WiMAX - A Study of Mobility and a MAC-layer Implementation in GloMoSim

Michael Carlberg Lax and Annelie Dammander

April 6, 2006
Master's Thesis in Computing Science, 2*20 credits

Supervisor at CS-UmU: Thomas Nilsson
nilsson@cs.umu.se
Supervisor at TietoEnator: Torbjörn Lundmark
torbjorn.lundmark@tietoenator.com
Supervisor at LTU: Christer Åhlund
christer.ahlund@ltu.se
Examiner: Per Lindström
perl@cs.umu.se

Umeå University
Department of Computing Science
SE-901 87 UMEÅ
SWEDEN

Abstract

Future generation networks will be characterized by variable and high data rates, quality of services, seamless mobility both within a network and between networks of different technologies and service providers. An important aspect of components in a global next generation network is standardization to allow vendor independence and interoperability. A technology developed to fulfill these characteristics, standardized by IEEE, is 802.16, also referred to as WiMAX.

This architecture aims to apply high data rates, quality of services, long range and low deployment costs to a wireless access technology on a metropolitan scale.

The technology and architecture of WiMAX is the focus of the thesis paper, and more specifically its mobility capabilities. The thesis investigates the handover and internetworking capabilities of WiMAX and then implements selected MAC-layer functionality in the GloMoSim network simulator. Through simulation attempts are made to identify MAC-parameters affecting performance during handovers.

Results indicate that the study needs to be extended to cover upper layer protocols and procedures. At the MAC-layer the most deciding factors are predicted to be procedures executed in preparation for handover, rather than the specific handover process.

Preface

This thesis marks the end of our MSc education, a journey we will never forget. The thesis work, from draft to completion, has been demanding, but at the same time fun and experiencing. We have learned both through mistakes and success, with the former being as important as the later.

Many thanks to our supervisors, Christer, Torbjörn and Thomas for guiding us. Salute to everyone supplying the gossip at the coffee breaks. Thanks to Martin and Mattias for making us feel at home and thanks to TietoEnator for letting us test our wings.

Thanks to Nellie for enduring those days home alone. Voff.

Contents

1 Introduction **1**

1.1 Thesis Outline 1

2 Problem Description **3**

2.1 Problem Statement 4

2.2 Purpose 4

2.3 Scope 5

2.4 Methods 5

3 Theory **7**

3.1 Fixed Wireless Access 7

3.2 Network Sizes 10

3.3 Handover 11

3.4 Network Composition 13

4 WiMAX **15**

4.1 802.16 15

4.2 802.16e 17

4.3 Network Architecture 17

4.4 Deployment 19

4.5 WiMAX Forum 21

5 Homogeneous Mobility **23**

5.1 Background 23

5.2 802.16e handover 24

5.3 WiMAX Handover 26

5.4 Related Work 30

5.5 Conclusions 34

6 Heterogeneous Mobility **35**

6.1 Background 35

6.2 UMTS 36

6.3 Wi-Fi 41
6.4 Related Work 44
6.5 Conclusions 46

7 Simulation Study 49
7.1 Environment 49
7.2 Scenarios 51

8 Conclusion 57
8.1 GloMoSim 58
8.2 Implementation 58
8.3 Simulation Results 59
8.4 Related Work 60

9 Abbreviations and Acronyms 63

References 67

A Sequence Charts 73
A.1 Network Entry 73
A.2 Hard Handover 74
A.3 Fast Base Station Switching 74

B 802.16 MAC-summary 75
B.1 Convergence Sublayer 75
B.2 Medium Access Control 77

List of Figures

3.1 Point to Point FWA . . . 8
3.2 Point to Multipoint FWA . . . 9
3.3 Mesh configured FWA . . . 9
3.4 Overview of the different network sizes . . . 10
3.5 Vertical and horizontal handovers . . . 12

4.1 Layering of IEEE Std. 802.16 . . . 16
4.2 The WiMAX Network Reference Model . . . 18
4.3 Available frequencies, world map [18] . . . 20
4.4 Layer Harmonization and WiMAX's role. . . . 21

