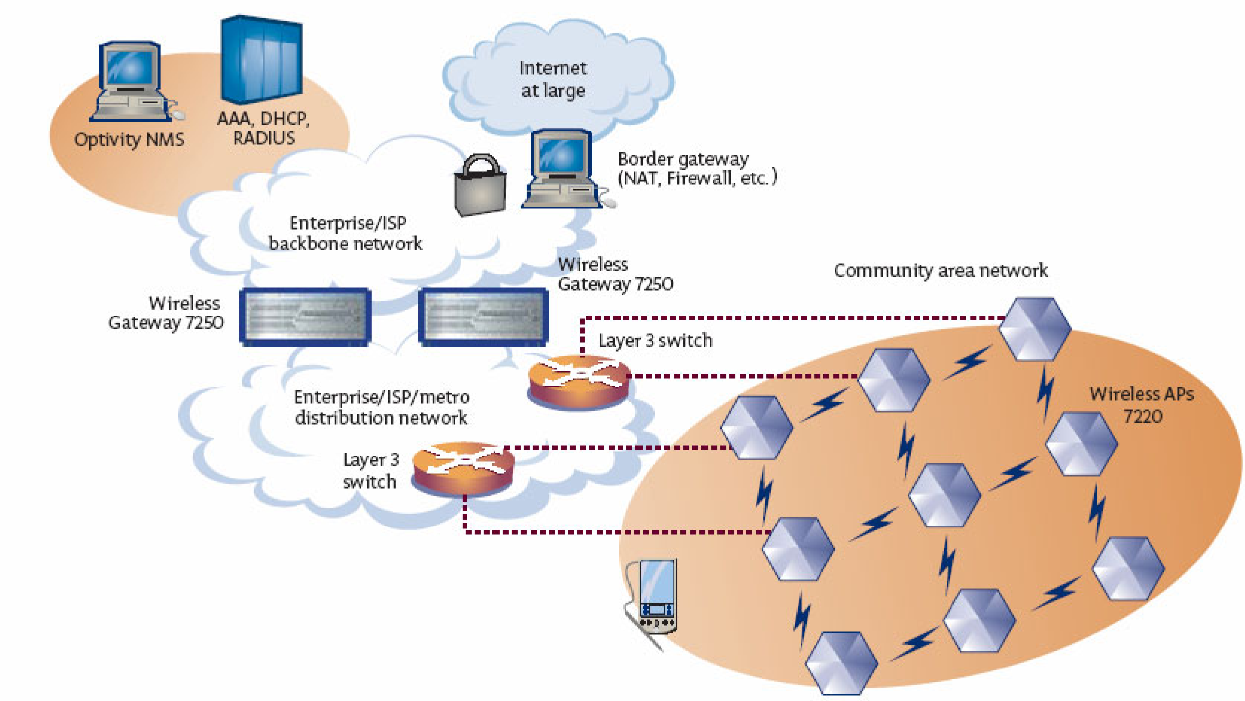


Fig. 10. Nortel's PWLAN solution architecture [10]

PWLAN solution architecture [9]

The following are the main components of the Nortel's PWLAN solution with their functions in brief:

1. Wireless Access Point 7220 (Wireless AP)
    - Routing
    - Security functions (validating connections, controlling access)
2. Wireless Gateway 7250
    - Reachability of IP subnets of subscribers and network entities
    - Data security for inter-AP links
3. Optivity Network Management System
    - Monitoring and managing network operations (discovery of APs and Gateways)
    - Fault management
    - Real-time performance metrics

AAA Server

The Nortel solution does not provide its own AAA server, so any AAA server can be integrated with the PWLAN

Services and features [10]

- Utilizes 802.11 technology — the interface of choice for high-speed wireless packet data
- Offers high-speed wireless packet data access across wider coverage areas
- Minimizes cost of capital, installation and commissioning
  - Utilizes low-cost 802.11 technology
  - Uses wireless links for backhaul to eliminate costs associated with installation of wired interconnect
  - Auto-configuration algorithms in Wireless AP eliminate costs associated with engineering and organization of the wireless backhaul network
- Minimizes cost of operations
  - Uses wireless links for backhaul to eliminate costs associated with ongoing leasing of facilities
  - Auto-configuration, self-organizing and self-healing are intrinsic to the Wireless Mesh Network
- Highly flexible in terms of capacity, coverage and availability
  - Increasing capacity, coverage and/or availability simply means deploying more Wireless Access Points
  - Wireless Access Points may be deployed indoors or outdoors

3.1.4 Juniper Networks PWLAN solution

Background

The present PWLAN approaches require gateways at every hotspot to perform the subscriber management and service functions that control users’ access to the Internet. These gateways increase the overall cost of each hotspot. The solution offered by Juniper Networks is based on removing these services & functions from the hotspot and grouping them together in the service operator’s backbone network. Therefore the overall hotspot cost decreases as gateways are no longer required [13].

Scenario: Identified as scenario 4 as it provides roaming capability**.**

Coupling: Loose coupling as the AAA server is being used and readily integrates into existing infrastructure, enabling providers to retain existing mobile GPRS/3G or DSL back-office systems [12],[14].

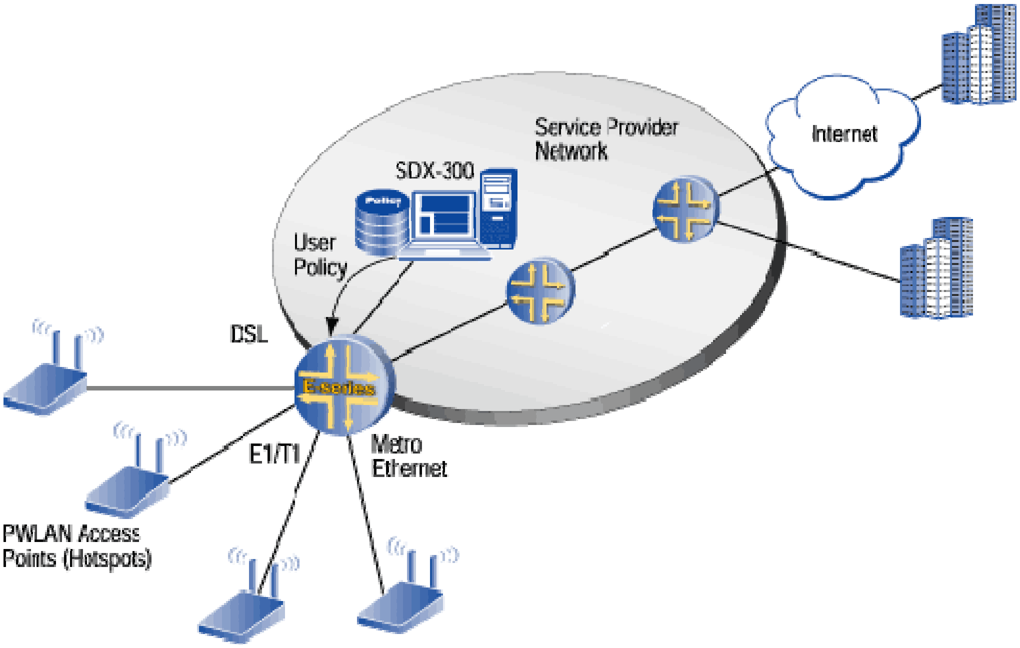


Fig. 11. Juniper Networks PWLAN solution architecture [12]

PWLAN solution architecture

The following are the main components of the Juniper Networks PWLAN solution with their functions in brief:

- E-series Broadband Service Routing Platform [15]
    - Collects output from AP's, provides session termination, enforces QoS policies and routes traffic into the IP backbone
    - MultiService Edge (protocols supported: Frame Relay, PPP, ATM, Ethernet and POS)
    - High Performance (wire speed frame processing)
    - Advanced QoS

- Service Deployment System (SDX) [16]

  Allows service providers to create and deploy new IP services to subscribers. These IP services include video on demand (VOD), IP television, and integrated voice and data. Services are offered over a variety of broadband access technologies: Wi-Fi 802.11 wireless hotspots, DSL, cable, Ethernet, ATM, Frame Relay, SONET, and fixed wireless.

Services [12]

- Services activated at service provider edge, rather than across thousands of PWLAN access points helps to reduce operational costs
- Multiple authentication options through support for clientless Web-based login and client based PPPoE, IPSec login
- Flexible solutions generate new PWLAN revenues to enable operators to move from flat-fee access to tiered and content-based fees
- E-series routers coupled with SDX-300 systems offer a unique solution deployed today for service delivery in DSL, cable and aggregation environments
- Readily integrates into existing infrastructure, enabling providers to retain existing mobile GPRS/3G or DSL back-office systems

3.2 Academic Solutions

In this section, a couple of the most recent solutions discussed in academic papers will be presented. Each section is focused on a single paper and describes a novel architecture that has been presented buy the author.

3.2.1 HCRAS: A novel hybrid internetworking architecture between WLAN and UMTS cellular networks

In this paper [27], a novel interworking architecture is proposed between WLAN and UMTS networks named Hybrid Coupling with Radio Access System (HCRAS). This architecture is proposed to overcome common inherent drawbacks in both tight and loose coupling architectures. These drawbacks are mentioned below.

1. All signaling and data transmission pass through either UMTS core networks or core Internet increasing the burden of core networks and resulting in bottleneck congestion or reconfiguration of network load when the interchanging traffic is too much.
2. When a coupling scheme is chosen, the routing for traffic is fixed or static. The routing schemes may, in some cases, result in communication disruption if the only routing path breaks down or congests.
3. Mobile IP protocols are adopted by both tight and loose coupling for terminal mobility management and handoff management. The binding process between Home Agent (HA) and Foreign Agent (FA) introduces latency and packet loss during intersystem roaming, especially during the process of vertical handoff.

This architecture is based on the loose coupling architecture as well as IPv6 technology along with the Mobile IP (MIP) protocol. The most important innovation in this architecture is the introduction of an IEEE 802.16 (WiMax) air interface. This new wireless link is named WLAN-to-UMTS Air Interface (WLUAI) and is used to communicate between a base station in the UMTS network and an access point or router in the WLAN network.

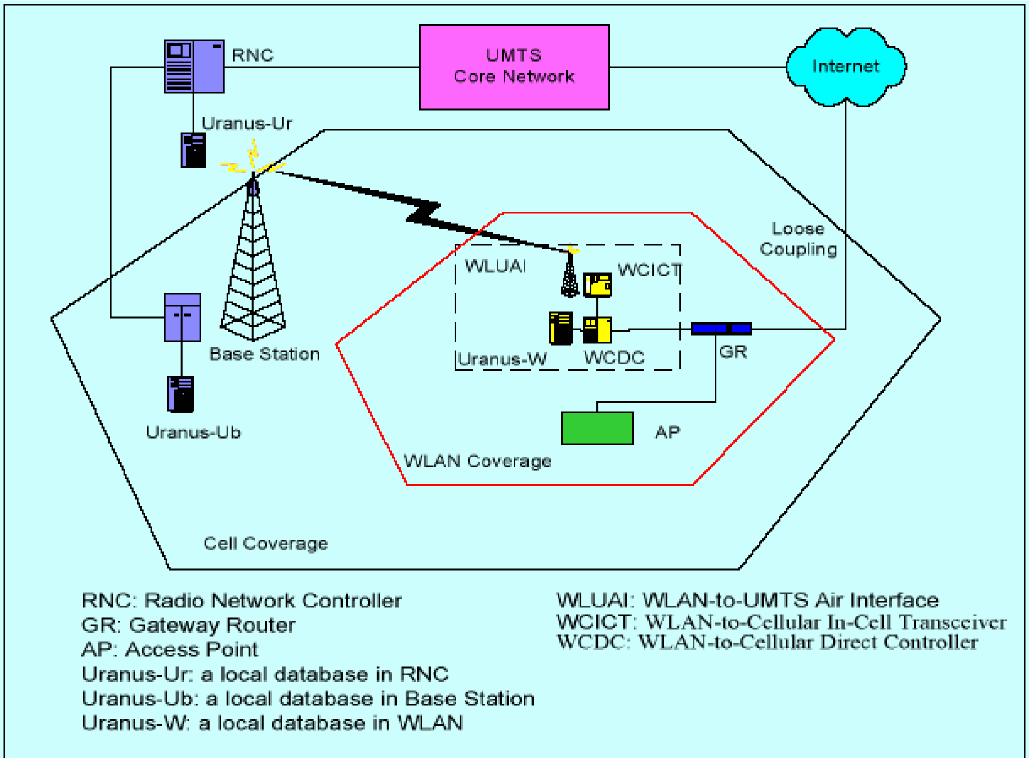

RNC
UMTS Core Network
Internet
Uranus-Ur
Loose Coupling
WLUAI
WCICT
Uranus-W
WCDC
GR
Base Station
Uranus-Ub
AP
WLAN Coverage
Cell Coverage
RNC: Radio Network Controller
GR: Gateway Router
AP: Access Point
Uranus-Ur: a local database in RNC
Uranus-Ub: a local database in Base Station
Uranus-W: a local database in WLAN
WLUAI: WLAN-to-UMTS Air Interface
WCICT: WLAN-to-Cellular In-Cell Transceiver
WCDC: WLAN-to-Cellular Direct Controller

Fig. 12. HCRAS architecture [27]

This new architecture enables us to design two new algorithms for intersystem communications and vertical handoff management. Thus both Intra-cell and Intra-RNC communications can use the wireless link to establish direct communication without routing the singling and data traffic to the core networks.

This architecture outperforms previous tight and loose coupling architectures in the areas of system routing efficiency, handoff latency, signaling cost and handoff cost. These are results are discussed in detail through mathematical analysis and numerical results.

3.2.2 An Improved Interworking Architecture for UMTS-WLAN Tight Coupling

This paper [28] is basically an extension to the work done in the previous paper. Here an interworking architecture is proposed named Tight Coupling with Wireless Access (TCWA) which improves on the performance of tight coupling. Although loose coupling is usually the preferred architecture, tight coupling is suitable for a private interworking system which requires seamless mobility, better QoS and high security. This architectures also tries to overcome the same drawbacks that the HCRAS does such as congestion in the core UMTS network and the core internet; high latency and packet loss inherent in MIP facilitated intersystem handoff and limited fault tolerance due to fixed or static routes.

TCWA as in HCRAS in based on IPv6 technology. In addition to the original wired connection between WLAN and UMTS intrinsic to a tight coupling architecture, a new wireless link named Direct Air Interface (DAI) is created utilizing the IEEE 802.16 standard. This link can be used to communicate with the BS in the UMTS network and the gateway router (GR) in WLAN. A new component not present in HCRAS is the inclusion of a control unit named SGSN Agent (SA) which is installed in the WLAN to make the WLAN network compatible with the UMTS network.

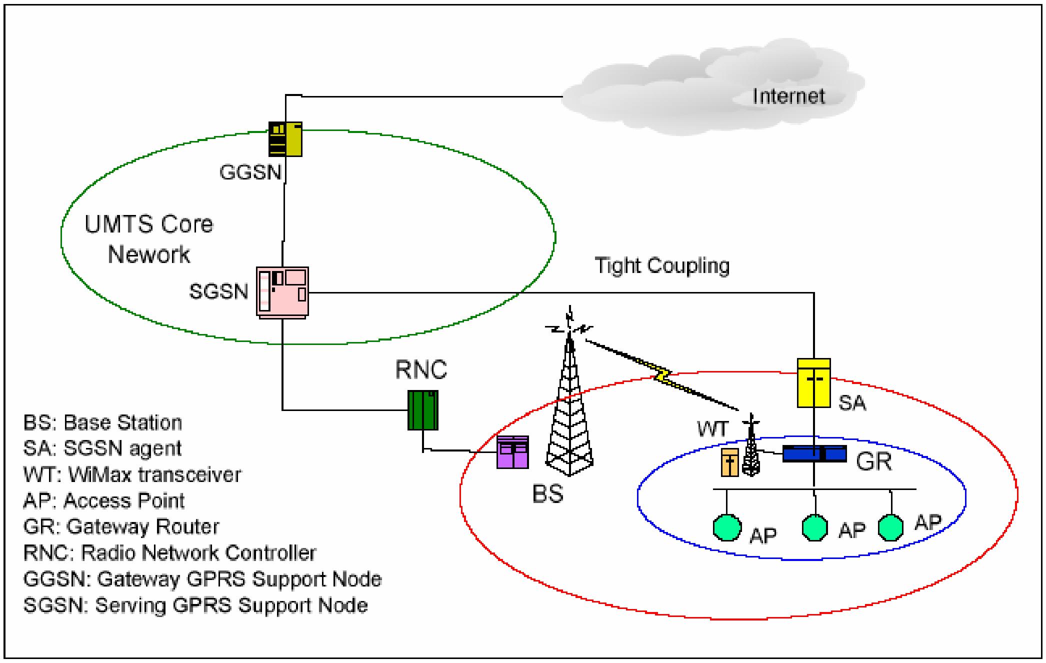


Fig. 13. TCWA architecture [28]

Based on this new architecture, algorithms for intersystem communication and vertical handoff management were developed. An important point to be noted here is that the signaling is routed using the original tight coupling link to preserve security and QoS. However the data traffic can be dynamically distributed utilizing the additional wireless link in internet-work communications.

TCWA also provides seamless handoffs by adopting a fast handoff algorithm to reduce vertical handoff latency. The fast handoff algorithm crosses three different layers, i.e., physical layer, MAC layer, and network layer. Detailed analysis and numerical results are provided to show that TCWA can reduce the traffic cost, relieve the burden of the core UMTS network, and enhance the fault tolerance.

3.2.3 Practical Considerations on End-to-End Cellular/PWLAN Architecture supporting Bilateral Roaming

This paper [29] discusses some practical aspects of Cellular/PWLAN architecture specifically in the areas of authentication and roaming mechanisms. The architecture proposed tries to reuse the existing mechanisms for user authentication, access control, billing and roaming handling procedures in mobile territory to construct a PWLAN integrated network.

There are two basic architecture styles are proposed. As this paper is geared towards the discussion of roaming mechanism, the architectures proposed consider a confederate of WISPs and cellular WISPs having roaming agreements between them thus allowing the expansion of PWLAN service territory.

1. Centralized Network with AC and AAA Server: Each hotspot is composed of one or more access points which point to a centralized AC and the core RADIUS AAA server. All AAA messages are transported within a private network. In such a construction as shown below, a central and larger Internet access backhaul is required.

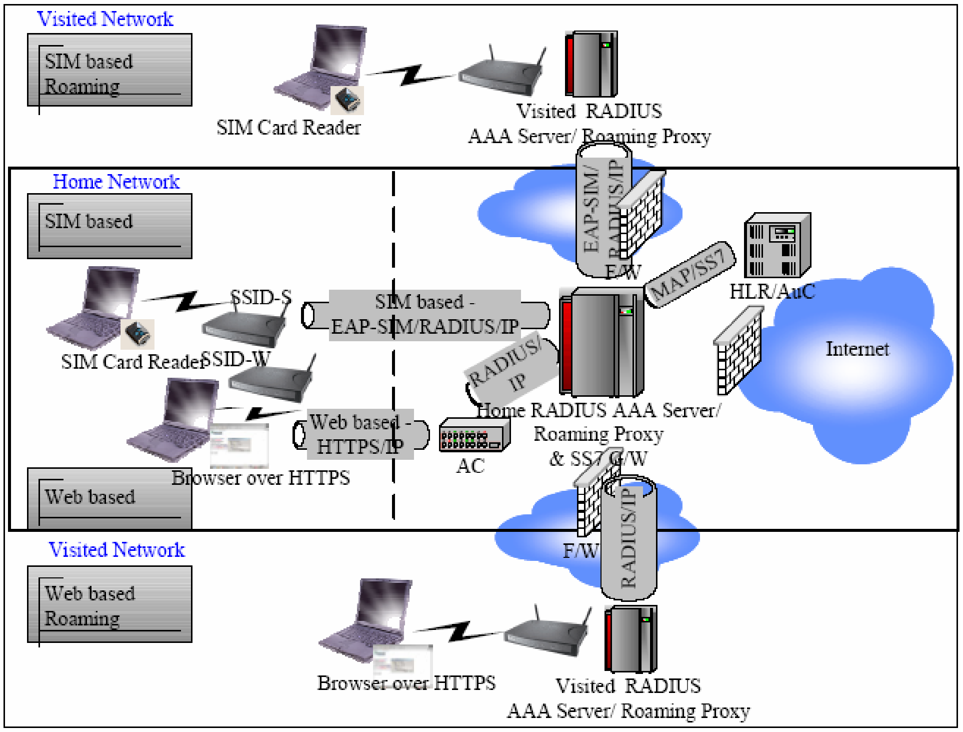


Fig. 14. Centralized Network [29]

2. Distributed Network with AAA Server over the Internet: Each hotspot is composed of one or more access points and a local AC. The AC routes user credentials over the Internet to the AAA server. In such a construction shown below, a Virtual Private Network (VPN) solution between the local AC and the AAA server is recommended. The Internet access backhaul is scattered.

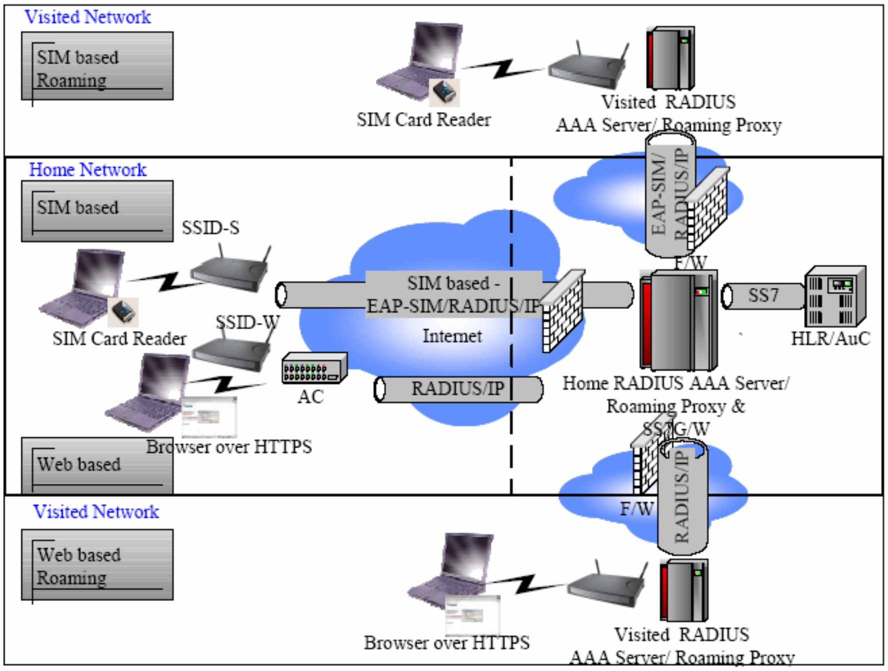


Fig. 15. Decentralized Network [29]

Two types of authentication schemes are proposed that can be used. 2G/3G SIM-based authentication (using EAP-SIM) for current mobile users can be used and Web-based authentication for ordinary users without SIM modules can be carried out simultaneously. A detailed discussion of both these schemes and practical considerations with each are provided.

Two roaming proxy solutions to exchange AAA information using the RADIUS protocol among the concerned parties are provided as well.

- Roaming with a Clearinghouse: Each party sets up one link to the roaming centre. Its advantage is one-time configuration, maintenance and cleared billing processing done between each party and the roaming centre.

- Mutual and Direct Roaming: Each party must set up many individual links directed to all the other parties. Its advantage is high profit margin without any extra charge by the roaming centre in the middle. However, individually configuration, maintenance and cleared billing procedure done by mutual operators will become complicated once the roaming group grows.

3.2.4 WLAN/3G Interworking Architectures for Next Generation Hybrid Data Networks

Cellular network operators realize that there is a strong need for integrating WLAN and 2G data networks. This paper [30] discuss some architectures capable of providing 3G/WLAN interworking for Scenario 1(3G-based common billing and customer care) and Scenario 2 (Access to 3G PS based services) situations.

Architecture for Scenario 1 Interworking:

The architecture discussed is based on the roaming case i.e. when the user is not directly connected to his 3G home PLMN. The figure below illustrates the proposed architecture. It should be noted that the user data traffic of the UTRAN UEs and the WLAN UEs follow completely different routes to the internet.

Also, in scenario 1, AAA signaling is only exchanged between WLAN and 3G PLMN for authenticating, authorization and charging purposes. The 3G AAA server in turn is interfaced to various components such as the HLR, HSS, and OCS e.t.c. As this is a roaming case, the 3G AAA server can also route AAA signaling to/from another 3G PLMN in which case it serves as a proxy as noted in the figure below.

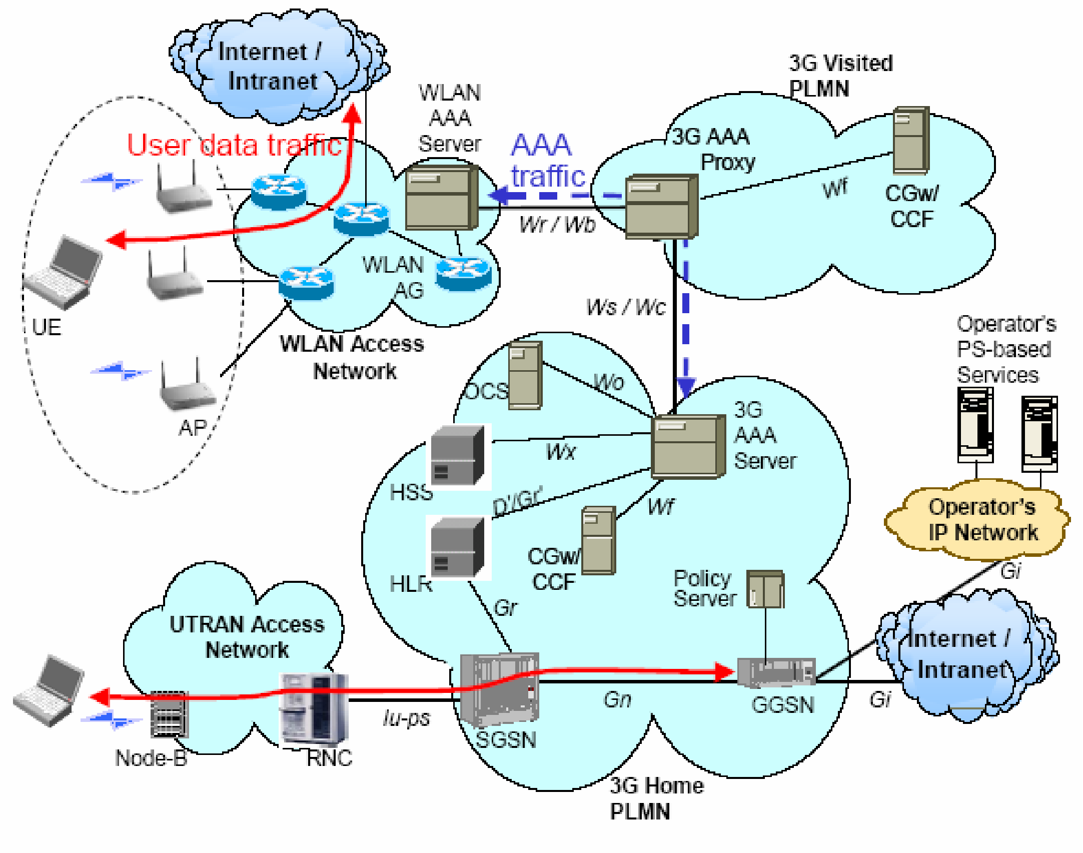


Fig. 16. WLAN/3G interworking architecture for scenario 1 (roaming) [30]

A detailed discussion of the WLAN interfaces i.e. Wr/Wb, Ws/Wc, Wf, Wo, Wx and D'/Gr' interfaces are provided. AAA signaling for scenario 1 is also discussed wherein the AAA server uses EAP-AKA or EAP-SIM to initiate and complete the authentication process.

Architecture for Scenario 2 based Interworking:

To satisfy the scenario 2 requirements, the data traffic has to be routed through the 3G system. The figure below illustrates such architecture. Note that the user data traffic is now also routed to the home PLMN to access the 3G PS-based services such as MMS, WAP e.t.c. in addition to the AAA signaling following the same route as in scenario 1.

Two new components are added called the *Packet Data Gateway* (PDG) which functions as another GGSN node in the 3G PS network and the *Wireless Access Gateway* (WAG) which is accessed in case of roaming and is in the visited PLMN.

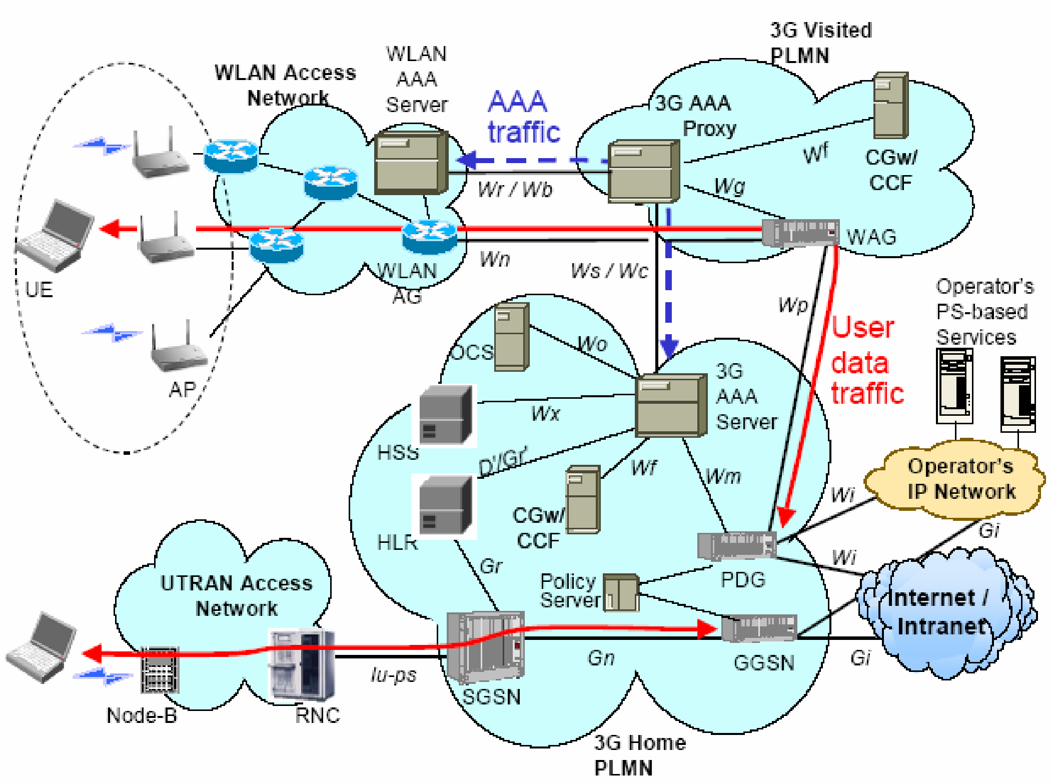


Fig. 17. WLAN/3G interworking architecture for scenario 2 (roaming) [30]

The various reference points i.e. Wn, Wm, Wi, Wg and Wp and the AAA signaling mechanism are discussed in detail. The key requirement of scenario two which is the routing of user data traffic between the UE and the PDG is enforced through establishment of appropriate tunnels. One possible method of tunnel establishment is MIP.

These architectures allows interworking across various radio access technologies such as IEEE 802.11, HiperLan/2, UTRAN, GERAN, cdma2000 e.t.c. and focus on the two scenarios which are generating the most market interest.

3.2.5 Efficient Mobility Management for Vertical Handoff between WWAN and WLAN

Current wireless networks provide different access bandwidths and coverage areas. Wireless local area networks (WLANs) such as IEEE 802.11 provide high bandwidth but have limited coverage whereas wireless wide area networks (WWAN) such as (GPRS/UMTS) have limited bandwidth but a large coverage area. A user can switch between these two networks via a procedure called vertical handoff. This paper [31] discusses in detail the methodology of a vertical handoff.

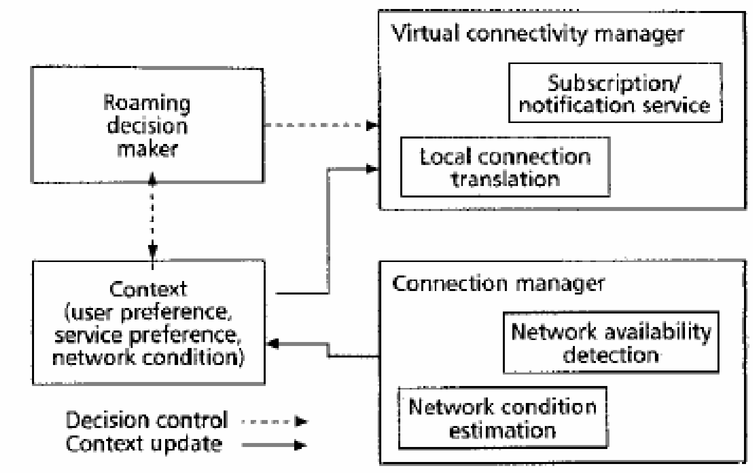

Virtual connectivity manager
Subscription/ notification service
Local connection translation
Roaming decision maker
Context (user preference, service preference, network condition)
Connection manager
Network availability detection
Network condition estimation
Decision control
Context update

Fig. 18. Vertical handoff architecture [31]

The following are the main components of the handover solution with their functions in brief:

1. Connection Manager (CM)

   The connection manager detects the availability and condition of different networks and deals with the bidirectional handover between the WWAN and WLAN. This is done using MAC layer and physical layer sensing as well as a Fast Fourier Transform (FFT) based signal decay detection scheme along with an adaptive threshold configuration approach.

2. Virtual Connectivity Manager (VC)

   This component maintains connection continuity during the handoff. It has a Local Connection Translation (LCT) which maintains a mapping between the original connection information and the current connection information. It also has a subscription/notification (S/N) service that acts as a bridge between the communicating entities.

3. Roaming decision maker and context database

   These entities interconnect the CM and the VC. The context database consists of user preferences and technical details needed to make roaming context-aware. The roaming decision considers the entire context and then makes the roaming decision.

The above components are described in detail which is followed by a performance evaluation of the constructed prototype. The evaluation results show that seamless roaming between WWAN and WLAN can be achieved with a high throughput.

3.2.6 Seamless Handover between WLAN and UMTS

This paper [32] first describes the key vertical handover challenges. These include optimum triggering and network access selection, reliability and transparency, interoperation with available services and technologies.

The handover process is divided into three phases as follows:

1. Handover Detection: This phase involves monitoring the situations when a handover is needed. For example, when the network load increases past a certain limit a handover is needed (load balancing). Another possible situation is when the user requests the handover for better QoS or when he moves out of the coverage area.
2. Handover Preparation: This phase requires checking the availability of resources and cost information of the network receiving the handover.
3. Handover Execution: Here the handover entity sends an execution message to trigger the handover

The following solution scenarios are discussed

1. Mobile IP based solution

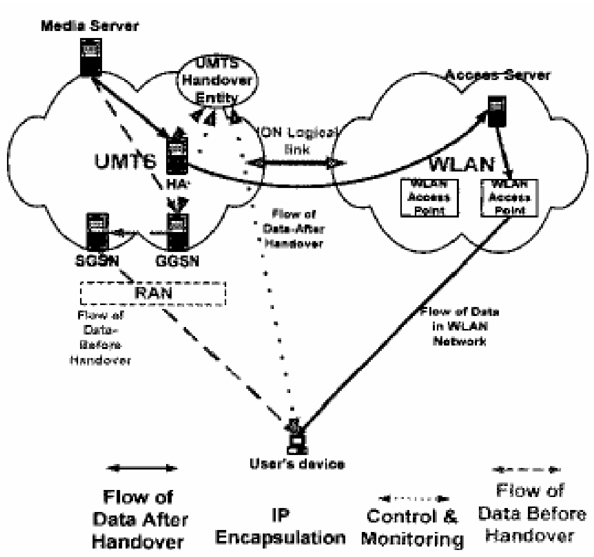


Fig. 19. Handover using Mobile IP [32]

2. SIP based solution

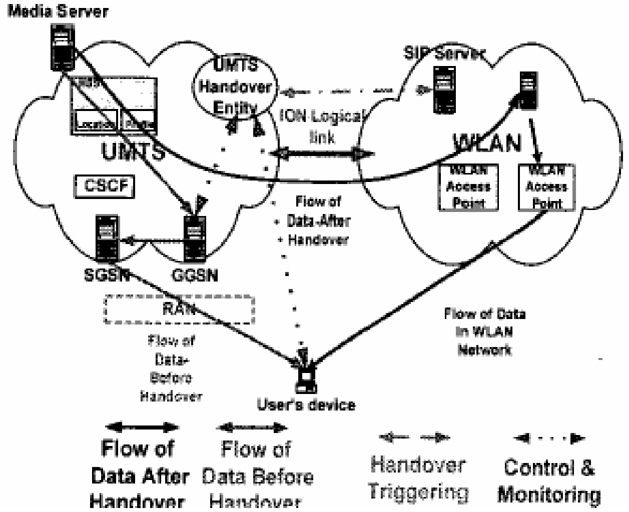


Fig. 20. Handover using SIP [32]

Mobile IP has certain drawbacks (encapsulation overhead, triangular registration with the home network etc). Therefore real time traffic cannot be handled well by Mobile IP. SIP also has drawbacks (breaking TCP connections) thereby making it unsuitable for non-real time traffic. The following are solutions that handle both real time and non-real time traffic.

1. Multi-layer Mobility Management Solution (Integrated Solution)
   In this solution, mobility between two domains is carried out by using SIP for real-time traffic and Mobile IP with location registers for non-real time traffic. Mobility within a domain is supported by micro-mobility management protocol (MMP).

2. Pure SIP with IP Encapsulation
   Here, real-time traffic is handled by using extended SIP while non-real time traffic is handled by SIP and IP encapsulation.

3. Hybrid Solution
   In this solution, both Mobile IP and SIP are for non real-time traffic and real-time traffic respectively. But the domain edge router first separates the two kinds of traffic.

Simulations are carried out on real time and non-real time traffic and a comparison based on delay and packet loss indicate that the hybrid solution is better than a pure SIP solution.

3.2.7 Design and Evaluation of UMTS-WLAN Interworking Strategies

This paper [33] presents the design and evaluation of three possible UMTS-WLAN interworking strategies, i.e. mobile IP approach, gateway approach, and emulator approach based on the current UMTS, WLAN and Mobile IP specifications.
The following are the important features of each of the interworking strategies.

1. Mobile IP approach
   This approach uses loose coupling. Mobile IP is supported by the network entities of both UMTS and WLAN as well as in the mobile terminal. When the terminal is in the UMTS network, roaming is carried out by the standard UMTS session management (SM) and GPRS mobility management (GMM). In the WLAN, the terminal uses standard IP. Mobility in the WLAN is carried out via Mobile IP. If a

handover from UMTS to WLAN is needed, the terminal disables its UMTS protocols and uses the IP stack.

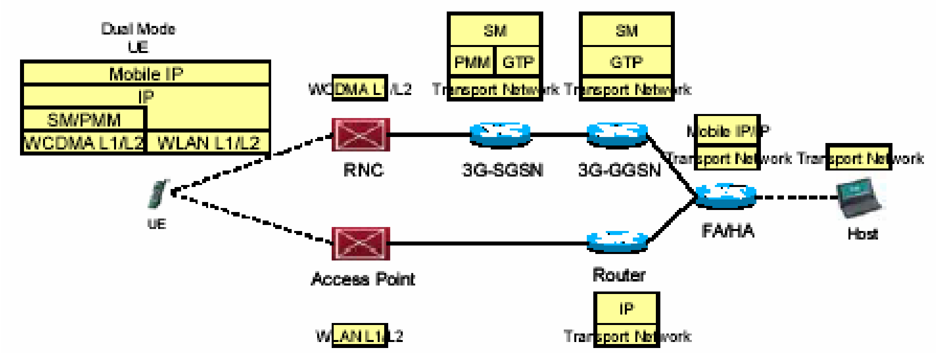


Fig. 21. Architecture for Mobile IP approach [33]

2. Gateway approach

   In this approach, the UMTS network and WLAN are connected together via a gateway. The mobile terminal uses SM and GMM in the UMTS network and standard IP in the WLAN. The interworking between the two is carried out via the gateway. It exchanges necessary information between the networks, converts signals, and forwards the packets for the roaming users. Therefore, the gateway allows the two networks to operate independently.

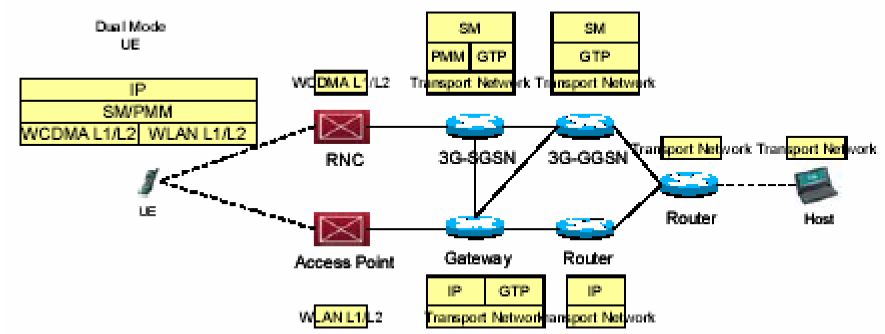


Fig. 22. Architecture for Gateway approach [33]

3. Emulator approach

   This approach uses tight coupling. From the UMTS point of view, the WLAN acts as a cell. The session management and the mobility management are handled by the UMTS SM and GMM. The mobile terminal cannot access Internet through the WLAN directly. The WLAN can be viewed as a slave network of the UMTS. The advantage of this approach is that the handover latency is much lower than other two approaches. However, every packet should to through GGSN, which becomes the bottleneck.

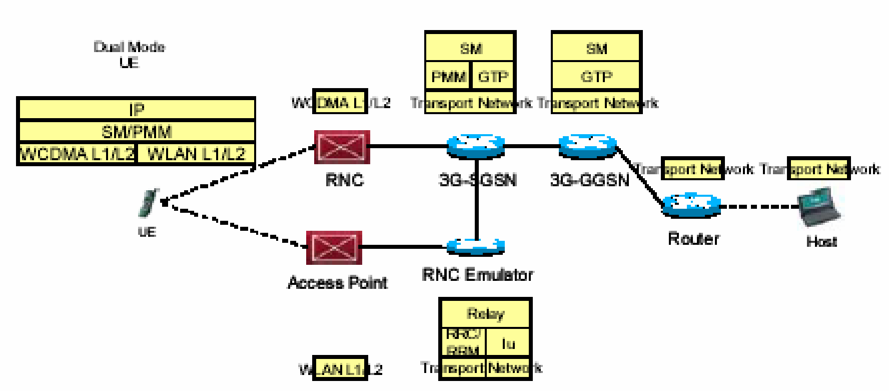


Fig. 23. Architecture for Emulator approach [33]

Simulations are carried out on the different approaches. Mobile IP suffers from long handover latency and might not be able to offer real-time services and applications. The gateway approach obtains a much lower latency than mobile IP approach. It helps the two networks to operate independently. The emulator approach achieves the best performance in terms of handover latency but lacks flexibility since the two networks are tightly coupled.

4 Mobility Management Protocols

The design of intelligent mobility management techniques is without doubt one of the research challenges facing the next generation all-IP-based wireless systems. These techniques should have capabilities that allow it to take advantage of IP-based technologies to achieve global roaming among various access technologies. At present, neither UMTS nor WLAN have any inherent functions for performing inter-technology handovers between UMTS and WLAN network technologies. Consequently, some sort of mobility protocol is required in order to execute such handovers.