5.1 ASN Decomposition . . . 26
5.2 Handoff function network transaction . . . 28
5.3 Intra ASN Handover . . . 29
5.4 Inter ASN Handover . . . 30
5.5 Wireless Access Network Model[7, p. 4] . . . 32

6.1 UMTS Reference Architecture Main Components . . . 37
6.2 UMTS And WiMAX Architecture Components . . . 38
6.3 WiMAX/UMTS Loose Coupling . . . 40
6.4 A Common WLAN Configuration . . . 42
6.5 IEEE Std. 802.21 Layering . . . 44

7.1 Network Topology . . . 50
7.2 Measured Network Entry Times . . . 53
7.3 Measured Handover Times (Varying Frame Rate) . . . 54
7.4 Measured Handover Times (Varying UCD) . . . 55
7.5 Handover: Throughput of BS1 and BS2 . . . 56
7.6 FBSS: Throughput of BS1 and BS2 . . . 56

8.1 Heterogeneous Technologies Enveloping Each Other . . . 57
8.2 RPS Outdoor Coverage Distribution[14, p. 179] . . . 61

A.1 Initial Network Entry Message Exchange 73
A.2 Hard Handover Message Exchange 74
A.3 Fase Base Station Switching Message Exchange 74

B.1 Layering of IEEE Std. 802.16 75
B.2 Classification and CID mapping[2, p. 22] 76
B.3 MAC SDU Format[2, p. 21] 77
B.4 TDD Frame Structure[2, p. 153] 78
B.5 OFDM Frame Structure[2, p. 450] 79
B.6 DL-MAP references to downlink bursts of varying types[2, p. 310] 80
B.7 MAC PDU Format[2, p. 35] 81
B.8 Sample UL-MAP with multicast and broadcast IE[2, p. 145] 85
B.9 Handover and initial network entry[3, p. 176] 86
B.10 MS in sleep mode with two Power Saving Classes[3, p. 166] 87

List of Tables

4.1 The 802.16 Alphabet Soup 17
4.2 Estimated timeline for WiMAX 20

5.1 BS and ASN GW roles 27
5.2 Re-anchor decisions and outcome 29

6.1 UMTS Data rates 36
6.2 UMTS QoS Classes 36
6.3 UMTS/WiMAX QoS Mappings 38
6.4 802.11 Flavors 41

7.1 GloMoSim Protocol Layers 49
7.2 Selected GloMoSim Layers 51
7.3 Tested Frame Rates 52

8.1 References to the Problem Statement 58

Chapter 1

Introduction

With the rollout of the third generation cellular networks, 3G, the aim is already set towards the next generation.

Future generation networks will be characterized by variable and high data rates, quality of services and seamless mobility both within a network and between networks of different technologies and service providers.

An important aspect of components in a global next generation network is standardization to allow vendor independence and interoperability.

A technology developed to fulfil these characteristics, standardized by IEEE[1][28], is 802.16[2]. Even though this IEEE standard is limited to the air interface, it is commonly referred to as Worldwide Interoperability for Microwave Access (WiMAX)[19]. WiMAX is the end-to-end network architecture built on the IEEE Std. 802.16 and amendments such as the IEEE Std. 802.16e[3].

This architecture aims to apply high data rates, Quality of Service (QoS), range and low deployment costs to a wireless access technology on a metropolitan scale.

The network architecture is developed by the WiMAX Forum, an interest group backed by technology companies, such as Intel, Fujitsu, Samsung, AT&T and Alvarion. See WiMAX Forum[21] for a complete member roster.

The technology and architecture of WiMAX is the focus of this thesis, and more specifically its mobility capabilities. These capabilities are evaluated in comparison with other technologies and are concluded with a practical study of a Medium Access Control (MAC) layer implementation.

The thesis work was made possible through cooperation between TietoEnator, Luleå University and Umeå University. The work took place at TietoEnator, Ursviken, Sweden.

1.1 Thesis Outline

This thesis is divided into the following areas. Chapter 2 presents the purpose for this thesis and the methods used. It explains in more detail why this is an area worthy of interest and to what extent the work has been limited.