Mobility protocols exist at different layers of the Internet protocol stack. Each layer has its distinct responsibilities and thus exhibits individual behavior. Likewise, the mobility protocols also exhibit individual behavior when comparing them across the layers. Present mobility management solutions usually concentrate on network layer or link layer.

Mobility management is generally considered to consist of two main components: location management and handoff management [42]. Location management allows the tracking of mobile hosts and consists of two major tasks i.e. location registration or location update and call delivery. Handoff management is the process by which a mobile host can keep its connection active as it moves from one area to another. Handoffs have been discussed in detail in previous section. A further classification can be done based on the roaming types of the mobile terminals. These are: intrasystem (intradomain) and intersystem (interdomain) roaming. Intrasystem roaming refers to moving between different cells of the same system. Intrasystem mobility management techniques are based on similar network interfaces and protocols. Intersystem roaming refers to moving between different backbones, protocols, technologies, or service providers. These are also referred to as micro-mobility and macro-mobility scenarios [43].

We shall initially talk about some widely discussed macro-mobility protocols such as the network layer protocol Mobile IP, the transport layer protocol Mobile Stream Control Transmission Protocol (mSCTP), and the application layer protocol Session Initiation Protocol (SIP). During the discussion of Mobile IP we shall also discuss some protocols related micro-mobility as most of these protocols use Mobile IP as their macro-mobility protocol and finally end with a brief comparison of the macro-mobility protocols.

4.1 Mobile IP

The Internet Protocol version 4 (IPv4) is a fundamental network layer protocol that contains addressing information and some control information that enables data packets to be routed. The addressing information consists of a 32-bit unique identifier called the IP address. Every client on the Internet is identified by a unique IP address. The routing information contains the IP destination address plus some extra information used for finding a route between the source and destination.

The dual-functions of the IPv4 create some mobility conflicts. In order for a client to be identified to others, it needs to have a stable IP address. However, if the IP address is stable, the routing to the client is also stable and the routing path essentially remains static which results in no mobility. Mobile IP version 4 (Mobile IPv4) is a supplementary network layer mobility protocol that addresses this problem by allowing the client to effectively utilize two IP addresses, one for unique identification, and the other for routing. This address allows the mobile terminal to specify its current location in the network. The home address is a static address used for identification, and the care-of-address is an address that changes at each new point of attachment. The upper transport and application layers keep using the home address, allowing them to remain ignorant of any mobility taking place, and thereby keeping the TCP connection alive. This means that Mobile IPv4 can provide transparency to the upper layers while providing seamless mobility using the care-of-addresses. [44][45][46].

When the client is in its home network, standard IP routing mechanisms deliver incoming and outgoing data packets to and from the client, respectively, using the home address. If the client changes location during the communication from its home network to another (foreign) network, standard IP routing mechanisms do not suffice, as it can no longer be reached by its home address. Instead it must obtain and register a care-of-address using Mobile IPv4 in order to continue communication [44]. Figure below illustrates Mobile IPv4 mobility management. [44] – [46] provide in depth detail about the Mobile IP protocol. However, the message exchanges that occur during a mobility scenario is briefly discussed below

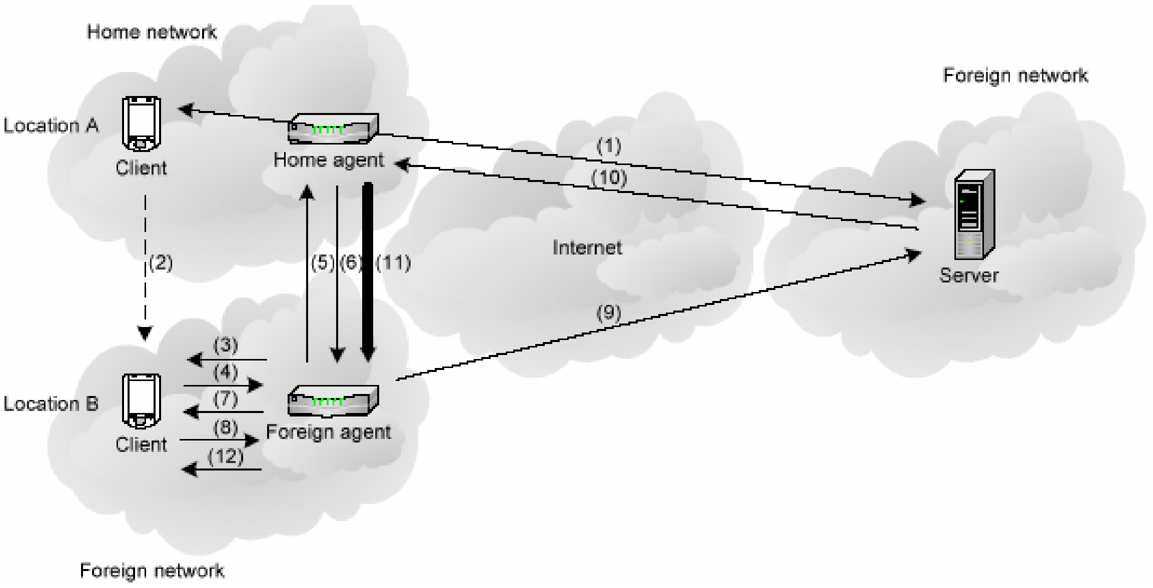


Fig. 24. Mobile IPv4 mobility management [56]

The client is first located in its home network at location A where it has established communication with a server in a foreign network through the Internet (1). The client then changes position from location A in the home network to location B in a foreign network (2). In order to sustain communication Mobile IPv4 now enters the picture.

Mobile IPv4 employs two new network elements: a home agent in the home network and a foreign agent in the visited network. These two types of agents are new network elements and are usually the border routers. The foreign agents as well as the home agent advertise their availability by attaching a special extension to the router advertisement, the so-called agent advertisements. The advertisements are usually broadcasted at regular intervals, e.g. once a second or once every few seconds. Alternatively, the client can send a router solicitation to ask the router to send router advertisements if impatient.

When the client receives an agent advertisement, it determines whether it is in its home network or a foreign network (3). If it finds out that it is in a foreign network, it registers its new care-of-address found in the agent advertisement, or alternatively found by contacting Dynamic Host Configuration Protocol (DHCP) or Point-to-Point Protocol (PPP), with its home agent through the foreign agent (4)(5). Whenever the client moves to a new foreign network, it registers its new care-of-address.

The home agent responds to this request by updating its routing table with the new care-of-address (typically performed though optional), approving and authorizing the request, and finally returning a registration reply to the client through the foreign agent (6)(7). The reply includes a registration lifetime, which specifies how long the registration will be honored by the home agent. The home agent associates the home address of the client with the care-of-address until the registration lifetime expires. The triplet of a home address, a care-of-address, and registration lifetime is called a binding for the client. A registration request is therefore considered a binding update sent by the client and the registration reply a binding acknowledgment.

After successful registration, the communication between the client and server can continue unaffected. When the client has data packets destined for the server, it sends the packets to the foreign agent, which then forwards them directly to the server (8) (9).
In the reverse direction, when the home network receives data packets destined for the client, it intercepts the packets and encapsulates them by preceding each packet with a new IP header, so-called IP-within-IP encapsulation (10). The new IP header contains among other things the destination address, which is the care-of-address.

The home agent then tunnels the encapsulated data packets to the foreign agent using the care-of-address (11). The foreign agent receives the data packets, de-encapsulates them, and forwards them to the ultimate destination (12). This asymmetric way of routing data packets to and from the client has given it the name triangular routing. [44][45][46].

Mobile IPv4 was originally defined as an add-on for IPv4. For the emerging IP version 6 (IPv6), the mobility support (Mobile IPv6) has been an integrated feature from the beginning. This means that some of the identified problems from Mobile IPv4 have been addressed in Mobile IPv6. The major problems with Mobile IPv4 are deployment, triangular routing, tunneling overhead, and security [45] [47]. Mobile IPv6 and its approach to tacking these problems are briefly discussed below.

A Mobile IPv4 deployment requires the implementation of foreign agents in each potential foreign network. However, the requirements for foreign agents means extensive network configuration. Mobile IPv6 deals with this problem by completely eliminating the foreign agents. It retains the ideas of a home network, home agent, and the use of encapsulation to deliver packets from the home network to the client's current point of attachment. The figure below illustrates Mobile IPv6 mobility management.

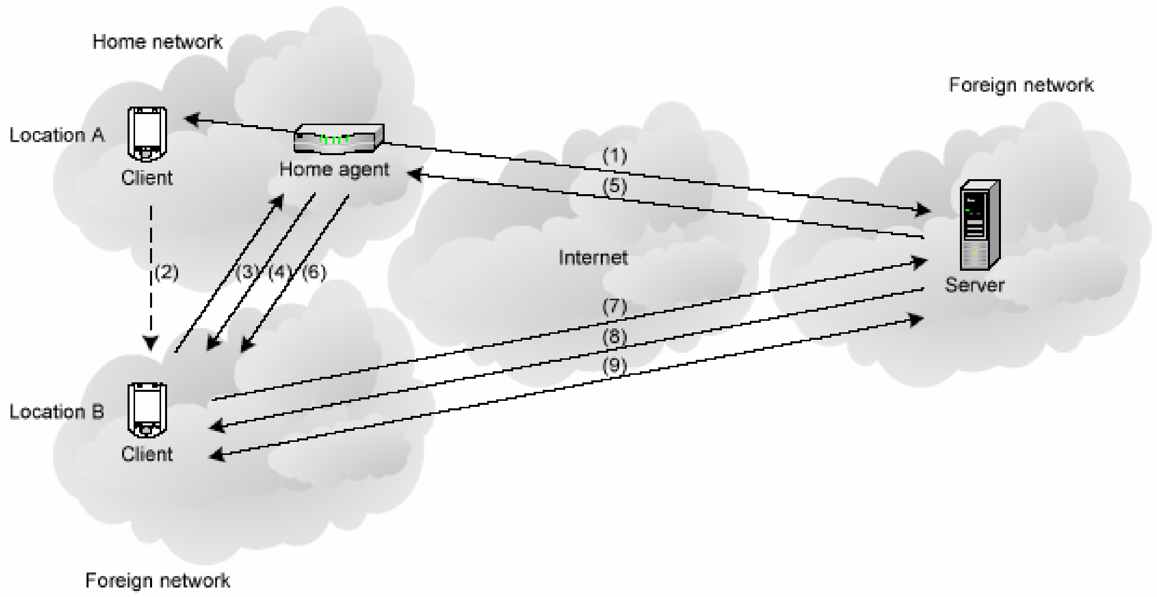


Fig. 25. Mobile IPv6 mobility management [56]

The scenario for Mobile IPv6 is similar to the Mobile IPv4 scenario. The client is first located in its home network at location A where it has established communication with a server in a foreign network through the Internet using standard IP routing mechanisms (1). The client then changes position from location A in the home network to location B in a foreign network (2).

Instead of listening to the foreign agents advertising their availability through agent advertisements, the client listens to the router advertisements. The router advertisements in IPv6 have been extended with some extra bits of which the prefix information option format has received one bit extra, allowing the router to efficiently advertise its global IPv6 address instead of the link local address. The client can determine whether it is in its home network or a foreign network from the network prefix in the router advertisement. If the network prefix equals the network prefix of the home address of the client, it is on its home network. If the client finds out that it is in a foreign network, it obtains a care-of-address and registers it with its home agent. The care-of-address is obtained using either state-full or stateless address auto-configuration. In the first situation, the client obtains the care-of-address by contacting e.g. a DHCPv6 server. In the latter, the client extracts the network prefix from the router advertisement and adds a unique interface identifier of the client to form the care-of-address.

When the client has obtained a care-of-address, it sends a binding update to its home agent (3). The home agent responds with a binding acknowledgement (4). The Mobile IPv6 registration process thus mainly differs from the Mobile IPv4 registration process by the lack of foreign agents. [45][47][48].

Triangular routing refers to the process of all data packets sent to the client being routed through the home agent which adds delay to the traffic towards the client (not from the client) and increases network load. This problem is addressed in Mobile IPv6 by implementing route optimization. Route optimization was initially specified as a nonstandard extension for Mobile IPv4. Route optimization is, however, a fundamental part of the Mobile IPv6 protocol and not just an extension.

In route optimization, the client first registers its care-of-address with the home agent as described above. It then sends a binding update directly to the server to notify the server of its new care-of-address (7). The server responds with a binding acknowledgment and the two nodes can continue their communication unaffected (8) (9). The home agent can also receive data packets from the server before the client has registered its care-of-address with the server (5). In this case, the home agent receives the data packets, encapsulates them, and forwards them to the client (6). When the client receives the first encapsulated packet from the home agent, it sends a binding update to the server, which is responded with a binding acknowledgement from the server (7) (8). After this, the server and client can continue communication on a direct path (9) without any interaction with the home agent. Thus, the transport delay is reduced and the network capacity preserved solving the triangle routing problem. [45][47][48].

When the server sends packets to the client, the packets go via the home agent that intercepts the packets and encapsulates them, and on to the foreign agent through a tunnel. The tunneling means that an overhead of typically 20 bytes is added to each packet (IP-within-IP encapsulation). Mobile IPv6 approaches the tunneling overhead problem simply by removing the tunneling function. [45][47][48].

Finally, there are some security concerns. When the client registers a care-of-address with its home agent, the home agent must be certain that the request was originated by the client and not some malicious node pretending to be the client. A malicious node could cause the home agent to alter its routing table in such a way that the client would be unreachable to all incoming communications, and in worst case that these communications would be directed to the malicious node. Mobile IPv4 employs a security association between the home agent and the client by using the Message Digest 5 algorithm with 128-bit keys to create digital signatures that cannot be forged for registration requests. Mobile IPv4 does not, however, require that the foreign agents authenticate themselves to the client or home agent. Mobile IPv6, in contrast, implements strong authentication and encryption features in all nodes using IP Security (IPsec). [45][47][48].

It might seem obvious just to implement Mobile IPv6 instead of Mobile IPv4 with its inherent problems. However, the decision is not an easy one. The main reason is that Mobile IPv6 has been implemented only on a small scale of test networks. Implementing Mobile IPv6 on a full scale will still take a few years. Nevertheless, many people foresee the arrival of Mobile IPv6 sooner rather later.

4.1.1 Micro-mobility management protocols

In the context of Mobile IP, micro-mobility management protocols deal with mobile hosts moving between subnets of one domain. These protocols aim to reduce the signaling traffic and the delay to the home network that occur during intra-domain roaming. Basically there are two main types of micro-mobility management protocols: tunnel- based and routing-based [5]. In tunnel-based protocols, the signaling traffic and delay are reduced by using local registration and encapsulation techniques. In routing-based protocols, routers store host-specific routes that are used to deliver packets. Routing information is updated on the basis of host mobility. This section discusses the following mobility management protocols: IDMP (tunnel-based), HAWAII and Cellular IP (both routing-based). These protocols usually use Mobile IP as the macro-mobility protocol but it is possible for these protocols to be used by other protocols such as SIP as is discussed in the subsequent sections.

4.1.1.1 Intra-Domain Mobility Management Protocol (IDMP)

The Intra-Domain Mobility Management Protocol (IDMP) is proposed to reduce the latency of intra-domain location updates and the mobility signaling traffic that is present in Mobile IP. Basically IDMP is two level generalization of the Mobile IP architecture.

IDMP offers intra-domain mobility by using two dynamically configured care-of addresses (CoAs) for routing the packets destined to mobile nodes [38]. IDMP can be combined with multiple global mobility protocols (Mobile IP or SIP). The following figure shows the IDMP logical elements and architecture

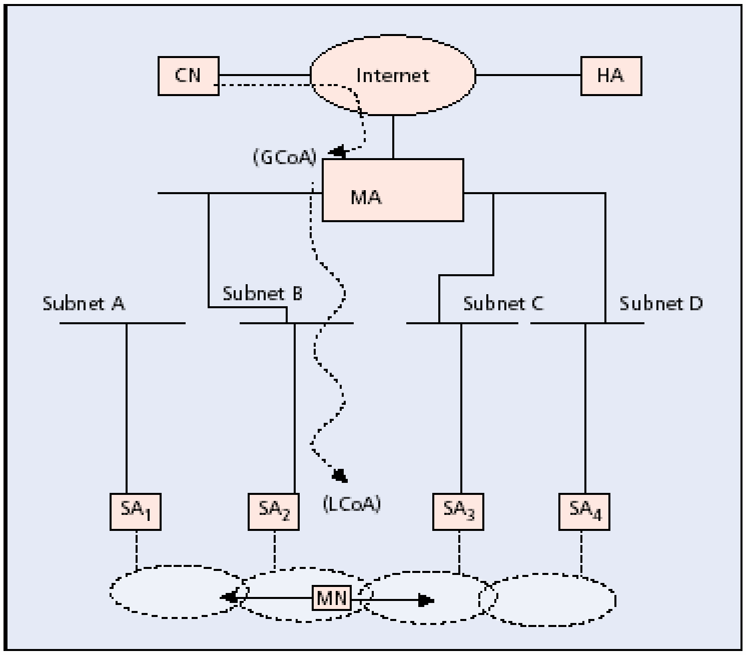


Fig. 26. IDMP logical elements and architecture [39]

The Mobility Agent (MA) acts as a domain-wide point for packet redirection. A Subnet Agent (SA) provides subnet-specific mobility services.

With IDMP, an MN obtains two concurrent CoAs [38]:

Local Care-of Address (LCoA): identifies the MN's attachment to the subnet level and changes every time the MN changes subnets. By updating its MA of any changes in the LCoA, the MN ensures that packets are correctly forwarded within the domain.

Global Care-of Address (GCoA): identifies the MN’s attachment to the current domain. This address resolves the MN’s current location only up to a domain-level and hence remains unchanged as long as the MN stays within a single domain.

The MN ensures that packets are routed correctly to its present domain by issuing global updates that contain this GCoA [38]. Under IDMP, packets from a remote CN are forwarded to the GCoA and are intercepted by the MA which then tunnels these packets to the MN’s current LCoA. Since global binding updates are generated only when the MN changes domains and obtains a new GCoA, this approach reduces the global signaling load.

When the MN first moves into a domain, a subnet-specific registration using IDMP is performed to obtain a local care-of address (SA_2's address in Figure above). During this subnet-specific registration process, the serving SA (SA_2 in this case) dynamically assigns the MN a Mobility Agent (MA). The MN then sends the LCoA to this MA by performing an intra-domain location update. The MA then sends an intra-domain location update reply containing either its address or a separate GCoA. It is also up to the MN to register with the necessary remote nodes (e.g., HA if Mobile IP is used) by sending a global location update. However this is excluded in the IDMP specifications [38]. The following figure shows the message exchange when the MN first moves into a new domain.

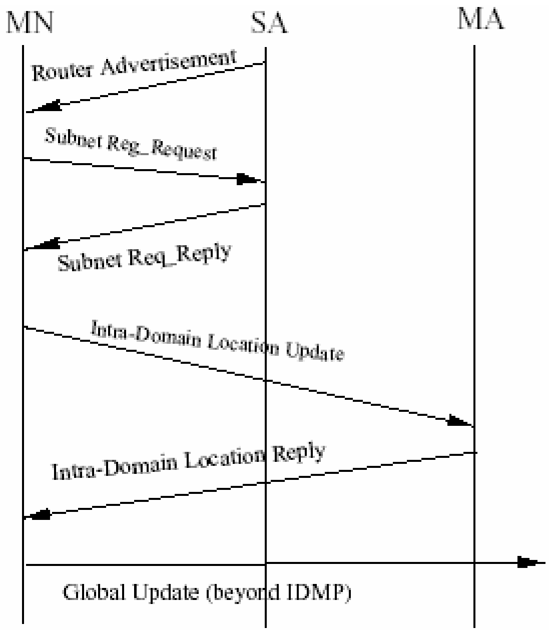

MN
SA
MA
Router Advertisement
Subnet Reg_Request
Subnet Req_Reply
Intra-Domain Location Update
Intra-Domain Location Reply
Global Update (beyond IDMP)

Fig. 27. Initial intra-domain location update message exchange [38]

When the MN moves to another subnet within the same domain, it performs a new subnet-specific registration with the new SA. Since the MN is already assigned a MA, a new MA address is not allocated by the new SA. The MN then sends the new LCoA to this MA by performing an intra-domain location update. As the global care-of address remains unchanged, no global messages are needed [38]. The following figure shows the message exchange during subsequent intra-domain movement.

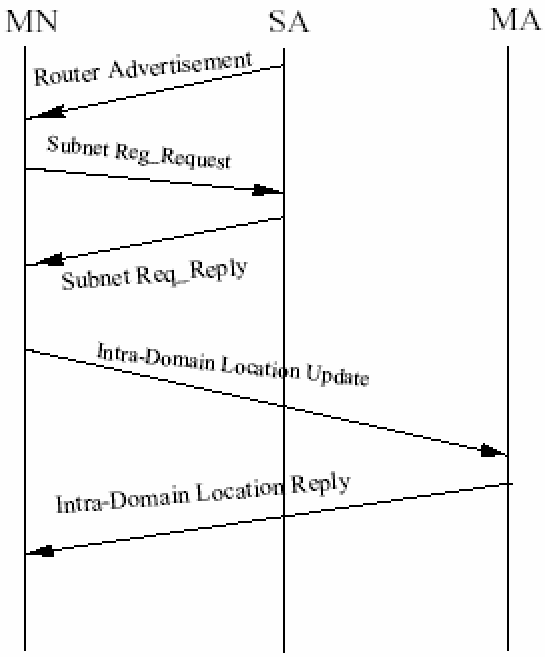


Fig. 28. Subsequent intra-domain movement message exchange [38]

4.1.1.2 Handoff Aware Wireless Access Internet Infrastructure (HAWAII)

HAWAII is a routing based protocol that supports mobility in wireless WAN's [40]. HAWAII divides the network into hierarchies of domains.

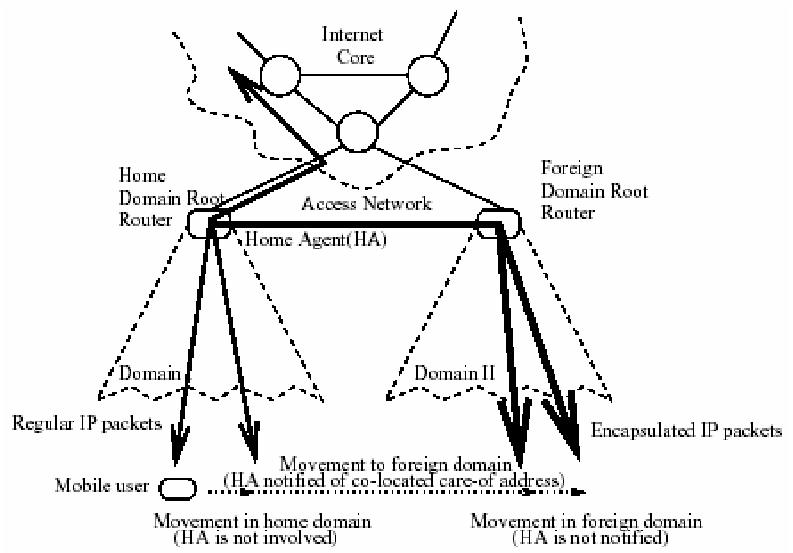


Fig. 29. Hierarchy with domains [40]

Each domain has a gateway called the domain root router [40]. Each host is assumed to have an IP address and a home domain. The mobile host retains its IP address when moving in its home domain. Packets that are intended for the mobile host first reach the domain root router based on the subnet address of the domain. The data is then sent to the mobile host over special dynamically established routes.

Mobile IP is used when a mobile host moves into a foreign domain. If the foreign domain is also based on HAWAII, then the mobile host is assigned a co-located care-of address from its foreign domain. The home agent in its home domain then tunnels packets to the care-of address. The care-of address remains unchanged when the mobile host moves within the foreign domain and dynamically established paths provide connectivity [40].

The protocol contains three types of messages for path setup: power-up, update and refresh. When a mobile host attaches to a domain, it sends a path setup power-up message. This establishes specific routes between the host and the domain root router as well as any intermediate routers on the path towards the mobile host.

When the mobile host moves within a domain, end-to-end connectivity is maintained by using path setup update messages [40]. The routing entries for the mobile hosts in specific routers in the domain are updated with these messages. Therefore packets arriving at the domain root router can reach the mobile host. A particular path setup scheme determines when, how and which routers have to be updated.

Periodic path refresh messages are sent to the base station by the attached mobile host to maintain the routing entries. Periodic aggregate hop-by-hop refresh messages are then sent by the base station and the intermediate routers towards the domain root router.

HAWAII ensures that data disruption during handoff is limited by using specialized path setup schemes. When the mobile host moves within its home domain, the home agent is not involved and packets are delivered to the mobile host without any tunneling. Therefore data transfer efficiency is maintained. Also, the home agents and correspondent hosts are not involved in the intra-domain movement of hosts. This helps in supporting a large number of mobile hosts, thereby improving scalability [40].

4.1.1.3 Cellular IP

Cellular IP is a micro-mobility protocol that provides local mobility and handoff support for hosts that move frequently.

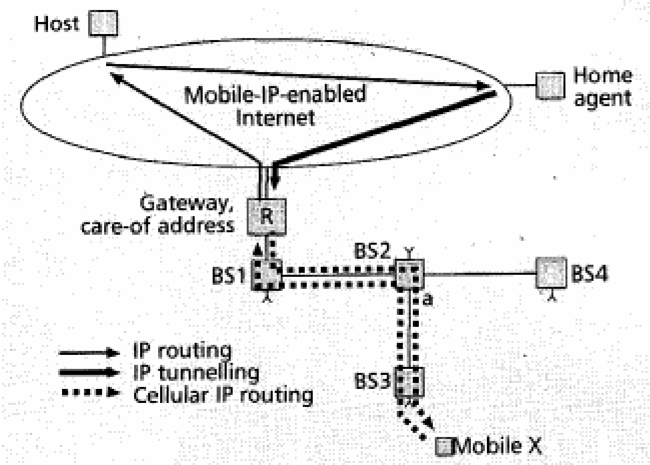


Fig. 30. Cellular IP access network [41]

The main component of Cellular IP access networks is the base station. It performs all mobility-related functions and acts as a wireless access point and routes IP packets. Cellular IP access networks are connected to the Internet thorough gateway which handles intra-domain mobility.
The IP address of the gateway is used by the mobile hosts as their Mobile IP care-of address to attach to the access network [41].

As shown in the figure, the correspondent node (Host) sends packets to the mobile host's HA which then tunnels these packets to the gateway router which then de-tunnels them and sends them towards a base station in the cellular IP network. The mobile hosts are identified by their home address, and data packets are routed without tunneling or address conversion. The Cellular IP routing protocol ensures that packets are delivered to the host's actual location. Packets transmitted by mobile hosts are first routed toward the gateway and from there on to the Internet.

In Cellular IP, location management and handoff support are integrated with routing. Mobile hosts transmit data packets regularly to refresh their location to reduce control messaging. Packets sent by a mobile host (uplink packets) are routed to the gateway on a hop-by-hop basis. The path taken by these packets is cached by all intermediate base stations. Packets addressed to the mobile host (downlink packets) use the same path but in reverse order. When the mobile host has no data to transmit, it sends special IP packets toward the gateway to maintain its downlink routing state.

In the above sections, we have discussed different micro-mobility protocols. However, their common aim is to limit most of the signaling traffic into one domain in order to decrease the global signaling load.

4.2 Mobile Stream Control Transmission Protocol (mSCTP)

Transport layer mobility is proposed as an alternative to network layer mobility for seamless mobility management. Mobility management in the transport layer is solely accomplished by use of Stream Control Transmission Protocol (SCTP) and its currently proposed Dynamic

Address Reconfiguration (DAR) extension. SCTP with its DAR extension is called Mobile SCTP (mSCTP).

mSCTP is a transport layer protocol similar to Transmission Control Protocol (TCP) that operates on top of the unreliable connection-less packet network. It provides unicast end-to-end communication between two or more applications running in separate hosts and offers connection-oriented, reliable transportation of independently sequenced message streams. The biggest difference between mSCTP and TCP is multi-homing i.e. the idea of having several streams within a connection (multi-streaming) and the transportation of sequence of messages instead of sequence of bytes.

mSCTP is capable of handling several multiple IP addresses at both endpoints while keeping the end-to-end connection alive. These addresses are considered as logically different paths between the endpoints. During initiation of the connection lists of addresses are exchanged between the endpoints. Both endpoints must be able to receive messages from any of the IP addresses related to the endpoints. One address is chosen as the primary address and is used as the destination for normal transmission. The other addresses are used for retransmissions only. The SCTP DAR extension enables the endpoints to add, delete and change the primary address dynamically in an active connection without affecting the established connection. [49][50][51][52][53][54].

The mSCTP scenario is assumed to start with a client located in its home network at location A where it has established communications with a server in a foreign network through the Internet. Figure below illustrates the mSCTP connection initialization.

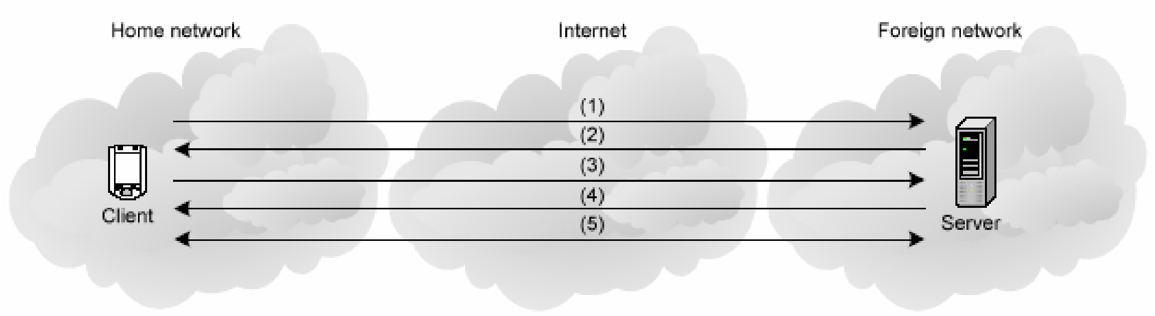


Fig. 31. mSCTP connection initialization [56]

In order to set up a transport layer connection with the server, the client first sends an init request to the server including a list of IP addresses and port number that will be used by the client (1). The server responds with an init acknowledgment including a state cookie and a list of IP addresses and port number that will be used by the server if it accepts the request (2). It can also contain indicate the primary IP address. The client must then return the state cookie from the init acknowledgement in what is known as a cookie echo (3). When the server receives the cookie echo, it moves to established state, and responds with a cookie acknowledgment (4). The exchange of cookies is basically a set of security enhancements. Finally, the client moves to established state (5) and communication can now take place between the nodes. Similar to TCP, mSCTP automatically produces an acknowledgement in between each sequence of messages. [52]

Let us now consider that the client then changes its position from location A in the home network to location B in a foreign network during communication. To keep the connection, mSCTP is now introduced. The figure below illustrates mSCTP mobility management.

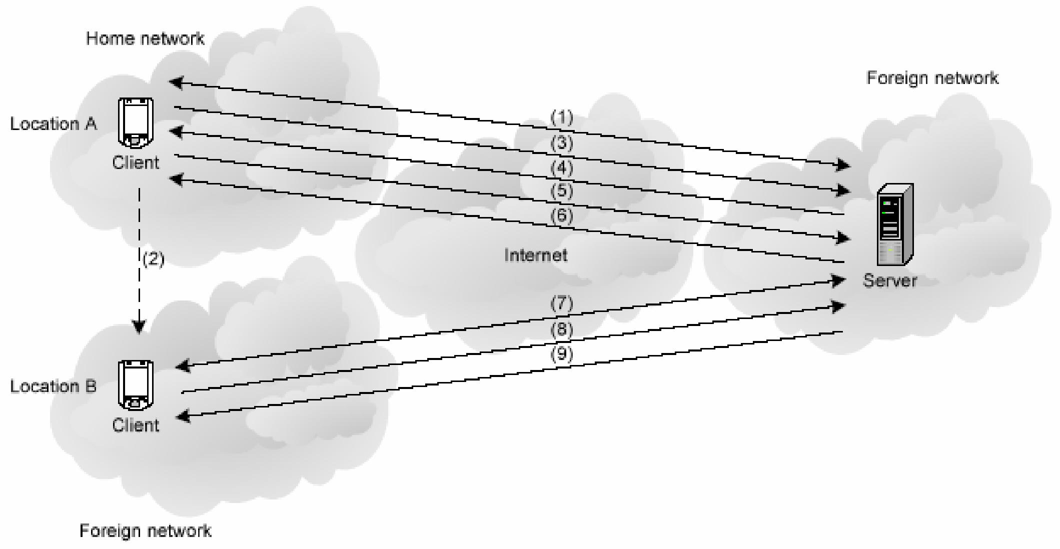


Fig. 32. mSCTP mobility management [56]

During ongoing communication between the client and server (1), the client moves from location A to location B (2). As the client moves into the coverage area of the foreign network, it receives an IP address from the local space at location B either by contacting DHCP or by IPv6 addresses auto-configuration. The client is now able to establish a link with the server with its second IP address and thus become multi-homed, i.e. reachable by the way of two different networks. The client therefore tells the server via the first link that it is reachable by a second IP address (3). In other words, it adds the newly assigned IP address to the association identifying the connection to the server. The server responds to this by returning an acknowledgment (4). As the client leaves the coverage area of the home network, the client tells the server to set the newly assigned IP address to the primary IP address (5), which the server responds to with an acknowledgment (6). The new primary IP address now becomes the destination address for further communication and the communication can continue unaffected over the new link (7). Finally, the client tells the server to delete the first IP address, i.e. remove it from the association (8), which produces an acknowledgment from the server (9). [49][50][52].

mSCTP, in contrast to other mobility protocols, does not require any additional entities or modifications of network entities. The only thing required is the support for mSCTP in the client and server. This makes the network architecture simple and easier to deploy. The users of the client and server are similar to Mobile IP responsible for downloading mSCTP and setting it up on the client and server, respectively. As with Mobile IP, mSCTP can be expected to be an innate part of the client/server operating system in some years.

4.3 Session Initiation Protocol (SIP)

An alternative to network and transport layer mobility is application layer mobility. A viable application-layer mobility protocol is the IETF-developed signaling protocol Session Initiation Protocol (SIP). SIP is a signaling protocol mainly used to establish, modify, and terminate multimedia sessions consisting of multiple media streams, unicast as well as multicast. The multimedia streams include audio, video, and any Internet-based mechanisms such as distributed games, shared applications, shared text editors etc.

SIP users are addressed using email-like addresses like user@host, where “user” is the user name and “host” is the domain name or numerical address. The SIP address changes when e.g. the user changes the network provider, moves to another job or changes organization, not necessarily when the user changes location. For temporary change of location purposes, the user can have multiple SIP addresses and redirect calls to the current location. A SIP user can thus be represented by multiple SIP addresses, each of which can furthermore point to multiple devices. A SIP address being able to concurrently relate to multiple devices is something no other signaling protocol currently does.

SIP defines four logical entities, namely user agents, registrars; redirect servers and proxy servers, and an abstract service known as the location service. The user agent has two roles: a user agent client that issues requests and receives responses and a user agent server that receives requests directed to it and issues responses by accepting, rejecting or redirecting the request.

The registrar is responsible for maintaining user agent access information based on incoming modification requests from the user agent. The registrar only manages requests targeted at SIP addresses within its managed domain. Typically, such requests concern the change of location of the user. All incoming requests are communicated onward to the location service that maintains this information.

The redirect server keeps track of the user's location and manages redirecting contacts to the user agents that are out of the registrar's domain. The redirect server returns only the location of the user; it does not relay any messages.

The proxy server is responsible for relaying the messages. The proxy servers are classified in two ways. The first classification is where the proxies are classified by the location of the proxy in the path from the source user agent to the destination user agent.

The closest proxy to the source user agent is the outbound proxy, while the closest proxy to the destination user agent is the inbound proxy. All proxies in between these two are the intermediate proxies. The second classification is state-fullness. Stateless proxies forward requests and responses without actively generating new types of requests and responses and thus without ensuring the request's reliability. State-full proxies respond to the user agent client requests with the response closest to the user agent client's requirements and maintain state for the transaction.

Finally, the location service is a database that contains location information of the user agents. The location service is used by the proxy and redirect services to locate the user agent client and user agent servers. Typically, a physical SIP server implements a redirect and proxy server with information provided by a built-in registrar. The location service can be stored either locally at the SIP server, or in a dedicated location server. [55][56][57].

The SIP scenario starts with a client located in its home network at location A where it has established communication with a server in a foreign network through the Internet. Figure below illustrates the SIP connection initialization.

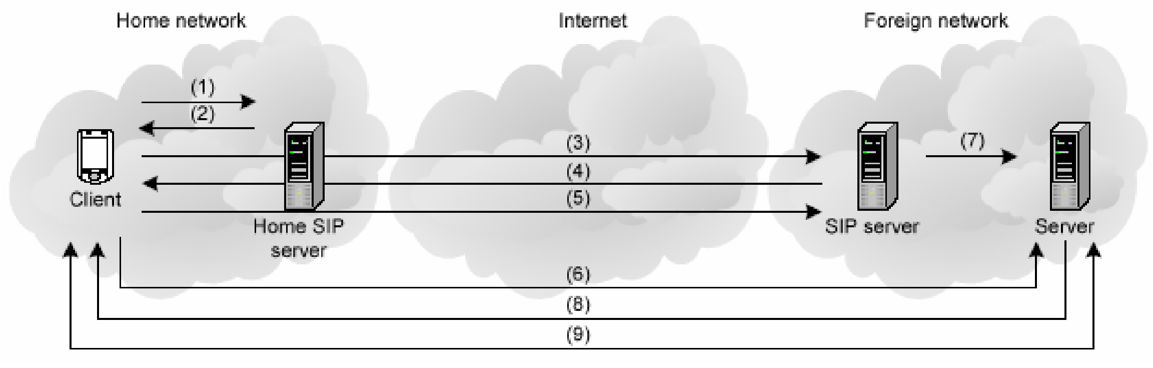


Fig. 33. SIP connection initialization [56]

As the client attaches to the home network, its user agent sends a location update to the registrar in the home SIP server (1) The registrar processes the update message and forwards it to the location service, which stores the information. The home SIP server in return sends an acknowledgement (2). The client now wants to communicate with a server located in a foreign network, so its user agent sends an invite request to its home SIP server. The home SIP server recognizes that the request is not meant for it and forwards it to the SIP server belonging to the server's domain (3). The redirect server in the server's SIP server receives the request and consults the location service to find the location of the server. Most often the location service is able to find the address in the registration table. In some cases the location service can however only return an address of another redirect server. The location service returns the address to the redirect server, which returns it to the client's user agent (through its home SIP server) (4). The client's user agent confirms the response with an acknowledgement (5). The client's user agent now has the latest address of the server and is able to send its invite request to the server's user agent (6).

Alternatively, instead of redirecting the request (4-6), a proxy server could forward the request to the server (7). The server's user agent acknowledges the request (8) regardless of it has gone through a redirect or a proxy server, and the two nodes can begin communicating (9) [55].

The client now changes its position from location A in the home network to location B in a foreign network during the communication. To maintain the connection, SIP implements a two-fold location update. Figure 8 illustrates SIP mobility management.

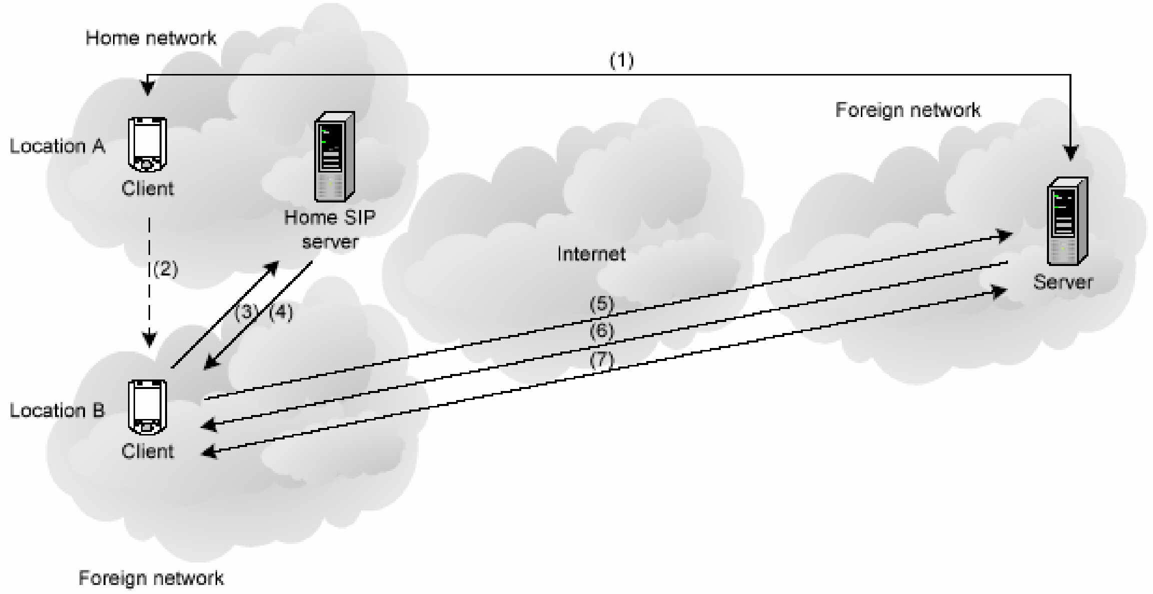


Fig. 34. SIP mobility management [56]

As the client moves from location A in the home network to location B in a foreign network during communication with the server (1) (2), it must update its location. Its user agent therefore sends a location update to the home SIP server so that new invite requests can be redirected correctly (3). The registrar processes the update message and forwards it to the location service, which stores the information. The home SIP server responds with an acknowledgment (4).

Then the client sends a new invite request to the server's user agent using the same call identifiers as in the original connection setup (5). The request contains the new address, which tells the server's user agent where it wants to receive future SIP messages. The server's user agent acknowledges the request (6) and the communication continues unaffected (7). [55].

Mobile IP has some shortcomings when it comes to delay-sensitive multimedia applications. The triangular routing adds handover delays and the tunneling overhead adds extra bytes to the packet header. It is more suited for long-lived TCP connections like telnet, ftp, etc. In comparison, SIP is much more suited for real-time communication over UDP. It is, however, less suited for TCP-based application. SIP is therefore often used either as a partially replacement for Mobile IP or as a complement, where SIP handles UDP connections and Mobile IP handles TCP connections. [55][56]. The table below summarizes and compares some major points between the various protocols that have been discussed.

Protocol	**Mobile IPv4**	**Mobile IPv6**	**mSCTP**	**SIP**
Layer	Network	Network	Transport	Application
Deployment requirements	(1) Home Agent in home network (2) Foreign agent in foreign network (3) Mobile IPv4 support by home agent, foreign agent and client	(1) Home Agent in home network (2) Mobile IPv6 support by home agent and client	(1) mSCTP is supported by client and server	(1) SIP server in home network (2) SIP should be supported by client and server as well as a SIP server

Table 2. Comparison between mobility protocols

Mobile IPv4 and Mobile IPv6 are both network layer protocols. They are also beneficial in the sense that the aid in furthering the vision of an all IP network. Mobile IPv4 however has various problems plaguing it such as triangle routing and security discussed previously which might not make it very attractive. It also requires new network equipment such as the home agent and foreign agent which makes its deployment costly. Mobile IPv6 is certainly an improvement over its predecessor as it solves problems such as triangle routing and security and also reduces deployment expenses as only a home agent is needed. However, it is not deployed as widely at the moment.

mSCTP is actually a Transport layer protocol which uses SCTP for communication and is designed to replace TCP and even UDP in the long run. Although it is better than both the protocols in many aspects, it is not widely deployed at the moment and hence might not be very easy to work with. The main benefit that mSCTP provides is that it does not require any new network equipment for deployment. All it needs is that the client and server support mSCTP which can be implemented in software.