Chapter 3 will introduce the theoretical concepts of handovers, homogeneous and heterogeneous networks. It also contains terminology and scope of the research area

[1] Institute of Electrical and Electronics Engineers

such as network sizes, e.g. Metropolitan Area Network (MAN), Local Area Network (LAN) etc, and the network family of Broadband Fixed Wireless Access (BWFA).

In Chapter 4 the IEEE Std. 802.16 and WiMAX Forum's network architecture is introduced in more detail, together with features and market deployment scenarios.

Chapter 5 marks the start of the in-depth study of WiMAX mobility and will specifically look at mobility within the same network technology. This is done in comparison with different versions of the IEEE Std. 802.16 and other standards such as IEEE Std. 802.20.

The theoretical study continues in chapter 6 with a look at the capabilities of heterogeneous mobility between a WiMAX network and a Universal Mobile Telecommunications (UMTS)[61] network. Heterogeneous mobility is also invesitaged between a WiMAX network and a Wireless Fidelity (Wi-Fi)[8] network.

After discussing mobility in theory, chapter 7 covers the simulation environment, setup of scenarios for simulation of the 802.16 MAC-layer. It also covers the results of the simulations.

Conclusions and discussions raised during the work on this paper together with suggestions for future work is presented in chapter 8. That chapter also highlights related work on implementations of the 802.16 MAC-layer and testing of WiMAX equipment.

Appendix B on page 75 contains a short summary of 802.16, the WiMAX air interface.

Chapter 2

Problem Description

There are many new wireless access technologies and standards under development today, looking to extend mobility, data rates and user services and thus filling a position in the future generation networks. The building blocks of these networks are however uncertain due to this technology diversity.

Mobile communications, originating from fixed circuit switched voice traffic in public phone networks, has evolved to support stringent QoS and packet based traffic together with global mobility. The telecom industry expects data rates to reach 100 Mb/s with the new Super 3G[32] networks as a part of 3GPP's[1] Long Term Evolution[12] of 3G networks. Before that Turbo 3G will enable data rates of 14 Mb/s as an extension to deployed 3G networks with High Speed Downlink Packet Access, HSDPA[15].

The Internet Protocol (IP) technology, with front figures such as IEEE and IETF[2][29], has gone from indoor, fixed high speed networks with 802.11, to outdoor wireless access with support for mobility with standards such as 802.16 for broadband wireless access, 802.21 for mobility between networks and 802.20, the all-mobility standard.

As both of these paradigms of mobile and computer communications strive to meet the market requirements of both high speed and quality of service, they seemingly also move to meet each others characteristics.

WiMAX, based on IEEE Std. 802.16 attempts to bring quality of services, high data rates and coverage to wireless computer networks and to work as a "last mile" solution for end user access. Although not designed for mobility from the start, the standard could lend itself for use in mobility scenarios and for this purpose the IEEE Std. 802.16e has been developed with requirements to support vehicular mobility and seamless handover while maintaining differentiated QoS.

This technology shows promise in filling the gap between 3G and Wireless LAN (WLAN). A gap where the superior QoS of cellular networks are combined with the flexibility and scalability of IP technology.

As WiMAX is currently under development with parts of the 802.16 standard being complete, this is an interesting case for an observational and comparative study. Will WiMAX meet the goals of its creators and the market, and what issues can be expected to arise during development and deployment?

This paper serves as an orientation on the subject of mobility in WiMAX and the IEEE Std. 802.16. This orientation is done in comparison with other technologies as a

[1] 3rd Generation Partnership Project

[2] Internet Engineering Task Force

frame of reference.

2.1 Problem Statement

Different degrees of mobility for communication devices is a capability becoming increasingly desired by end users together with emerging services for mobile devices such as streaming audio/video through packet data. The methods for supporting various degrees of device mobility, e.g portability, roaming, full mobility, often varies between technologies and defines the mobility characteristics of each.

What are the mobility capabilities of WiMAX and what are the issues when this mobility is set in motion within a WiMAX network and between other types of technologies?

How does WiMAX suggest intra-network mobility and can it be compared to similar contemporary technologies?