Finally, as we discussed SIP is application layer protocol and its primary benefit is that it needs to be implemented only in the end hosts as SIP defines message exchanges between processes running on different hosts. However, a SIP server in the server's network is also required as part of the deployment for mobility to take place and establish communication. Another factor

As we can see, no one protocol can be considered as the best solution as each has its advantages and disadvantages. These pros and cons must be weighed against each other and the selection of the mobility protocol to be used may eventually depend on the best possible tradeoffs for a given scenario.

5 KFIA WLAN-Cellular Integration for Operators

The principle objective of this project undertaking was to develop a detailed integration architecture and interworking strategy. Cellular operators can then develop and enhance their infrastructure based on this ‘framework’, so as to tap into an important market which will continually grow into the future. KFIA has been selected as the first place to setup this technology and the interworking framework proposed has been developed with this in mind. However, the architecture and interworking framework proposed is not restricted to KFIA or airports alone but can be incorporated in all types of hotspot regions such as restaurants, hotels convention centers and so on. The only difference will be in the PWLAN deployment considerations and related aspects.

To this end, a complete interworking framework has been developed which addresses various aspects of accomplishing this integration between WLAN and cellular technologies. Initially, an overview of the integration framework is provided by discussing various components and interfaces. A reference model is presented that will form the basis of our discussion. This is followed by a detailed discussion of the integration architecture and discussion of how various technologies will inter-work together to provide mobility for the end user. In the next section, we enhance the reference model by adding mobility management and hence providing service continuity. In essence, a scenario 4 interworking architecture is developed. This is followed by a detailed discussion of common security related aspects of the interworking architecture for all the mobility management protocols. Thus, a common access architecture to support security is developed. Finally a brief discussion of the number of users and equipment that might be needed in establishing a PWLAN in KFIA and other KFIA specific information is briefly discussed and analyzed. Thus a complete framework for developing integrated WLAN network for the KFIA referred to as KFIA WLAN-Cellular Interworking Framework (KWIF) will be provided.

5.1 KFIA WLAN- Cellular Interworking Framework (KWIF)

Various types of interworking architectures have been developed and discussed in literature. A brief discussion on this was provided in the academic and vendor solutions sections. In adopting a reference architecture for our purposes, we choose an architecture that consisted of mostly standardized components and interconnections so as to allow the operator to easily integrate and operate these two networks.

The architecture adopted is illustrated in the figure below. This architecture is based primarily on Salkintzis's [30] scenario 3 architecture and the 3GPP interworking standardization Release 6 [59]. However, the design here does not take into consideration the roaming scenario as it is assumed that the operator will own the WLAN hotspots as well. Also we are trying to provide service continuity which is a scenario 4 requirement and has not been addressed by either of these publications. Consequently, further enhancements will be incorporated to allow mobility management which in turn will lead to service continuity or scenario 4 based integration. It should also be noted that since scenario 4 does not guarantee QoS, there is a chance of some services not surviving the handover process due to a large amount of packet loss.

Initially we will discuss various components in this reference architecture and then discuss the interconnections between these components. Hence, a comprehensive discussion of the reference model will be provided.

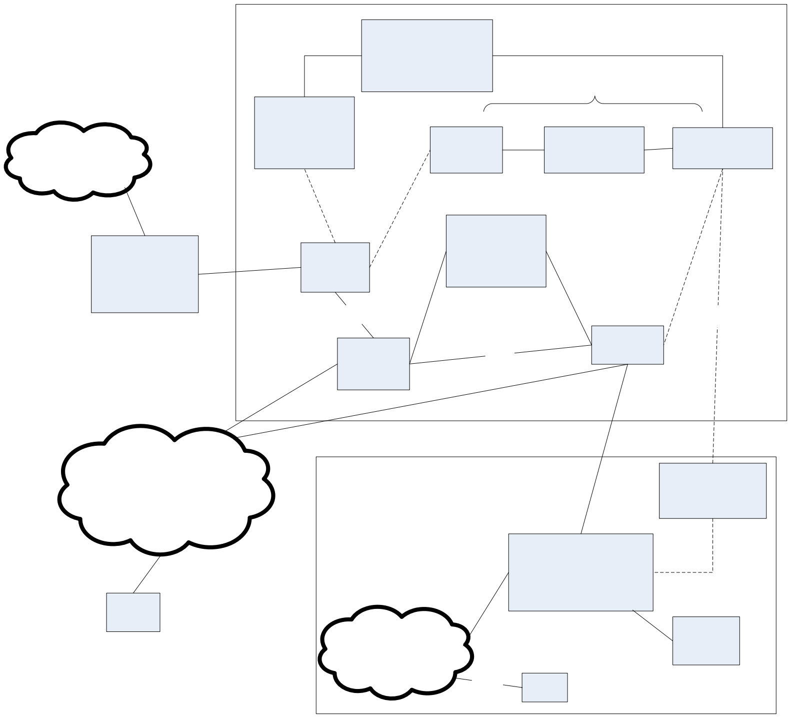


Fig. 35. WLAN-Cellular Reference Interworking Architecture

Network Components

The figure above illustrates an architecture where access to 3G-based packet services is provided [30]. A number of new components are added to operator's network to achieve this. Most of these components have already been standardized by the 3GPP [55, 59] or are used in WLAN networks separately. Some of these components are discussed in detail below.

WLAN Access Router Gateway (WARG): In order to provide access to 3G based PS services to the mobile node, data traffic to or from the WLAN goes through this gateway to the operator's network. It ensures that the packets are routed through the PDG. It also filters out packets based on the unencrypted information in them. Packets are only forwarded if they are a part of the existing tunnel or are expected messages from the mobile nodes in the WLAN. It implements routing policy enforcements after a tunnel is established.

Packet Data Gateway (PDG): Access to 3G based PS services are provided through the PDG. It contains routing information for WLAN-3G users and routes packets to/from the external packet data network from/to the end users. It performs address translation and mapping as well as encapsulation/de-capsulation of packets. It also has packet filtering functions to prevent unauthorized or unsolicited traffic. It also sends the mapping between a user identifier and a tunnel identifier to the AAA server and provides charging information related to user data traffic.

IP-MAP Gateway: This component will perform the translation between WLAN EAP or 802.1X and GSM SIM authentication mechanisms. This component basically establishes a bridge between the RADIUS-based authentication used in WLAN networks and the SIM-based authentication used in GSM networks. Thus, it behaves as another VLR from the HLR point of view and another RADIUS server from the 3G AAA server point of view. More details on this component and how it contributes to interworking security are discussed in 5.2.

3GPP AAA Server: All the AAA traffic from the WLAN comes to this server in the operator's network. It also communicates with the other 3G elements such as the HLR. The server gets subscriber information (authentication and authorization details) from the HLR and uses it to authenticate the subscriber. It also generates accounting information for each user that is reported to the charging system.

WLAN AAA Server: This server deals with AAA signaling for the WLAN only. It sends this information to the 3G AAA as discussed above.

Reference Points:

As can be seen in the figure above, various interconnections between different components exist. All the labeled interfaces have been standardized by 3GPP and implementation details for these interfaces and components can be found in [55, 59]. We briefly discuss some of these interfaces below that are important to the discussion of KWIF.

Wn: This is the interface between the WLAN and the WARG. Tunneled user data between these two is transported via the Wn interface.

Wg: This is an AAA interface between the 3GPP AAA Server and the WARG. It transports routing policy enforcement information to the WARG. Using this information, the WARG can identify which mobile node the data traffic belongs to and then apply the routing policy.

Ww: This is the interface connecting the mobile node to the WLAN according to the IEEE 802.1x specifications. It carries the authentication parameters exchanged between the mobile node and the 3GPP AAA server.

Wr/Wb: AAA signaling messages between the 3G network and the WLAN are securely transported over this interface. Authentication and authorization messages are carried over **Wr** while **Wb** transports the accounting messages.

Wp: This is the interface between the WARG and PDG that delivers tunneled user data between them.

Wm: This is the interface between the 3G AAA server and the PDG. Mainly, it allows the 3G AAA to get the tunneling attributes and a mobile node’s IP configuration parameters from the PDG. Charging information for 3GPP PS based services is also provided to the PDG by the 3GPP AAA via this interface. The authentication between the mobile node and the 3G AAA also requires message exchanges between the PDG and the 3G AAA that are carried over Wm.

Wi: This is similar to the **Gi** reference point provided by the packet switched domain. Packet Data Networks can work together via this interface based on IP

D'/Gr': This interface is located between the 3GPP AAA Server and the HLR. Its main function is to allow communication between the WLAN AAA and the HLR. It allows the authentication vectors to be retrieved from the HLR and registers the 3G AAA server of an authorized WLAN user in the HLR. It enables changing of subscriber profile within the HLR and can be used to retrieve service related information. It can also enable fault recovery procedures between the between the HLR and the 3GPP AAA server.

Gn': This reference point facilitates the implementation of the PDG by using a subset of the procedures provided by the Gn interface. This allows the interoperation between the PDG and the GGSN. The GGSNs do not have to be upgraded and therefore access from a WLAN is supported by reusing existing infrastructure.

There are two important aspects that will be discussed later. These include how access to the operator’s services can be secured so that only authorized users are allowed access. This will be discussed in detail in section 5.2. Secondly, to achieve service continuity mobility management has to be added to the network. This will be discussed with reference to three separate mobility management protocols: MIP, SIP and mSCTP in the subsequent sections. The mechanism of these protocols has been discussed in detail in section 4. In the next section we will discuss only how these protocols will provide mobility management and service continuity when incorporated in KWIF.

5.1.1 Mobile IP – based Mobility Management Architecture

In this section, an interworking architecture supporting scenario 4 that is based on a Mobile IP approach is presented. Initially, the new components that are added to enable MIP are discussed. This is followed by an explanation of the handoff procedure between the GSM network to WLAN and vice versa. The figure below depicts the interworking architecture of the Mobile IP based approach. Scenario 3 (access to 3G PS- based services) is supported by routing the user data to the 3G operators network which is also the user's home network. The requirement now is to provide Scenario 4 (access to 3G PS- based services with service continuity) support which is done by introducing the HA and FA (shown in red). This allows the same IP session to be maintained which consequently results in service continuity.

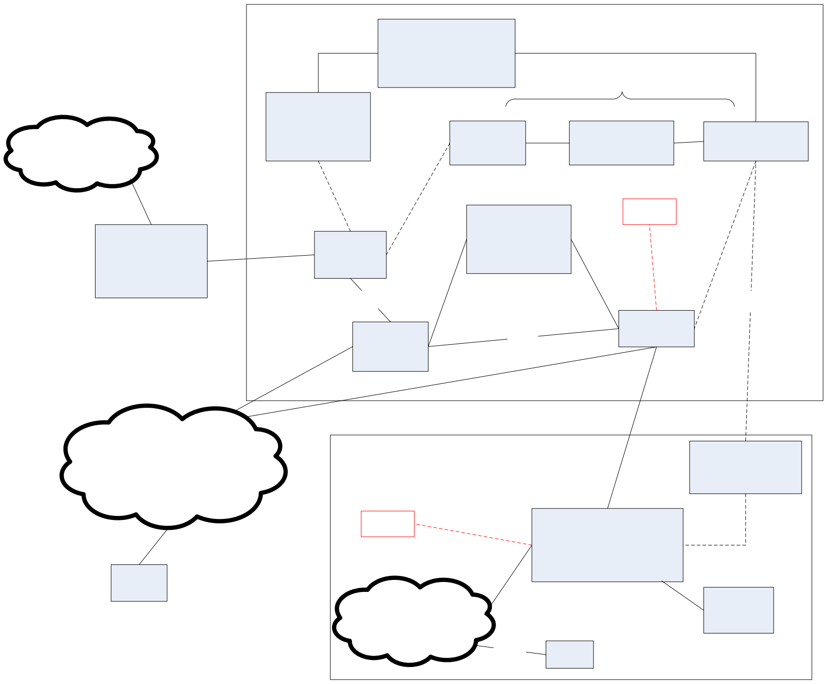


Fig. 36. Mobile IP based Interworking Architecture

The Mobile IP HA, which is a new component included specifically to allow mobility management using the MIP approach, is located in the operator's network and is connected to the PDG. Basically it performs mobility management functions for the mobile node when it is in its home network (the operator's network in this case). The other new component; the Mobile IP FA is located in the KFIA network and is connected to the WARG. It performs mobility management functions for the mobile node when it is in its foreign network (the airport network in this case).

The handover of a MN from the Cellular to the WLAN network or vice versa is the core function of any mobility management protocol and this is also true for MIP. The message exchange that takes place to achieve this is illustrated in the figures below. The handover procedures have been developed using [33], [44] and [43] which discuss various aspects of Mobile IP and its interworking in a cellular network. The important thing to note is that a tunnel is created between the PDG and the UE to transport user data over the network to the internet or vice versa. The HA and FA help in establishing this tunnel and in routing this traffic. The authentication procedures for MN are discussed in detail in section 5.2 and thus are not elaborated here.

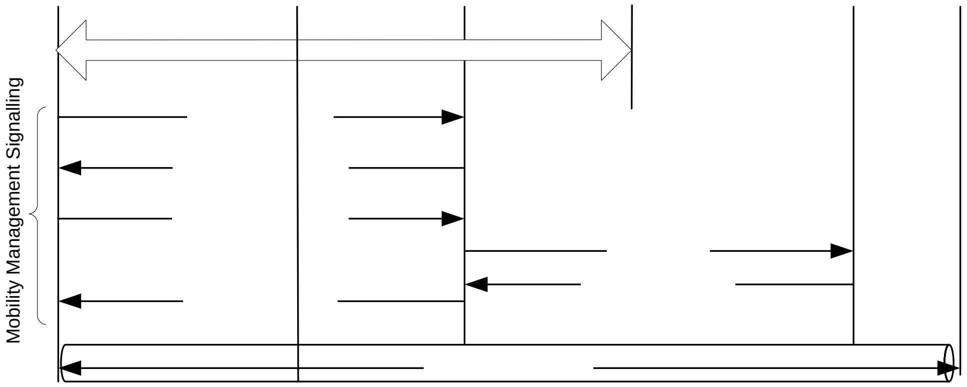


Fig. 37. Cellular to WLAN handoff for Mobile IP

Here, mobile IPv4 has been applied although mobile IPv6 can also be applied with minimal changes. Two different handover scenarios are illustrated. The figure above depicts the Cellular to WLAN handover. We assume the operator’s network is the MN’s home network. Initially, the MN is sending or receiving data packets from a Cellular network using GPRS for example. Once the MN decides or the network decides to handover to a WLAN based on the signal strength or some other algorithm, it starts with a set of handover procedures which only involves the physical and link layer at first. It will also be authenticated and authorized by the operator’s network at this stage. This will include communication with AAA servers as well as operator’s HLR and AuC if necessary. Section 5.2 discusses various details related to this aspect. After authentication the MN can be assigned an IP from the foreign networks DHCP server. Also the MN will be detached from the operator’s cellular network.

However, this can be avoided by sending PDP connect standby message to its SGSN. This will help to reduce the re-attach effort on the MNs’ part.

After these initial steps, MN can access the IP network and sends an *Agent Solicitation* to locate a local FA. The local FA replies to the MN, and then the MN can send *Registration Request* to its home agent through the FA. After updating the CoA in the HA, the packets sent to the home network will be forwarded to the WLAN network.

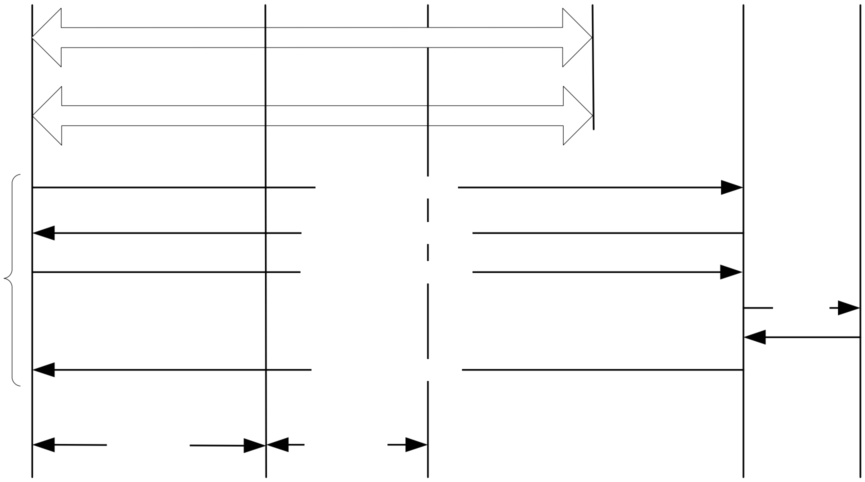


Fig. 38. WLAN to Cellular handoff for Mobile IP

As for the WLAN to Cellular handovers, it is similar to the case discussed previously. The figure above illustrates this scenario. The only difference is that if the MN did not attach the network or activate the PDP session before, it should attach to the network and activate a session before getting access to the cellular network. As mentioned earlier, if the MN has already had a session which has not timed out or it sends a PDP/MM context standby message to the UMTS before, the original PDP session can be kept using. Another point to note is that the HA will send an update message to the FA to delete the entry that included the MN’s MIP association. The AAA signaling between the WLAN network and the operator’s core network is discussed in detail in section 5.2.

MN

5.1.2 SIP – based Mobility Management Architecture

The figure below depicts the interworking architecture of the SIP based approach. As before Scenario 3 is supported by routing the user data to the 3G operator's network. Scenario 4 support is done by introducing a new entity which is the SIP server (shown in red). This allows the same IP session to be maintained thereby resulting in service continuity. The SIP server is located in the operator's network and is connected to the PDG.

SIP is capable of providing node mobility during a session when a MN moves from one network to the other. This is done by informing the CN of the MNs new IP address and session parameters via a SIP INVITE message. As soon as the CN gets the INVITE message, its starts sending the data to the new network. However, before the INVITE message can be sent, other procedures are needed to attach the MN to the network. These would be the GPRS attach and PDP context activation procedures in case of the cellular network and DHCP messages in case of the WLAN network.

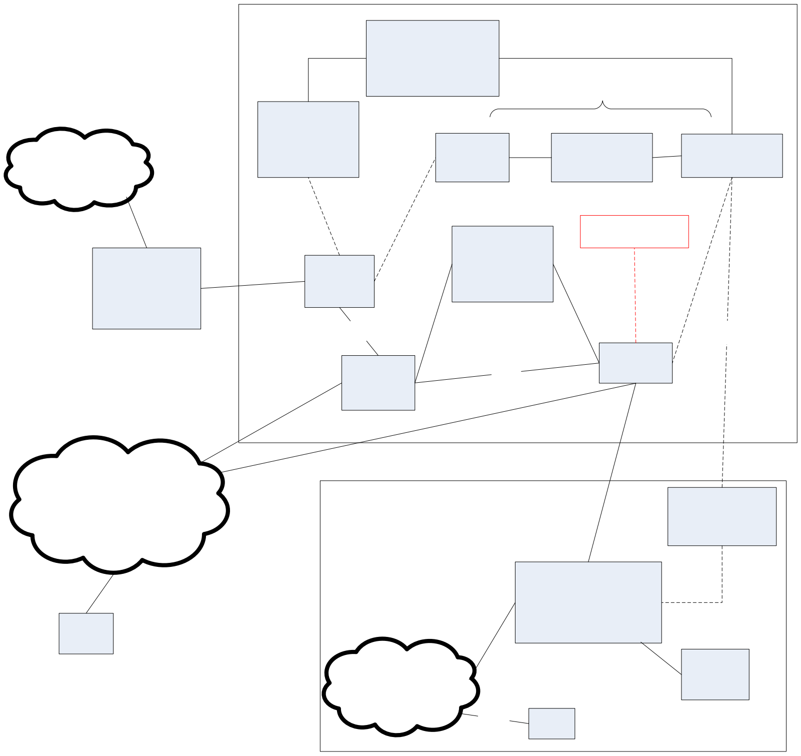


Fig. 39. SIP Based Interworking Architecture

The handover message exchange between the Cellular network and WLAN and vice versa is illustrated in the figures below. The handover procedures have been developed using [60] which discuss various aspects of SIP and its interworking in a cellular network.

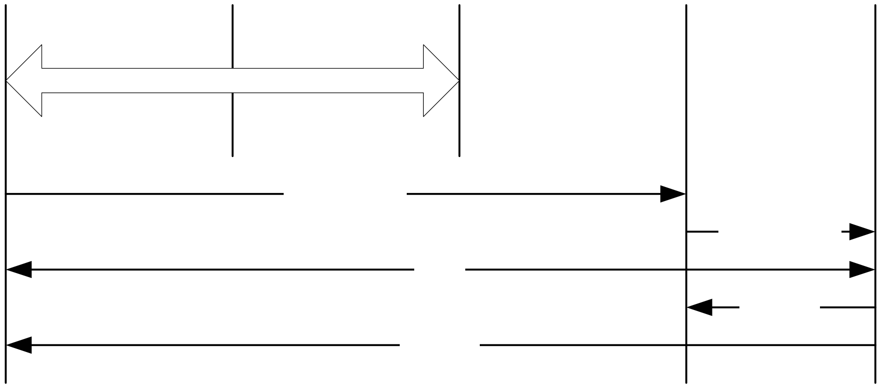


Fig. 40. Cellular to WLAN handoff for SIP

When a MN moves to the WLAN, it must first get a new IP address. Although not shown in the figure, this is done by the DHCP server. The MN will then be authenticated and authorized by the operator's network. This requires communication with AAA servers as well as operator's HLR and AuC if necessary. A more detailed explanation of these aspects is provided in 5.2. As discussed above, the MN will now send a SIP INVITE message (via the SIP server) to the CN informing it of the new IP address. The CN will then send the data to the WLAN and a SIP OK message after negotiating various network parameters.

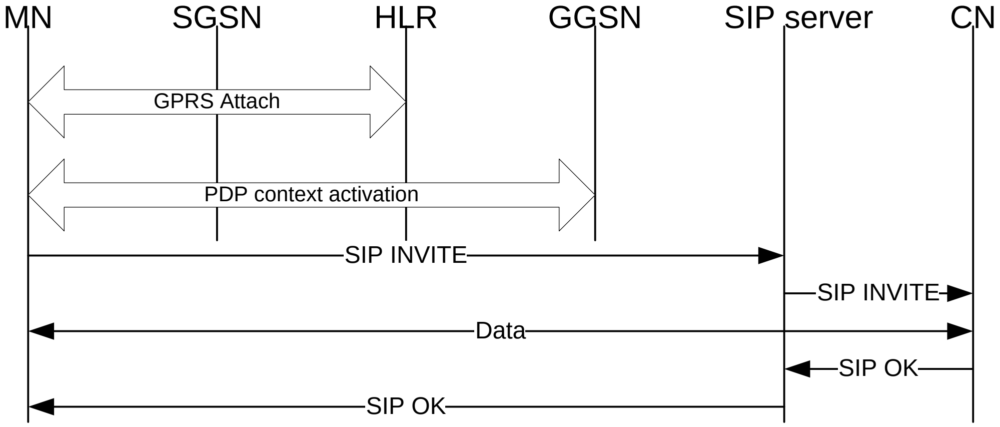


Fig. 41. WLAN to Cellular handoff for SIP

When a MN moves to the cellular network, it must first attach to it and activate a PDP session to begin receiving packets. As before, the MN then invites the CN to its new address and the CN replies by sending an OK message and negotiates network parameters before sending user data.

5.1.3 mSCTP – based Mobility Management Architecture

The figure below depicts the interworking architecture of the mSCTP based approach. As before Scenario 3 is supported by routing the user data to the 3G operator’s network. However, Scenario 4 support is provided without introducing any new components to maintain service continuity.

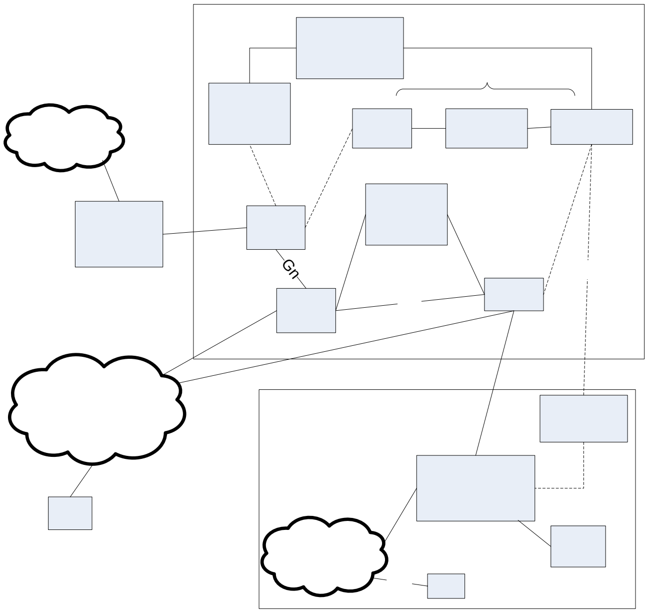


Fig. 42. mSCTP Based Interworking Architecture

The handover message exchange between the Cellular network and WLAN and vice versa is illustrated in the figures below. The handover procedures have been developed using [61] which discuss various aspects of mSCTP and its interworking in a cellular network. The MN and CN both implement mSCTP. For the handover in each direction, the procedure uses three basic steps: add IP address, vertical handoff triggering (set IP address), and delete IP address. This addition, setting and deletion of IP addresses during an ongoing SCTP association takes place via address configuration (ASCONF) messages [61].

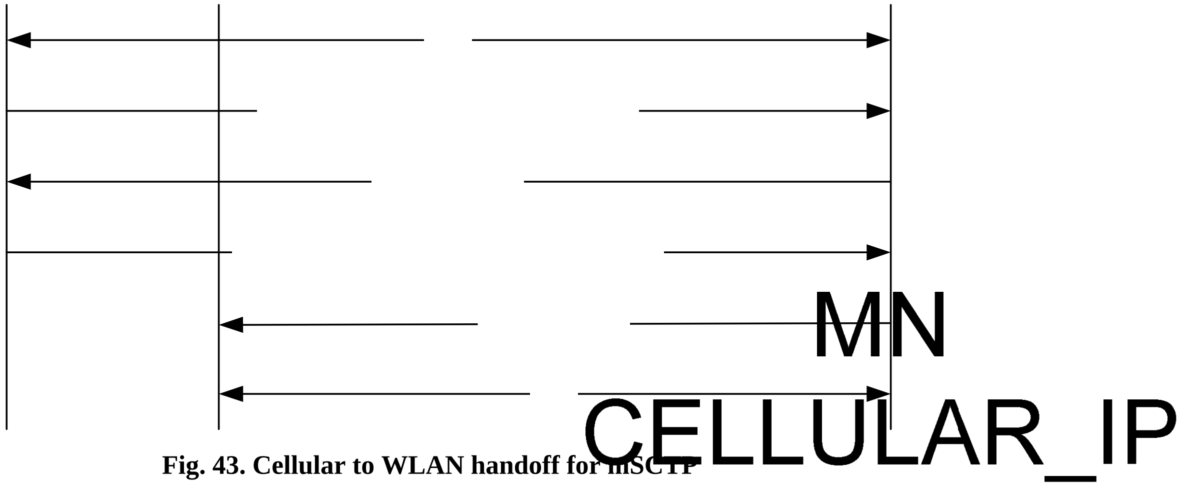


Fig. 43. Cellular to WLAN handoff for mSCTP

In the above figure, the MN with IP address ‘Cellular_IP’ is communicating with the CN (IP address: CN_IP) using mSCTP. Now, after moving into the WLAN area, the MN is given a new IP address (WLAN_IP) after authentication from a DHCP server. The MN then starts the IP addition process by sending an ASCONF message to the CN with “add IP address” and WLAN_IP as parameters. The WLAN_IP is now added to the SCTP association. However, the data is still being received from the cellular network i.e. the cellular interface is still active.

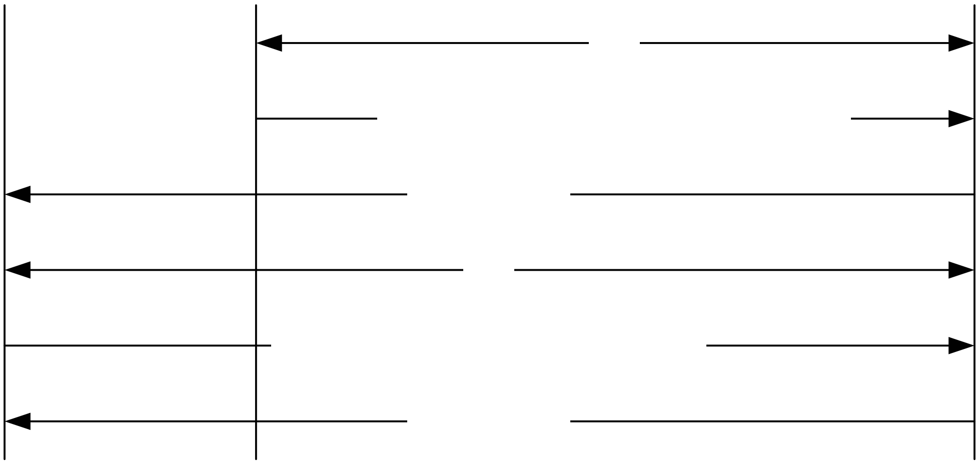


Fig. 44. WLAN to Cellular handoff for mSCTP

Now the signal strength of the WLAN increases and in order to trigger the handover, the MN then sends an ASCONF message with “set primary address” and WLAN_IP as parameters. When the CN replies with an acknowledgement (ACK), the data is routed through the WLAN and the handover process is complete.

In this case also, the MN triggers the handover by sending an ASCONF message "set primary address" with Cellular_IP as parameters. After receiving the ACK from the CN, the MN is now associated with the cellular network through which the data is now routed. Now as the signal from the WLAN becomes weak, the MN begins the IP deletion process by sending to the CN an ASCONF message "delete IP address" with WLAN_IP as parameters. After CN replies with an ACK the WLAN_IP is deleted from the SCTP association.

The above procedures are for single-homing configuration of mSCTP i.e. the CN is assumed to have a single fixed IP. In cases where the CN can have two or more IPs, the procedures are similar. For dual-homing configuration, it has been seen that an overall improvement in the delay is observed. A more detailed explanation of this can be found in [61].

5.2 WLAN-Cellular Integration Security

Providing users access to WLAN hotspots in a WLAN-3G interworking environment requires an access model that should be able to address several aspects. It should use secure authenticating procedures to allow the operator to protect user credentials and data transfer over the wireless connection. AAA functionality in a network by using AAA server should be included to facilitate the management and delivery control needed to offer customized services that can increase customer loyalty and generate revenue. The protocols and authentication schemes used to communicate these aspects should also be specified. The access model can also specify whether users pay on-demand or are billed against an existing provider account or allow both options. [29]

In this section, a WLAN-Cellular access architecture is proposed for the KFIA interworking architecture referred to as KFIA WLAN-Cellular Access Architecture (KWAA). The KWAA will only be responsible for AAA management which is a key element in an integrated WLAN-Cellular network. Mobility management and other features in an interworking scenario will not be addressed. In essence, it will verify a claimed identity (Authentication), determine whether a particular right can be granted, for example, an access to a resource, to the presenter of a particular credential (Authorization) and collect resource consumption data for the purposes of performing trend analysis, capacity planning, billing, auditing, and cost allocation (Accounting). These principles are applicable to all kinds of networks. [57]

However, Cellular and WLAN networks have different databases and AAA procedures. The proposed architecture takes this into account and reuses existing standards for user authentication, access control and billing which ensures that the architecture is standards based allowing apparent advantages for operators. It also offers a key advantage of being able to work with any underlying mobility management protocol such as MIP, SIP or mSCTP. In the following sections, we detail various features of KWAA.

5.2.1 KFIA WLAN-Cellular Access Architecture (KWAA)

KWAA basically integrates access of the WLAN network into the operators GPRS network. This is done by allowing AAA messages to be transported from the WLAN network over IP network to the operator’s core network for AAA purposes. The AAA protocol used here can be either DIAMETER or RADIUS both of which have been discussed in detail in previous sections. However, for this architecture RADIUS will be used due to its wider acceptance and product availability even though DIAMETER has many advantages over it. This will also allow interworking with legacy systems. Nevertheless, RADIUS alone is not sufficient. Wireless interface security between MN and the WLAN network should also be implemented which is provided by 802.1X based protocols and Extensible Authentication Protocol (EAP). A number of protocols have been developed such as EAP-SIM, EAP-AKA, EAP-TTLS and so on that provide complete security against known hazards of wireless computing.

The 802.1X is a Port-Based Network Access Control standard that uses the Extensible Authentication Protocol (EAP) between the MN and access point to perform per session user authentication. Consequently, a user is not allowed access to the public network until he or she has been successfully authenticated against a customer database where authentication can be based on user name and password, SIM-based credentials or an operator allotted certificate. EAP messages are encapsulated using EAPoL (EAP over LAN). On a wireless link, it is referred to as EAPoW (EAP over Wireless). These EAP messages are then decapsulated at the access point, and then re-encapsulated using RADIUS. IEEE 802.1x uses standard security protocols, such as RADIUS or DIAMETER, to provide centralized user identification, authentication, authentication, dynamic key management, and accounting. [29]

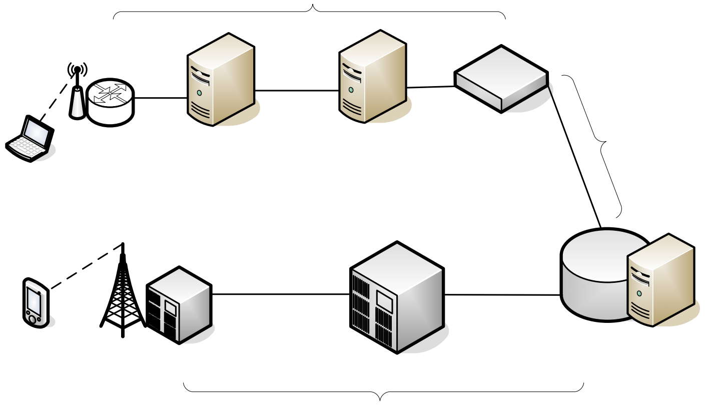


Fig. 45. KFIA WLAN-Cellular Access Architecture (KWAA)

The figure above illustrates KWAA. The lower branch depicts the authentication route followed by MNs being authenticated by the GSM network. The WLAN network integration is shown in the upper branch. Interfaces between each of these components have been standardized by the 3GPP and will be discussed in detail in earlier. The WLAN Access Router Gateway (WARG) is connected to the WLAN AAA server which is located in the WLAN network. The WLAN AAA is in turn connected to the 3G AAA server located in the operator's core network. However, to connect this to the operators' HLR and AuC, a translator has to be inserted so that the 3G AAA server can communicate with the HLR/AuC. [29, 30]

This is accomplished by a separate entity know as an IP-MAP gateway. This component serves an important purpose in this architecture. It principally acts as a translator by converting RADIUS messages on an IP network MAP messages on an SS7 network. This allows the reusing of existing subscriber database and hence allows users to receive a common bill as well as reduce the number of databases for storing subscriber information. Thus the operator, when implementing WLAN hot spots into the network, can continue using the subscriber and authentication infrastructure that is already in place for cellular services. [29]

5.2.2 Access Authentication Schemes

The KWAA is proposed to consist of two authentication schemes: web-based and SIM-based. The KWAA incorporates both of these schemes. When an MN moves into a PWLAN the access authentication scheme for this can be of two types. The Web-based scheme [29] is used only for those users who are unable to authenticate themselves using SIM-based authentication. This would include “walk-in” users who would like to avail of the PWLAN for a short while. These users can get access to the network by buying a prepaid card to access the facility. For users who are already registered with the operator and have access to a WLAN-Cellular interworking service, SIM-based authentication [29, 56] is used.

This decision for selecting the authentication scheme could be implemented in the following way. Once a WLAN AAA requests the MN for its IMSI using EAPoW, two cases are possible. Either the MN will respond with its IMSI in which case the decision will be taken to go through SIM-based authentication route. On the other hand, if IMSI is not received then the user will be given the option to access the network using web-based authentication. This functionality is proposed to be implemented in the WARG so as to keep traffic related to this selection process confined to the WLAN access network. Alternatively the WLAN AAA can also be involved. The flowchart below demonstrates how this decision would take place. In the following sections we detail the SIM-based and Web-based authentication schemes.

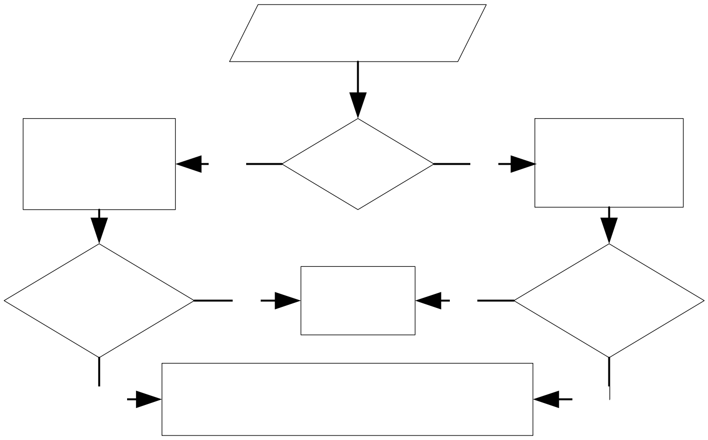


Fig. 46. Authentication Scheme Selection Tree

5.2.2.1 SIM-based authentication

The SIM-based authentication is based on an IETF draft [56] which is a work still in progress. The protocol proposed for this uses a combination of Extensible Authentication Protocol (EAP) and 802.1x protocols and is known as EAP-SIM. This provides security over the wireless interface in a similar way as is done in the GSM security architecture by reusing existing subscriber information stored in the operator's network.

Since complete authentication with EAP-SIM in this way takes time due to the use of the A3/A8 algorithms, EAP-SIM also includes a provision for fast re-authentication [56] which is done using the vector triplets received in a previous full authentication. Authentication will occur much more quickly and without involving the HLR/AuC at every authentication procedure. This will reduce the time required for authentication and reduce the handoff delay overhead. Of course there will have to be an upper limit to the number of fast re-authentications allowed before a full authentication has to be done again. This option will be of particular benefit to those MNs that have to move in and out of WLAN coverage frequently.

The figure on the next page illustrates the message exchange that would take place during such an authentication procedure. The SIM card is used to provide user credentials (IMSI) as well as to calculate SRES based on the operator specific implementation of the A3/A8 algorithms stored in it. The important thing to note is that EAP-SIM provides authentication MNs without compromising the GSM security architecture. [56]

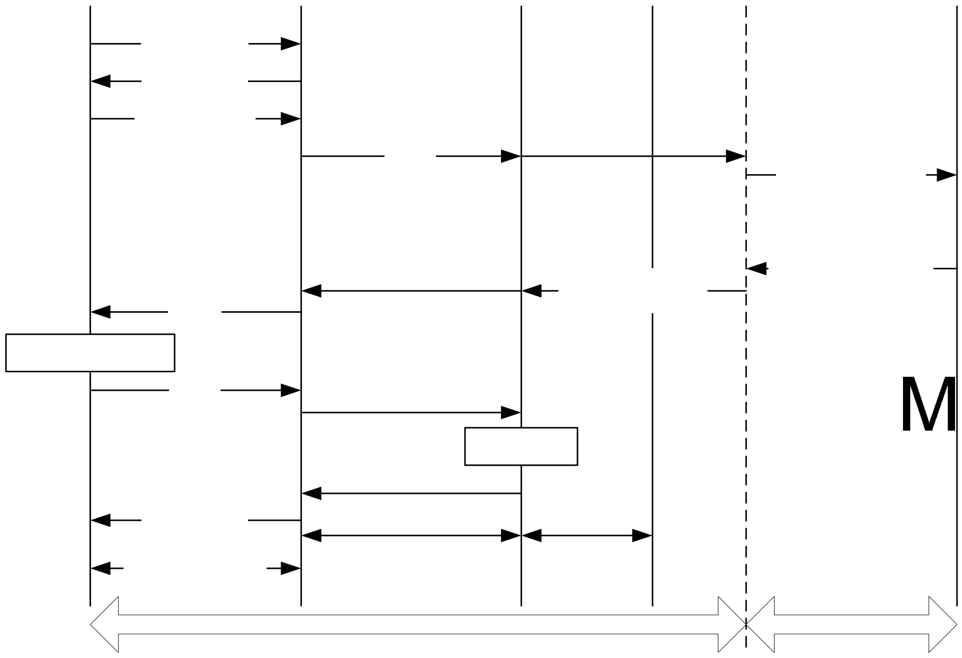


Fig. 47. SIM-based Authentication message exchange

It is worthy to note that the IP-MAP gateway acts as an interface between IP and SS7 networks. This component, as discussed earlier, translates RADIUS messages on the IP network into MAP messages in the SS7 network [29]. Thus it provides common access to the subscriber database (HLR) and also reuses the GSM authentication scheme by communicating with the AuC. Another important benefit of using this method is that it is very secure and if re-authentication is also implemented, it will prevent attackers from hijacking user sessions [29].

5.2.2.2 Web based authentication

This type of authentication utilizes a browser for authentication purposes. A user can be given a prepaid scratch card with a username and password. The MN would initialize a browser and receive a login page to enter this information to get access to the network. The WLAN AAA server validates the request using RADIUS. Once the authentication is complete, the WARG will deliver accounting information continually until the session is terminated. [29]

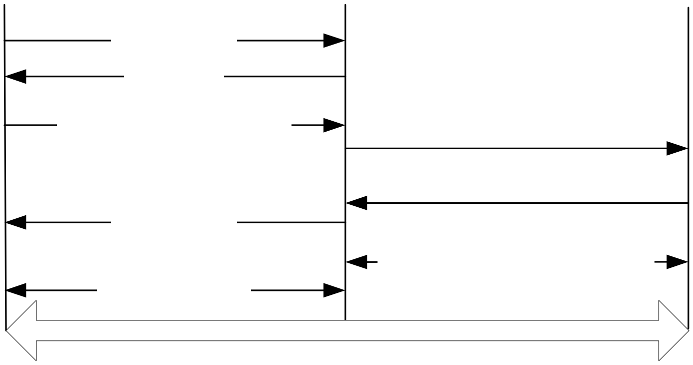


Fig. 48. Web-based Authentication message exchange

Note that in this scheme the entire process can occur over the IP network as the HLR is not involved. The figure above illustrates this method by specifying the message exchange that takes place in KWAA. Another important comparison with the SIM-based authentication scheme is that there is a low level of security as EAP based protocols is not utilized. A summary of some of the differences between the two authentication methods is given in the table below.

	SIM-based authentication	**Web-based authentication**
Device Needed	No extra device	SIM card reader
Login security protocol	Https	EAP-SIM
Complexity	Low	High
3G-WLAN handovers allowed	No	Yes
User access procedure	Username and password are typed, user is involved	Seamless access, user need not be involved
Overall security	Low	High

Table 3. Comparison of SIM-based and Web-based Authentication Schemes

Consequently, KWAA authentication scheme is recommended to consist of a hybrid version that incorporates both these authentication schemes as discussed previously. The SIM-based authentication (EAP-SIM) is used for users that are already subscribed with the operator network. Whereas the web-based authentication is used for allowing unregistered users access to the PWLAN as any other hotspot facility. However handoff capability and session continuity will be available only to registered users due to obvious reasons.

To simplify the authentication process it has been assumed that all users subscribed to the operator’s network are also subscribed to the WLAN-Cellular interworking service. This can be accomplished by allowing free access to this service for subscribed users. This is justifiable because the operator will be able to generate more revenue as people will tend to spend more time accessing these networks due to their high performance and low cost. This will also reduce the load on the GPRS network by allowing a unique way of load balancing. Another benefit is that this would act as a value added service for GPRS users, which will attract new customers and keeps old customers loyal.