These questions hope to highlight the tools available for devices moving both within WiMAX networks or crossing borders between WiMAX and other technologies.

As WiMAX is a fairly new and hyped context in the debate on wireless networks and broadband an interesting study is also the WiMAX MAC-layer, IEEE Std. 802.16.

What tools would be suitable for building and simulating a complete WiMAX network? GloMoSim[40] is a library for parallel simulation of wireless networks open for own implementations and familiar to the authors through previous experience. Is GloMoSim a suitable platform for a WiMAX simulator?

To summarize, this thesis will attempt to answer the following questions:

- What are the mobility capabilities of the WiMAX network architecture with regards to both heterogeneous and homogeneous networks?
- How efficient is this mobility support with regards to handover delays and overhead?
- Is GloMoSim suitable for developing a WiMAX simulator environment?

2.2 Purpose

As the standard is new and still under development at the writing of this thesis, there is a need to gather and filter information on this technology and to present it in a collected fashion.

This thesis contributes with a theoretical and practical part.

The theoretical part presents the following.

- Theoretical background of the subject of WiMAX
- Study on heterogeneous and homogeneous mobility
- Summary of the 802.16 MAC-layer functionality and its key components.

The practical study presents a suggested framework for implementing the 802.16 MAC-layer and the WiMAX architecture in a simulator environment.

2.3 Scope

This thesis has limited the scope to mobility and will not focus on the fixed wireless aspects of WiMAX. Since the WiMAX network architecture and parts of the 802.16 standard is still undergoing changes, this paper will not be able to in-depth study certain aspects of WiMAX within the mobility scope, but should rather open up to discussion where applicable.

The thesis reflects the status of WiMAX and its underlying standards during the period of which this study is being performed, and the end results can come to differ with later revisions of the WiMAX architecture.

The implementation is limited to scenarios developed during the analysis and design of the framework. It primarily focuses on movement of subscribers in a WiMAX network, and leaves other fundamentals like scheduling and QoS for future work.

2.4 Methods

A study of literature will precede the theoretical summary on mobility in chapter 5 and 6. Since there is little practical experience with WiMAX, the literature study will through related research on inter-networking, mobility and handovers build a base for evaluating WiMAX mobility capabilities. Other technologies, such as UMTS, Wi-Fi and IEEE 802.20 will form a frame of reference for this comparison.

In order to evaluate the suitability of GloMoSim as a platform for a WiMAX simulator environment, selected aspects of the WiMAX MAC-layer model will be implemented.

The practical simulation of the WiMAX MAC-layer will be based on discussions with experienced people from the industry and market driven issues. These discussions govern within what areas of interest a simulation will performed and how extensive the implementation will be. The implementation will be driven by scenarios defined during this process.

Chapter 3

Theory

The theory in this chapter defines the scope of the paper. It will highlight areas of research that serves as the base for the literature study, and thus also limits the conclusions and discussions introduced ahead.

The concepts covered here will help the reader place WiMAX itself in a broader perspective and also explain terminology used throughout the paper.

3.1 Fixed Wireless Access

Fixed Wireless Access (FWA) refers to a wireless infrastructure replacing regular cable. It is also known as Radio Fixed Wireless or Wireless Local Loop, and has the main purpose of replacing the "last mile" connection between a user and backbone, thus serving as a wireless backhaul[37].

In contrast to many other wireless communications networks, the sender and receiver in a FWA system are both stationary, rather than mobile, hence the name Fixed. This mitigates many of the challenges found in mobile wireless networks, such a highly dynamic user topology and bandwidth limitations. The FWA network does not need to adapt to fast changes brought on by mobility, but can still be very flexible in comparison to regular cable connections.

Deploying a FWA network can give new users access immediately without having to dig cable, and requires only hardware installed at the user end if the system is already deployed. This enables very fast rollout of FWA networks compared to regular cable access networks where digging requires resource to be invested in something other than the access technology itself. The biggest competitor to FWA systems is the Digital Subscriber Line (DSL)[11] technology as this already has a widespread infrastructure through the Public Switched Telephone Network (PSTN)[66].

There are three different types of FWA topologies.

– Point to Point (PTP) systems has two base stations with the communication statically configured for a link between these two base stations only. This dedicated link is characterized by higher bandwidth compared to PTP systems and can utilize directed antennas. Figure 3.1 on the next page shows a typical PTP deployment of transmitters[37, p. 1]. These types of systems are often used as backhaul[1].

[1] Transporting data to the network backbone

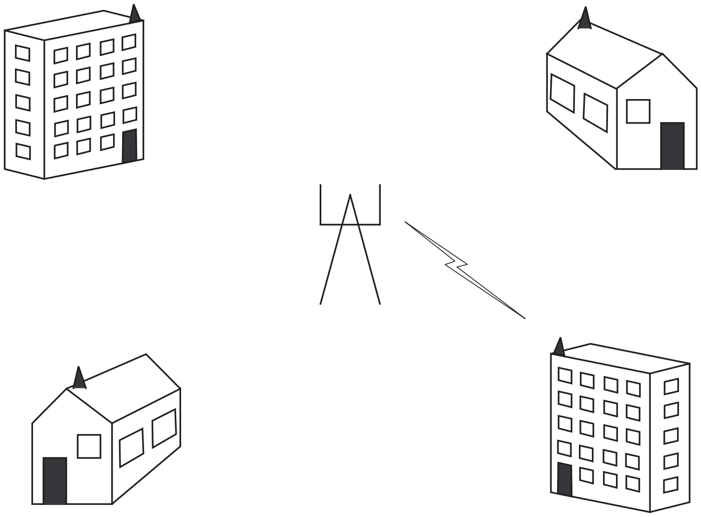


Figure 3.1: Point to Point FWA

- Point to Multipoint (PMP) is an access network with one more powerful base station and many smaller subscriber stations. Users can get immediate access to the network after installing only user equipment. Here the subscriber station can deploy directed antennas towards the base station, where the base station has omni-directional or a cluster of directed antennas. See figure 3.2 on the facing page for a general PMP configuration, typically serving as broadband replacement[37, p. 1].
- Multipoint to Multipoint also called mesh network, is where there is no centralized base station[37, p. 2]. As more user transmitters join the mesh, the covered area is increased. This architecture adds complexity in routing, QoS and node discovery amongst other things.

FWA systems are also characterized by Line of Sight (LOS) or Non Line of Sight (NLOS) transmission capabilities. LOS systems usually operate in frequencies above 20 GHz where the particle characteristics of radio waves makes transmissions more susceptible to reflection and blocking.

As all wireless communication systems, FWA has to consider the environmental effects on the radio waves. Since FWA often operate outdoors over MAN or smaller areas the signal can suffer from interference, fading, multi-path propagation, radio noise, path loss and many other factors. There are numerous measures that can help deal with these issues.

Modulation techniques[59, p. 46] can be used to adapt the data rates according to needs in QoS while mitigating signal impairments. Higher modulations such as 64QAM[2], coding 6 bits in every symbol, are used for high bit rates, but since there are more bits in every symbol it is also more sensitive to interference. This can be dealt with by switching to a lower modulation scheme like 16QAM.