As a result of implementing the proposed WLAN-Cellular Access Architecture for KFIA, users registered with the operator’s network as well as “walk-in” users who would like to use the PWLAN service at KFIA are targeted. This strategy allows the operator to fully utilize its resources and increase its revenue base. It should also be noted that this architecture is a separate entity in itself in that it can be installed separately and does not depend on the underlying management protocol being used in the network. Hence, the KWAA is a common feature in MIP, SIP and mSCTP-based architectures.

5.3 Protocol Architecture and Packet Traffic Study

The protocol architecture is an important component needed to implement functionalities into the various entities that make up the interworking structure. It also specifies how interfacing between various entities takes place and which protocols are involved at different layers of the network. In this section, a simplified version of the protocol architecture is presented. For the sake of brevity, this section focuses on the AAA protocol architecture or the *Signaling Path* and the *Data Path*.

The signaling path protocol stack is designed to transport information that performs AAA functions. The figure below depicts a protocol architecture that applies to the interworking scenario discussed in the previous sections. It is based on a similar discussion in [62] and more specifically applies to the KWAA and how the access architecture can be implemented. Note that the method used to authenticate the MN; in this case EAP-SIM needs to be implemented only at the end station and 3G AAA server. The intermediate nodes such as WARG need to support only the generic implementation of EAP and EAPoL as specified in 802.1x.

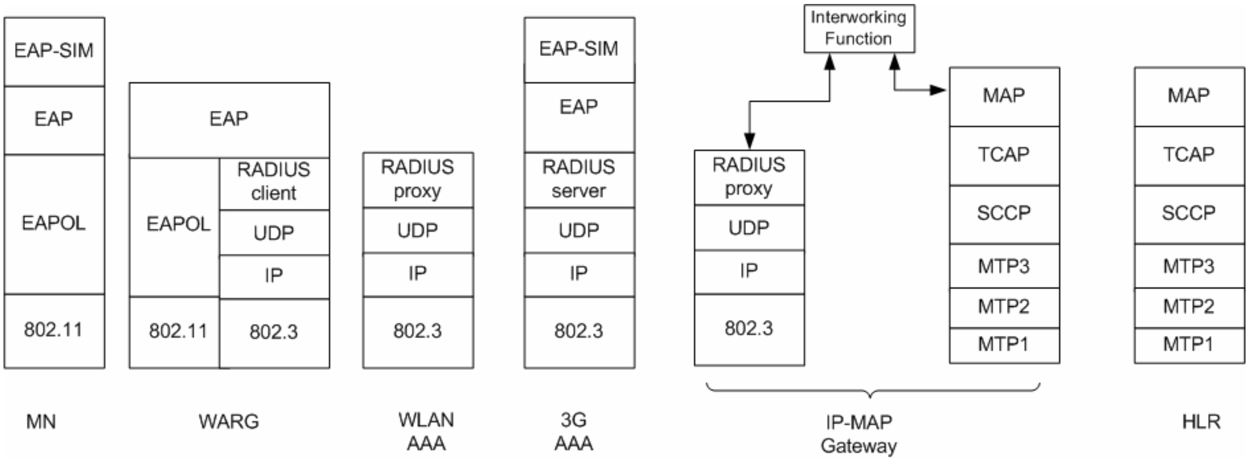


Fig. 49. Signaling Path Protocol Architecture

Another important thing is the interworking function in the IP-MAP gateway. As can be seen from the figure, it provides translation from the RADIUS/IP stack to the MAP/SS7 stack. This allows the operator to unify the authentication process at the HLR as discussed previously.

Similarly, a protocol architecture for the Data Path is developed. This is the path that data packets emerging from and going to the MN have to traverse through. This architecture was based on recommended 3GPP standards which are discussed in [59]. The Remote IP layer is used by the MN to be addressed in the external packet data networks. The lower IP layers are basically the Transport IP layer that is used to address the MN within the WLAN network and to transport the upper layers to their destination. For example, the transport IP could be a private IPv4 address assigned using a NAT. GTP is used as the tunneling protocol of choice to provide end-to-end tunneling between MN and PDG. Note that the PDG is responsible for decapsulating and encapsulating IP Packets it receives from outside based on the protocol stack implemented.

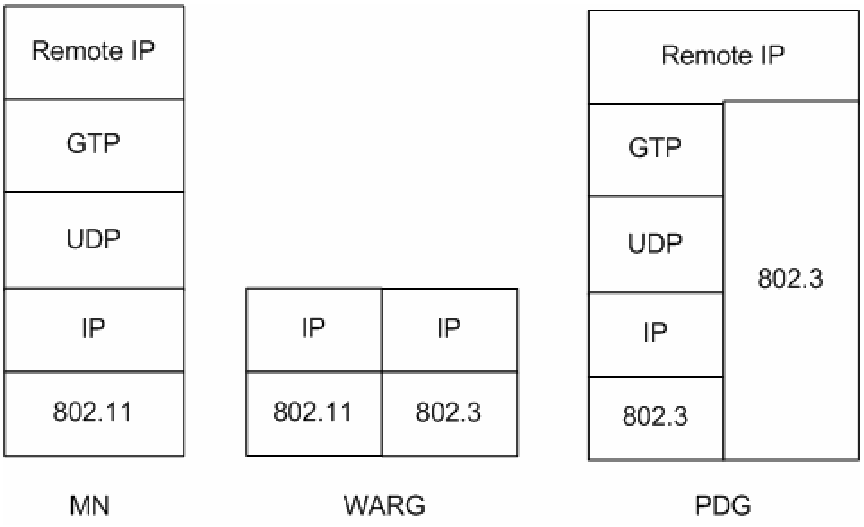


Fig. 50. Data Path Protocol Architecture

Packet Traffic Study

The purpose of this study was to estimate the amount of traffic being created between the various nodes. The information in this study can be further used to develop models that can be used to analyze the performance of the proposed interworking architectures. Both the Data Path as well as the Signaling Path protocol architectures were considered to calculate the traffic generated per user per session. The table below lists some protocols and their respective header lengths which were obtained from [63].

Protocol	Header Length	Protocol	Header Length
UDP	8 bytes	IP	20 bytes
TCP	20 bytes	802.11	34 bytes
GTP	20 bytes	EAPoL	6 bytes
RADIUS	23 bytes	802.3	26 bytes

Table 4. Protocol and their respective header lengths

It should be noted that both 802.3 and 802.11 frames must carry a maximum payload of 1500 bytes to maintain compatibility even though both can support varied lengths of payload. The payload is in addition to their on overhead bytes in the MAC layer. For example 802.3z has a seven byte preamble, four byte frame check sequence and six bytes each for the source and destination address as well as one byte start-of-frame delimiter giving a total of 26 bytes. A similar exercise with 802.11 gives an overhead total of 34 bytes. [63]

The application level messages that make up the body of these packets for EAP-SIM was obtained using the EAP-SIM IETF draft [52]. The figure below illustrates this.

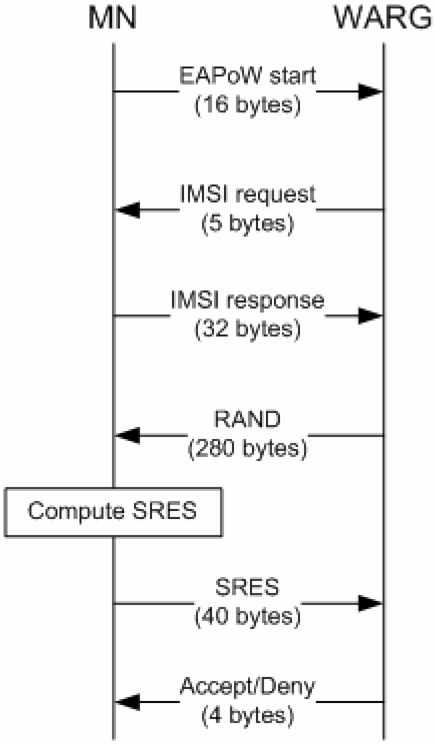


Fig. 51. EAP-SIM message exchange with length of messages

It can be seen that for the signaling path, even after adding the overhead costs for each message as well as the length associated with each EAP-SIM message. The total comes out to be less than 1500 bytes. This is true for other situation as well where communication takes place between WARG and WLAN AAA and WLAN AAA and 3G AAA. However, the minimum size for each frame has to be 1500 bytes and hence this size is used for each message exchange. The figure below presents the amount of control traffic that traverses between various components to authorize a single user in each session.

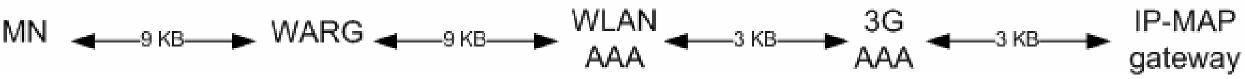


Fig. 52. Data flow between various components

In the case of the data path, we use the protocol stack in the figure shown previously to see that a total header length of approximately 100 bytes. The header overhead includes the headers of protocols such as Transport IP, UDP, GTP, and Remote IP and so on. However, since the minimum packet size is once again 1500 bytes. The effective user data being transferred is approximately 1400 bytes per packet.

5.4 PWLAN Deployment

In this section we will delve into the details of providing WLAN access to KFIA. In particular, we will estimate the number of users that the WLAN will have to handle and follow up with a numerical analysis of this estimate to find how much equipment is required. The following is a list of variables that we have defined and used in our calculations.

NAME	DEFINITION
$\mathbf{U_{PEAK}}$	The number of passengers on a peak day
$\mathbf{P_R}$	Penetration ratio
$\mathbf{U_{WiFi}}$	The number of users that have WiFi capability
$\mathbf{D_{USER}}$	Data rate assigned to each user
$\mathbf{U_F}$	The number of users per floor
$\mathbf{D_{AP}}$	The data rate provided by an access point (AP)
$\mathbf{N_{AP}}$	The number of access points
$\mathbf{A_F}$	The area of a floor
$\mathbf{A_{AP}}$	The coverage area of an AP
n	The number of users served by a single AP
α	Efficiency factor

Table 5. Variable Names & Definitions

5.4.1 User Analysis

$\mathbf{U_{PEAK}} = 38{,}000$ [58]

We assume that $\mathbf{P_R} = 20\% = 0.2$

$\therefore\ \mathbf{U_{WiFi}} = \mathbf{P_R} \times \mathbf{U_{PEAK}} = 38{,}000 \times 0.2 = 7{,}000$

We assuming here that we are providing each user with 512 kbps $\Rightarrow$ $\mathbf{D_{USER}} = 512$ kbps

We are providing WLAN access on three floors; Mezzanine, Departure and Arrival. We are assuming a break down of 20%, 40% and 40% for the above mentioned floors respectively.

$\therefore$ The number of users on the Mezzanine floor = 20% × $\mathbf{U_{WiFi}}$ **=** 1,400

Similarly, the number of users on each of the Departure and Arrival floors = 40% × $\mathbf{U_{WiFi}}$ **=** 2,800

5.4.2 WiFi Deployment Analysis

We assume that all access points are configured at their maximum speed of 54 Mbps ($\mathbf{D_{AP}}$). However, because of factors like interference etc. an AP will not always perform at 100% efficiency. To compensate for this we introduce an efficiency factor α that ranges between zero and one. The equation below represents a simple way of estimating the number of access points ($\mathbf{N_{AP}}$) required given certain parameters. This is followed by some sample calculations.

$$N_{AP} = \text{maximum}\left(\left\lceil \frac{Uf \times Duser}{\alpha \times Dap} \right\rceil, \left\lceil \frac{Af}{Aap} \right\rceil\right)$$

In the ideal case, an AP functions at 100% efficiency (α = 1) i.e. every access point is capable of providing 54 Mb/s. In reality, the available bandwidth is a function of many parameters; the most important being the number of stations served by that AP. Taking this into consideration we have come up with the following derivation.

The curve below shows α as a function of n, the number of terminals in a certain service area.

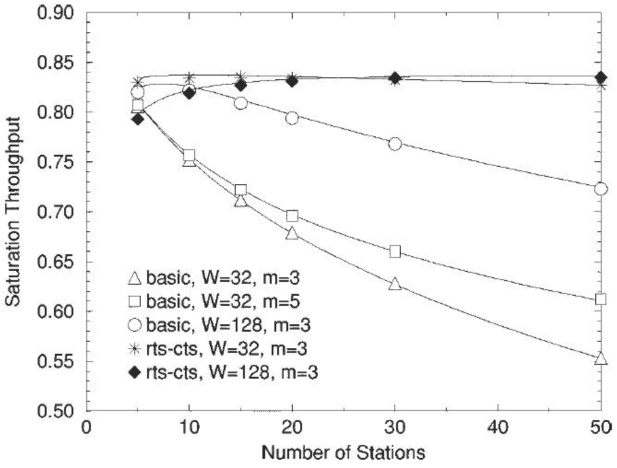


Fig. 53. Alpha vs n [64]

Curve fitting for this graph results in the following equation between α and n

$$\alpha = 0.9853 - 0.1047 \ln(n) \qquad \ldots (1)$$

Also, from above

$$n = \frac{Uf}{Nap} \qquad \ldots (2)$$

$$\frac{Uf}{Nap} = \frac{\alpha \times Dap}{Duser} \qquad \ldots (3)$$

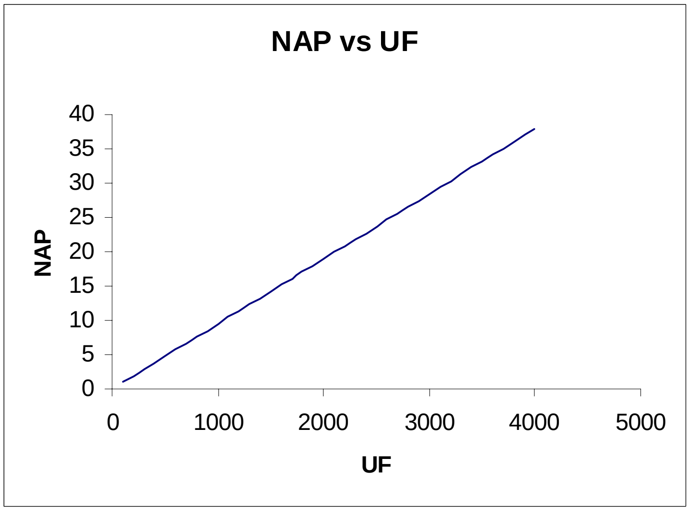
Therefore from (1), (2) and (3)

$$\alpha = 0.9853 - 0.1047 \ln\left(\frac{\alpha \times Dap}{Duser}\right)$$

In the Access Point Estimator (see Appendix 8.1) we solve this equation to determine α given a certain D_{AP} and D_{USER}. We then use this α to determine N_{AP}. APE can of course accommodate a different equation if a model other than the one used here [64] needs to be used.

For the user analysis figures, D_{USER} and D_{AP} values discussed above the estimator gives an AP efficiency of approximately 55% i.e. $\alpha = 0.55$

The following three graphs show how $\mathbf{N_{AP}}$ varies with $\mathbf{U_F}$, $\mathbf{D_{USER}}$ and $\mathbf{A_F}$ respectively.



$\mathbf{D_{USER}}$ = 512 kbps
$\mathbf{D_{AP}}$ = 54 Mbps
$\alpha = 0.55$

Fig. 54. N_{AP} vs U_F

In the following graph, D_{USER} varies and α varies with it accordingly as shown.

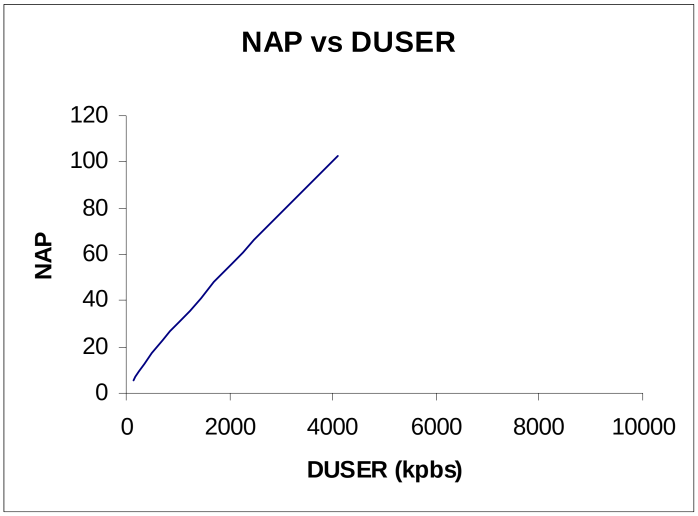
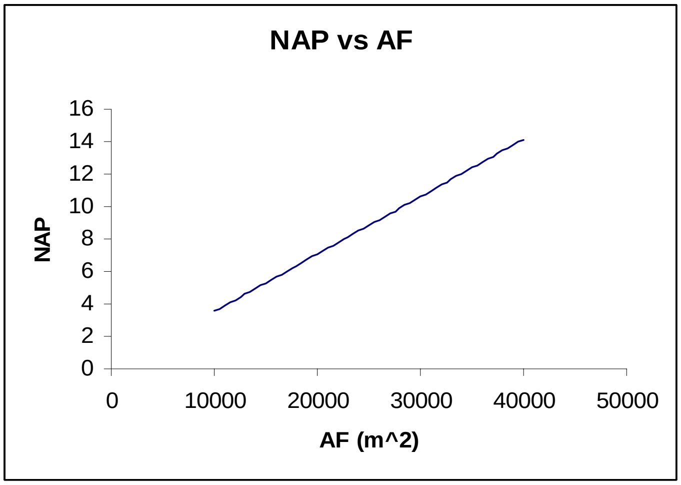


Fig. 55. N_{AP} vs D_{USER}

$\mathbf{U_F}$ = 1,000
$\mathbf{D_{AP}}$ = 54 Mbps

$\mathbf{D_{USER}}$	α
64	0.38
128	0.43
256	0.49
512	0.55
1024	0.62
2048	0.68
4096	0.74
8192	0.80



$\mathbf{A_{AP}}$ = 2,827 m^2
α = 1

Fig. 56. N_{AP} vs A_F

$\mathbf{A_{AP}}$ **=** 2,827 m^2, calculated as ($\pi \times R^2$) where R is the coverage radius of an AP (30 m @ 54 Mbps)

A set of simulations was conducted using APE to study the traffic between KFIA and the operator's core network. The rationale to focus on the PDG – WARG link and the WLAN AAA – 3G AAA link is that these links tend to be the most expensive in terms of both the distance covered and the cost of leasing a link. The simulations were done with the similar parameter values as used above and the results from section 5.3.

The graph below depicts the change in signaling traffic generated per second between WLAN AAA and 3G AAA as a function of the number of sessions per second.

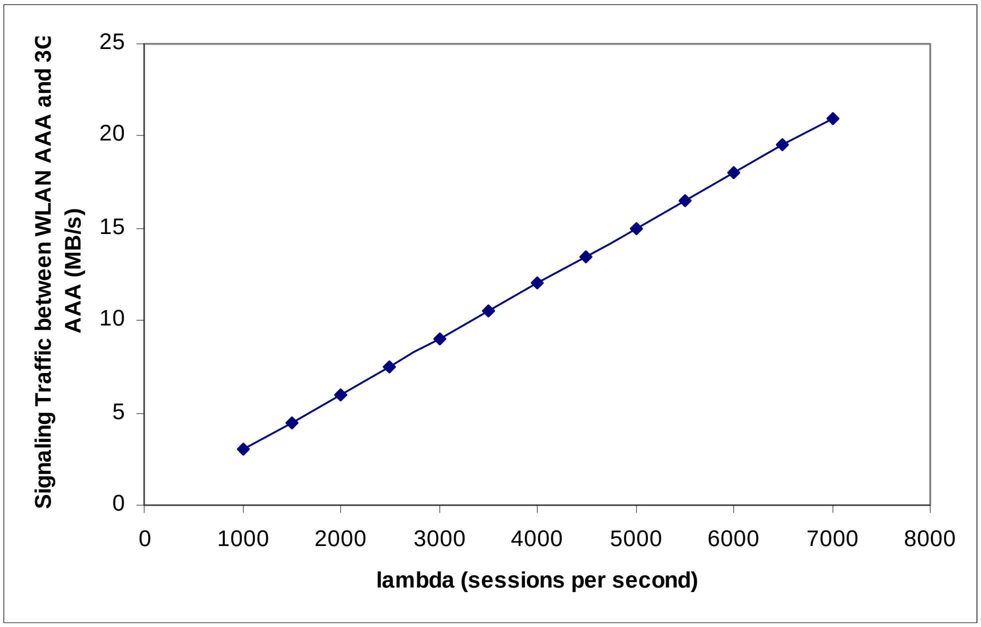


Fig. 57. Signaling Traffic vs Sessions per second

The graph in Fig. 58 depicts the relation between the data traffic generated as the number of users or the number of active session's increases.

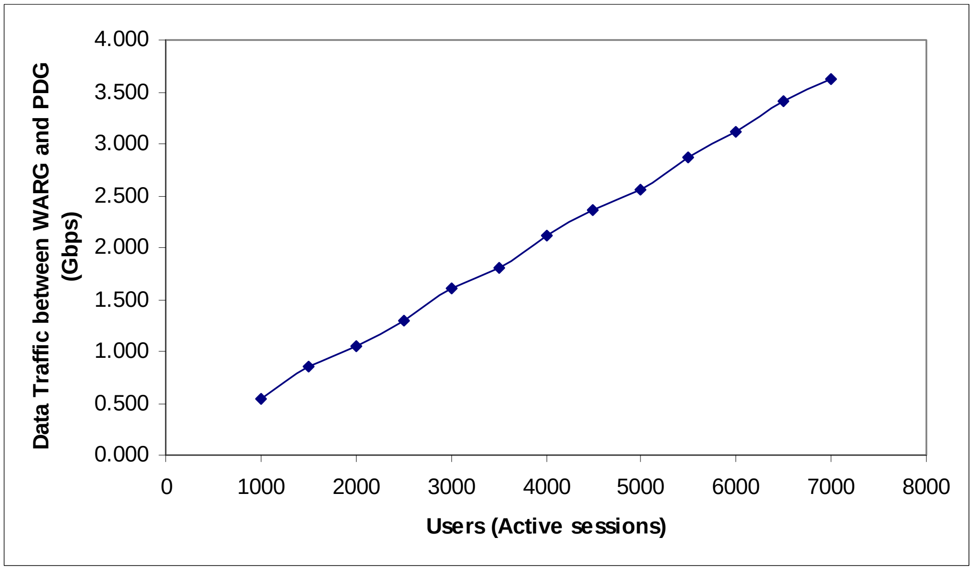


Fig. 58. Data Traffic generated vs Active sessions

5.4.2.1 Backbone Alternatives

There are two types of backbone connections: wired and wireless. In our design of the KFIA PWLAN we recommend a wired backbone for the following reasons:

- A wireless backbone would require the use of equipment such as wireless bridges or repeaters and this would increase the cost of the PWLAN deployment.
- A wired backbone would be ideal as the cabling required (telephone wires) is already present. Therefore, there is no installation cost added.
- Wireless backbone equipment suffers from overhead in the form of routing decisions. This overhead is not present in a wired backbone due to the presence of dedicated paths.