[2] Quadrature Amplitude Modulation

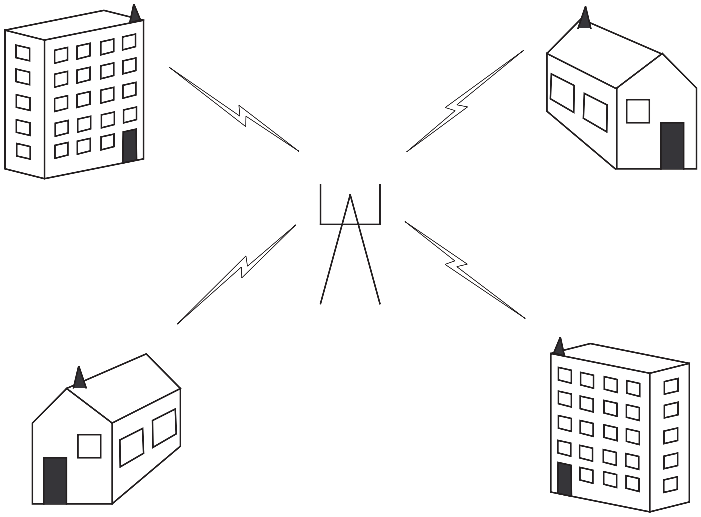


Figure 3.2: Point to Multipoint FWA

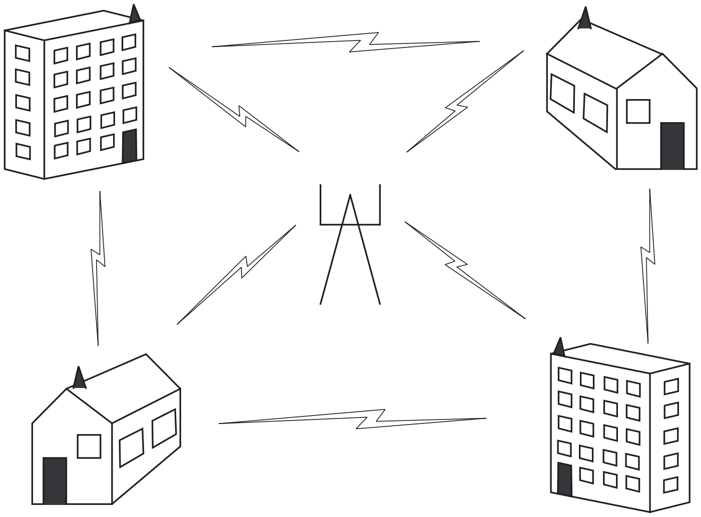


Figure 3.3: Mesh configured FWA

To increase throughput and QoS in bad links, Forward Error Correction (FEC) can be applied. This adds redundant information in the transmitted data that enables corrupted data to be recovered with only partially correct packets. This redundancy has a negative effect on link throughput with little interference, but can raise through put on a bad radio link.

Interference can also be handled with frequency allocation techniques, with channels being separated by frequencies as in Frequency Division Duplex (FDD), or by time as in Time Division Duplex (TDD).

The versatility and development of BFWA systems, makes them suitable for consideration in future networks. The wireless aspect of BFWA, originally intended for fixed end points[2], can be adapted to support mobility in different degrees, such as nomadic mobility and roaming[3].

3.2 Network Sizes

As means to consider the scale of management and deployment of communication networks, this section briefly recapitulates different network sizes relevant to the thesis.

Metropolitan Area Networks (MAN) are data networks built from high speed infrastructure to service a whole town or city, usually with fiber cable, bridges and routers serving as a backbone for smaller LANs.

Local Area Networks (LAN) are smaller networks built to cover a building or group of buildings. A common LAN technology is Ethernet or IEEE Std. 802.3[6], utilizing an Internet Service Provider (ISP) for internet access.

As a complement to LAN, Wireless Local Area Networks (WLAN) has become a common sight. This type of networks is of the same scale as a LAN, but uses wireless radio transmissions instead of cable media, enabling some mobility and deployment without cable installation, and thus suffering from the impairments of radio transmissions. The IEEE Std. 802.11[1] is often associated with WLAN and is sometimes referred to as Wi-Fi, a radio technology based on different parts of 802.11. This technology is promoted and certified by the Wi-Fi Alliance[8].

Figure 3.4 shows the different networks in relation to coverage.