5.4.2.2 Channel Allocation

An important requirement in placement of AP's is channel allocation. The 802.11 b/g standard provides fourteen channels. The channel represents the center frequency that the AP uses (e.g., 2.412 GHz for channel 1 and 2.417 GHz for channel 2). There is only 5 MHz separation between the center frequencies. As a result, the signal overlaps with several adjacent channel frequencies. Out of the fourteen channels, only three are non-overlapping. Therefore, to minimize interference, APs that are within range of each other have to be set to non-overlapping channels.

However, due to restrictions on obtaining floor plans for the KFIA, the exact locations to place the AP's could not be determined. The table below summarizes the approximate equipment list that would be used in establishing a PWLAN at KFIA.

Equipment	**Basic Function**
Wireless Access Router Gateway (WARG)	Acts a Gateway to the Internet and Operator's network
Two 48-port switches	To connect the APs to the WARG
One WLAN AAA server	To authenticate and charge the PWLAN users
One DHCP server	To assign IPs to the PWLAN users
120 Access Points	To provide PWLAN coverage

Table 6. Equipment List and Functionality

6 Conclusion

In this report we have proposed a WLAN-cellular interworking framework for network providers, preceded by a detailed discussion of the background information related to developing interworking architectures, a comprehensive survey of various WLAN-cellular solutions and an all inclusive coverage of mobility management protocols.

There is a proliferation of research and development in the concept of interworking between heterogeneous wireless networks as future generation networks are envisioned to be an integration of such networks. These entirely packet-switched networks will provide interactive multimedia services (e.g. teleconferencing), wider bandwidths and global mobility at a low cost. Therefore subscribers will have access to high quality video and audio and will be able to main the same multimedia sessions with the same quality of service as they move across the different access networks.

7 References

1. "PWLAN for Service Providers solution overview" Cisco Systems Inc., 2004 http://www.cisco.com/warp/public/cc/so/neso/mbwlso/pwlan_br.pdf
2. "Cisco Mobile Exchange Solution Overview", Cisco Systems Inc., 2003 http://www.cisco.com/warp/public/cc/so/neso/mbwlso/ns278/cmxru_ov.pdf
3. "The Business Case for Cisco IP Solutions for Mobile Operators" 2005 http://www.cisco.com/application/pdf/en/us/guest/netsol/ns177/c654/cdccont_0900aecd80311888.pdf
4. Farrington S., "Head for the Hotspot – How Service Providers can deploy Profitable Wireless Networks", Packet Magazine, Cisco Systems, 4th Quarter 2004.
5. "Cisco ITP MAP Gateway for Public WLAN SIM Authentication and Authorization" 2003 http://www.cisco.com/warp/public/cc/pd/witc/itp/prodlit/mapga_wp.pdf
6. "Public Wireless LAN (PWLAN)" http://www.cisco.com/en/US/netsol/ns341/ns396/ns177/ns436/networking_solutions_solution_category.html
7. Olivier S., Poiraud P., "Public WLAN for Mobile Operators", Alcatel Communications Review 4th quarter 2003/ 1st quarter 2004.
8. Drevon N., Baudet S., Sigle R. and Vriendt J., "Interworking strategy and service continuity between GSM-UMTS-WLAN Radio Technologies" Alcatel Communications Review 4th quarter 2004
9. "Extending Coverage Areas for Wireless Internet Service Providers" http://www.nortel.com/solutions/wrlsmesh/collateral/nn110400-113004.pdf
10. "Wireless Mesh Network" http://www.nortel.com/solutions/wrlsmesh/collateral/nn106481-042805.pdf
11. "Nortel: Wireless Advanced Technologies" http://www.nortel.com/solutions/wireless/features/2005/wireless_advanced_technologies.html
12. "Wireless Routers: Wi-Fi Access via WLAN from Juniper Networks" http://www.juniper.net/solutions/mobile/pwlan.html

13. “Public Wireless LANs”
http://www.juniper.net/solutions/literature/solutionbriefs/351024.pdf
http://www.nortel.com/solutions/wrlsmesh/collateral/nn106481-042805.pdf

14. “SDX-300 Advanced Services Gateway”
http://www.juniper.net/solutions/literature/tech_note/552007.pdf

15. “Edge Router - Juniper Networks E-series”
http://www.juniper.net/products/eseries/

16. “Advanced IP Service Deployment System - Juniper Networks SDX-300”
http://www.juniper.net/products/sdx/

17. Rigney C., Willens S., Rubens A. and Simpson W.,"Remote Authentication Dial In User Service (RADIUS)", RFC 2865, June 2000
http://www.ietf.org/rfc/rfc2865.txt

18. “TACACS (and TACACS+): Terminal Access Controller Access Control System”. http://www.javvin.com/protocolTACACS.html

19. “DIAMETER”
http://en.wikipedia.org/wiki/DIAMETER

20. Calhoun P., Loughney J., Guttman E., Zorn G., and Arkko J., “ Diameter Base Protocol”, RFC 3588, September 2003
http://www.ietf.org/rfc/rfc3588.txt

21. “Introduction to Diameter: White Paper”
http://docs.hp.com/en/T1428-90011/T1428-90011.pdf

22. “Internet accounting - Diameter”
http://ing.ctit.utwente.nl/WU5/D5.1/Technology/diameter/

23. Walke B., “Mobile Radio Networks”, 2nd Edition, John Wiley, 2001.

24. Holma H. and Toskala A.,”WCDMA for UMTS”, 2nd edition John Wiley, 2002.

25. Walke B., Seidenberg P. and Althoff M.P.,”UMTS – The Fundamentals”, 1st edition, John Wiley 2003.

26. “Cisco Mobile Exchange (CMX) solution Guide”, Cisco Systems, pg2-29
http://www.cisco.com/application/pdf/en/us/guest/netsol/ns278/c2001/ccmigration_09186a008011b522.pdf

27. Liu C. and Zhou C, "HCRAS: A novel hybrid interworking architecture between WLAN and UMTS cellular networks", *Consumer Communications and Networking Conference, CCNC*, Second IEEE 2005, pp. 374-379.

28. Liu C. and Zhou C., "An Improved Interworking Architecture for UMTS-WLAN Tight Coupling", *Wireless Communication and Networking Conference*, IEEE 2005, vol.3, pp. 1690-1695.

29. Leu J. et al, "Practical Considerations on End-to-End Cellular/PWLAN Architecture in support of Bilateral Roaming", *Wireless Communication and Networking Conference*, IEEE 2005, vol. 3, pp. 1702-1707.

30. Salkintzis A., "WLAN/3G Interworking Architectures for Next Generation Hybrid Data Networks", *IEEE International Conference on Communications* 2004, vol. 7, pp. 3984-3988.

31. Q. Zhang et al., "Efficient Mobility Management for Vertical Handoff between WWAN and WLAN," *IEEE Communications Magazine*, vol. 41, no. 11, Nov. 2003, pp. 102 – 108

32. N. Sattari, P. Pangalos, and H. Aghvam, "Seamless Handover between WLAN and UMTS, " *Vehicular Technology Conference*, vol. 5, 17-19 May 2004, pp. 3035 - 3038

33. S. Tsao and C. Lin, "Design and Evaluation of UMTS-WLAN Interworking Strategies," *Vehicular Technology Conference*, vol. 2, 24-28 Sept. 2002, pp. 777 – 781

34. "Introduction to Wireless LANs" http://www.wlana.org/learn/intro.pdf

35. L. McKeag, "WLAN Roaming – the basics," *Techworld Online Magazine*, Mar. 2004 http://techworld.com/mobility/features/index.cfm?featureid=435&Page=1&pagePos=16

36. A. Salkintzis and N.Passas, "A New Approach for Fast Handovers in Mobile Multimedia Networks," *Vehicular Technology Conference*, vol. 5, 17-19 May 2004, pp. 2972 – 2976

37. A. Salkintzis, C. Fors, and R. Pazhyannur, "WLAN-GPRS Integration for Next-Generation Mobile Data Networks," *IEEE Wireless Communication Magazine*, vol. 9, no. 5, Oct. 2002, pp. 112-124

38. S. Das et al., “Intra-Domain Mobility Management Protocol (IDMP),” http://www1.cs.columbia.edu/~dutta/research/idmp.pdf

39. A. Misra et al., “IDMP-Based Fast Handoffs and Paging in IP-Based 4G Mobile Networks,” *IEEE Communications Magazine*, vol. 40, no. 3, Mar. 2002, pp. 138 – 145

40. R. Ramjee et al., “HAWAII: A Domain-based Approach for Supporting Mobility in Wide-area Wireless Networks,” *Seventh International Conference on Network Protocols*, 31 Oct.-3 Nov. 1999, pp. 283 – 292

41. A. Campbell et al., “Design, Implementation, and Evaluation of Cellular IP,” *IEEE Personal Communications*, vol. 7, no. 4, Aug. 2000, pp. 42 – 49

42. I. F. Akyildiz et al., “Mobility Management for Next Generation Wireless Systems,” Proc. IEEE, vol. 87, no. 8, Aug. 1999, pp. 1347–84.

43. I.F. Akyildiz et al., “Survey of Mobility Management in next-generation all-IP-based wireless systems”, *IEEE Wireless Communications*, August 2004

44. Perkins, C.E.: “Mobile IP”, IEEE Communications Magazine, 1997, vol. 35, issue 5, p.84-99.

45. Perkins, C.E.: “Mobile Networking through Mobile IP”, IEEE Internet Computing, 1998, vol. 2, issue 1, p. 58-69.

46. Perkins, C.: “IP Mobility Support”, RFC 2002, 1996, p. 1-79.

47. Nokia: “Introducing Mobile IPv6 in 2G and 3G mobile networks”, Nokia, 2001, p. 1-16.

48. Johnson, D.: “Mobility Support in IPv6”, RFC 3775, 2004, p. 1-165.

49. Riegel, M. and Tüxen, M.: “Mobile SCTP: Transport Layer Mobility Management for the Internet”, SoftCOM 2002, 2002.

50. Stewart, R. *et al.*: “Stream Control Transmission Protocol”, RFC 2960, 2000, p. 1-134.

51. Stewart, R. *et al.*: “Stream Control Transmission Protocol (SCTP) Dynamic Address Reconfiguration”, Internet draft, draft-ietf-tsvwg-addip-sctp-08.txt, 2003, p. 1-37.

52. Koh, S.J. *et al.*: “Mobile SCTP for Transport Layer Mobility”, Internet draft, draftsjkoh-sctp-mobility-03.txt, 2004, p. 1-14.

53. Wedlund, E. and Schulzrinne, H.: “Mobility Support using SIP”, ACM/IEEE, *International Conference on Wireless and Mobile Multimedia*, 1999, p. 1-7.

54. Schulzrinne, H. and Wedlund, E.: “Application–Layer Mobility Using SIP”, IEEE Service Portability and Virtual Customer Environments, 2001, p. 1-9.

55. Rosenberg, J. *et al.*: “SIP: Session Initiation Protocol”, RFC 3261, 2002, p. 1-269.

56. H. Haverinen, J. Salowey, “EAP-SIM Authentication”, draft-haverinen-pppext-eap-sim-16.txt, December 2004.

57. S. Glass, T. Hiller, S. Jacobs, and C. Perkins, “Mobile IP Authentication, Authorization, and Accounting Requirements,” Internet draft, 2000

58. http://www.the-saudi.net/kfia/kfia.htm#GENERAL%20PARAMETERSAndSITE%20DATA

59. 3GPP TS 23.234 v6.6.0, “3G System to WLAN Interworking; System Description (Release 6),” Sept. 2005.

60. W. Wu, N. Banerjee, K. Basu and S. K. Das, “SIP-based Vertical Handoff between WWANs and WLANs”, *IEEE Wireless Communications*, June 2005.

61. Li Ma, FeiYu and V.C.M. Leung, “A new method to support UMTS/WLAN Vertical Handover using SCTP”, *IEEE Wireless Communications*, August 2004.

62. A. Salkintzis, “Interworking Techniques and Architectures for WLAN/3G Integration towards 4G Mobile Data Networks,” *IEEE Wireless Communication Magazine*, June 2004, pp. 50-61

63. http://www.javvin.com

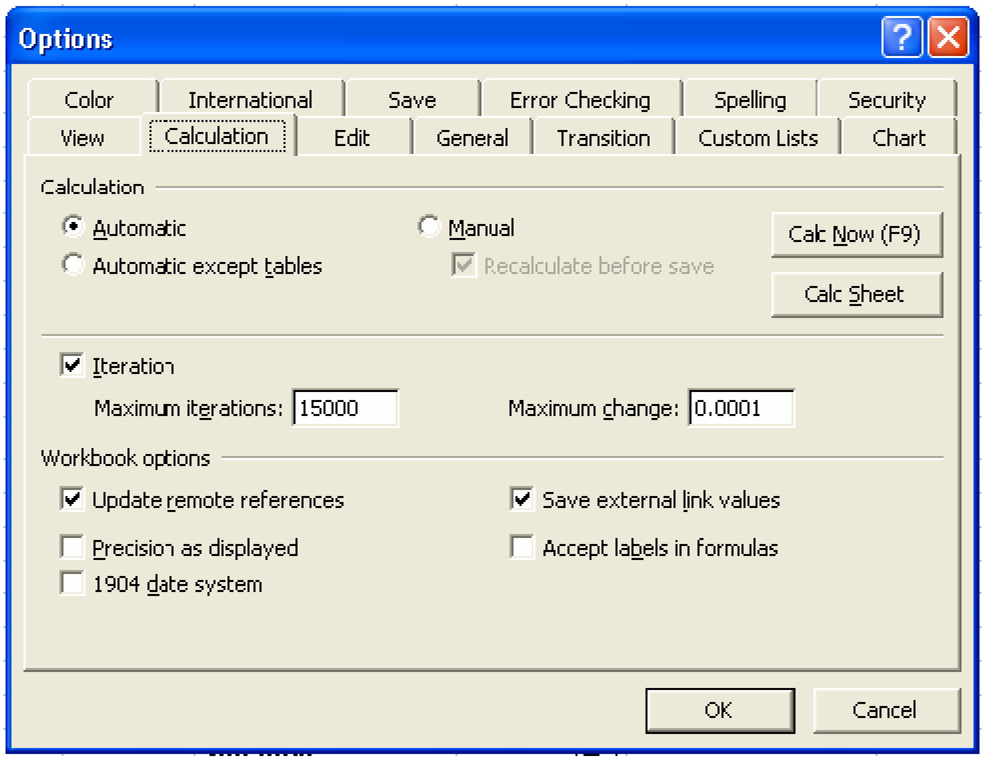
64. G. Bianchi, “IEEE 802.11—Saturation Throughput Analysis,” *IEEE Communications Letters*, vol. 2, no. 12, Dec. 1998

8 Appendix

8.1 Access Point Estimator (APE)

After opening the APE.xls file, make sure that you have done the following changes

1. Go to ‘Tools’ and select ‘Options’
2. Choose the ‘Calculation’ tab
3. Set ‘Maximum iterations’ to 15000 and ‘Maximum change’ to 0.0001
4. Check the ‘Iteration’ checkbox



Calculation

● Automatic ○ Manual
○ Automatic except tables ☑ Recalculate before save

☑ Iteration

Maximum iterations: 15000 Maximum change: 0.0001

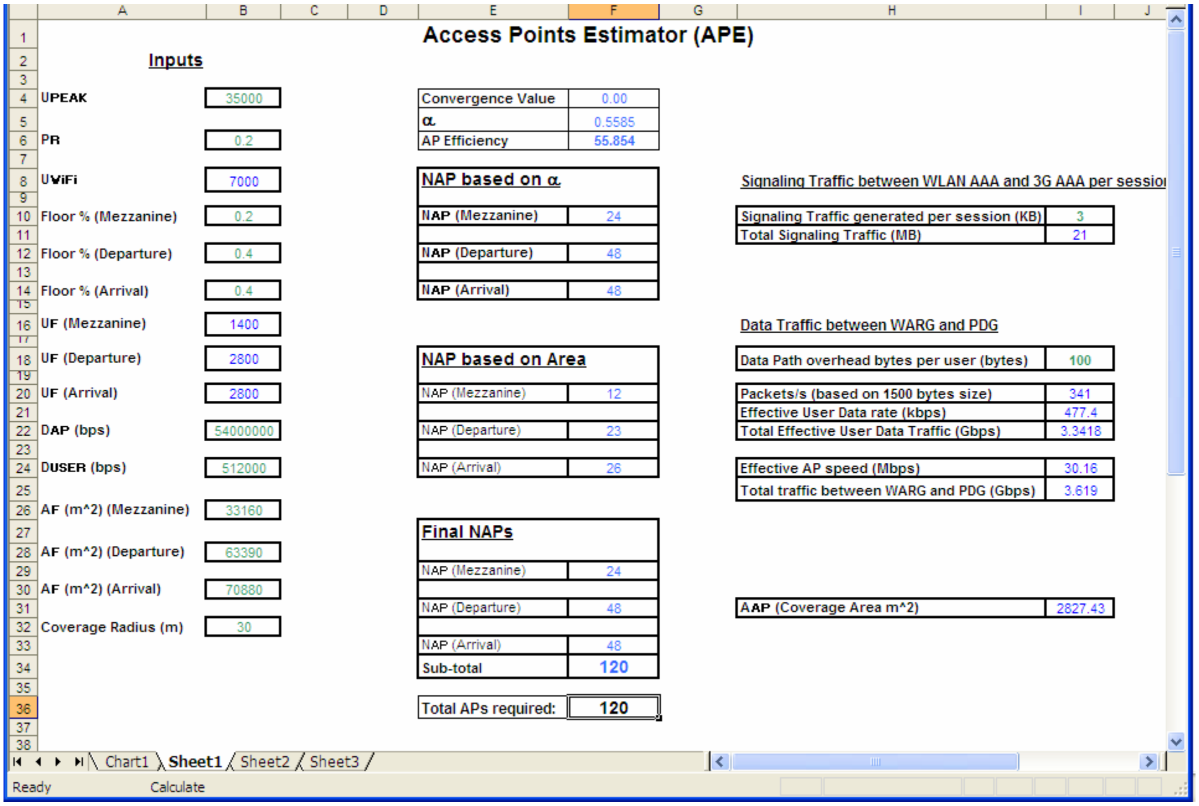
Workbook options

☑ Update remote references ☑ Save external link values
□ Precision as displayed □ Accept labels in formulas
□ 1904 date system

Fig. 59. Changes in the Options Pane

Now in the ‘Inputs’ column, input the values as desired and the estimator will show how many APs are required and at what efficiency they will function. A snapshot of APE is given below. The various inputs that have been used are discussed in detail in Table 5 of the report. A discussion of the calculations and analysis used in this estimator is also given in section 5.4.

Important Note: The user can only make changes where the text is green i.e. the inputs. The rest of the values (in blue) are calculated by Excel and hence cannot be changed by the user. The final value i.e. the total number of APs needed is calculated and displayed in black at the bottom. Also note that the convergence value should be as close to zero as possible. So if for the given inputs the value is not zero, press F9 to re-calculate until it approaches zero.



Access Points Estimator (APE)

Inputs

UPEAK	35000
PR	0.2
UWiFi	7000
Floor % (Mezzanine)	0.2
Floor % (Departure)	0.4
Floor % (Arrival)	0.4
UF (Mezzanine)	1400
UF (Departure)	2800
UF (Arrival)	2800
DAP (bps)	54000000
DUSER (bps)	512000
AF (m^2) (Mezzanine)	33160
AF (m^2) (Departure)	63390
AF (m^2) (Arrival)	70880
Coverage Radius (m)	30

Convergence Value	0.00
α	0.5585
AP Efficiency	55.854

NAP based on α	
NAP (Mezzanine)	24
NAP (Departure)	48
NAP (Arrival)	48

NAP based on Area	
NAP (Mezzanine)	12
NAP (Departure)	23
NAP (Arrival)	26

Final NAPs	
NAP (Mezzanine)	24
NAP (Departure)	48
NAP (Arrival)	48
Sub-total	120

Total APs required:	120

Signaling Traffic between WLAN AAA and 3G AAA per sessio

Signaling Traffic generated per session (KB)	3
Total Signaling Traffic (MB)	21

Data Traffic between WARG and PDG

Data Path overhead bytes per user (bytes)	100
Packets/s (based on 1500 bytes size)	341
Effective User Data rate (kbps)	477.4
Total Effective User Data Traffic (Gbps)	3.3418
Effective AP speed (Mbps)	30.16
Total traffic between WARG and PDG (Gbps)	3.619

AAP (Coverage Area m^2)	2827.43

Fig. 60. Access Point Estimator in action

8.2 Publications arising from this work

1. Ashraf Mahmoud, Marwan Abu-Amara, Tarek Sheltami, Ejaz Rahman and Junaid Jaffar, “WLAN Integration For Future Generation Mobile Network Operators – A Case Study”, 18th Saudi National Computer Conference, March 2006. *In progress*