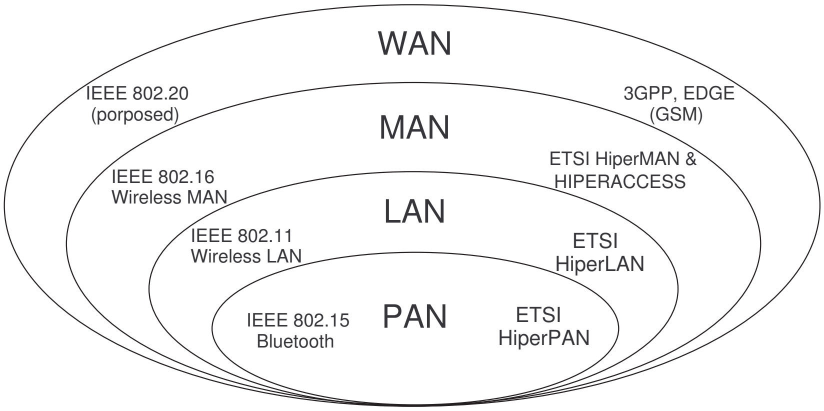


Figure 3.4: Overview of the different network sizes

3.3 Handover

The handover is an important process in mobile systems and it is defined by the migration of a Mobile Station (MS) between air-interfaces belonging to different Base Stations (BS). The reason for such a change could be that a cell is overloaded or that the MS gets out of the BS transmission range [65]. The BS associated with the mobile station before the handover is often called the serving BS while the new BS is referred to as the target BS as can be seen in [3, p. 5].

A handover can be divided into two parts, the pre-registration phase and the actual handover[7, p. 2]. The pre-registration phase includes messages such as a handover request and a list of possible target BSs. During this phase the MS can also measure the signal strength from adjacent BSs to help in the decision about which BS to use as target BS. When the actual handover takes place the mobile station will close the connection to the serving BS and open a new to the target BS.

There are two different kinds of handovers and they differ in the way they handle the connections to serving and target BS in the actual handover phase. The two types are hard and soft handover.

3.3.1 Hard Handover

In a hard handover the mobile station will interrupt the communication with the serving BS and make a transition to the target BS. The MS have to register with the target BS and can then begin to listen and send to the new BS instead. This kind of handover is often referred to as a break-before-make[74, p. 43] handover since the communication with the serving BS is broken before the actual handover. The hard handover is used in cases where the soft handover is not suitable or can not be employed, e.g. in handovers between different networks.

Hard handovers are the most common way to perform a handover and it is used in systems like UMTS[59, p. 154] and WiMAX [19, p. 174]. A common scenario when a hard handover is used could be a crowded city area covered by several BSs. If a BS observes that its cell is overloaded it can decide to transfer one of its mobile nodes to a BS working on another frequency. A handover request is sent to the mobile node which starts performing the handover. In this case the handover is of the hard type since the mobile station stops talking to the serving BS, changes the frequency and then connects to the target BS.

3.3.2 Soft Handover

During a soft handover the MS will keep contact with both base stations through out the second phase of the handover, often called make-before-break[74, p. 43]. The mobile station will in this way always be connected during the handover[73]. This makes the handover more robust to fast fading, shading and multi path propagation since the MS will receive several signals at once [59, p. 154]. There are how ever some drawbacks with the use of two connections, the overhead in the system will increase and the MS will use more network resources since there will be e.g. several channels reserved.

The soft handover is more complex than the hard but is often preferred since it is more stabile and gives better performance and a smoother transition. The soft handover will make the handover less visible for higher layer protocols such as Transmission Control Protocol (TCP). This kind of handover is common in UMTS[73] systems and will also be applied in 802.16e[3, p. 186].

When a MS is leaving the cell of the serving BS it can make a request for a handover to find a better BS. If soft handover is being used the mobile station will commence the registration phase with the target BS before aborting the connection to the serving BS. During the handover the MS will receive signals from both BSs. When the set up of the new connection is completed the transmissions from/to the serving BS will end.

3.3.3 Architectural Handover

When looking on a handover from an architectural point of view there are two different types, vertical and horizontal. The horizontal handover is a handover between base stations belonging to the same type of network technology while the vertical handover is made between base stations attached to different network technologies[59, p. 449]. See figure 3.5 for an illustration.

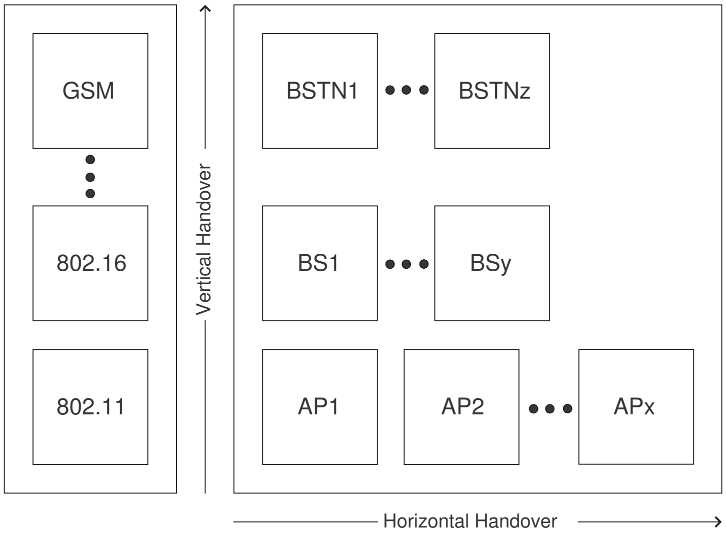


Figure 3.5: Vertical and horizontal handovers

Both the horizontal and vertical handover can be of the types hard or soft. Since the soft handover is more complex and demand more collaboration between components in the network it is more common in horizontal handovers than vertical ones.

Research in the area of vertical handover is closely related to reaserch in heterogeneous networks, as this type of handover is required to cross between different types of networks. Previous work in the area[74] has presented a model containing the technical aspects of vertical mobility in heterogeneous networks. It can be divided into three fundamental parts.

- Resource Management covers the aspect of resources in form of end-to-end QoS and network capacity etc.

- Mobility engineering models how the networks are integrated and what protocols are shared between them.

– Service Management deals with location management, user positioning and Authentication, Authorization and Accounting (AAA).

3.4 Network Composition

Today a lot of new wireless technologies are developed and they all have their special area of use. Some systems offer long range and low bandwidth radio transmissions while others offer a short transmission range and high bandwidth. Contrasting systems allows for wired high/low data transfers in the means of fiber optic cables or copper wire. When all these technologies cooperate to provide services they form a heterogeneous network. The opposite of a heterogeneous network is a homogeneous network. This section will explain the concepts of homogeneous and heterogeneous networks in a greater detail.

3.4.1 Homogeneous

In a homogenous network the technology used for the different components are all the same [51]. An example could be a mobile phone system based on the Wideband Code Division Multiple Access (WCDMA) technique. If the network only consists of units using this technique it is a homogeneous network. The network can in this case not be connected to a network that uses any other technique than the WCDMA.

Homogeneous networks often have a simpler structure and are easier to maintain than heterogeneous networks. A drawback that comes from this simplicity is that the network is forced to use only this technology. The users of the WCDMA phone system can only call other members of the network meaning that they can not call e.g. users of the public switched telephone network since the two different networks are not connected. In reality these two networks are connected to each other and form a heterogeneous network.

3.4.2 Heterogeneous

Homogenous networks of different types connected to each other form a heterogeneous network [74, p. 33]. There are a lot of examples of heterogeneous networks in today's society. Consider a big company situated at several floors in a building. The company network is connected to Internet with fiber optic cables provided by a local ISP. To connect offices on different floors to each other the company use twisted pair cables and the technology Ethernet. Internally within each office they use 802.11 to let the personnel get wireless access to the company network and internet. Some of the employees use Bluetooth to synchronize their PDA contacts and calendars with the computer. All these different technologies can be regarded as sub networks to the company network and together they build up a heterogeneous network.

In a heterogeneous network the architecture is more complex since sometimes completely different technologies have to work closely together. Problems that arise are both related to hardware and software. When connecting two different technologies together some of the components in the network will have to support both technologies, in hardware as well as software.

Different network types often support various types of QoS and this can lead to difficulties when connecting them to each other [74, p. 51]. If a network is supplying real time services to the customers and relay data through a network not supporting real time applications it can no longer guarantee its customers real time data to be delivered in time, compromising end-to-end QoS.

Even tough this kind of network introduces a lot of challenges it also gives more flexibility. Users of the network are not bound to a single technology and the possible applications and services that can be used increases.

Chapter 4

WiMAX

The technology commonly named as WiMAX is a BFWA system with the goal of delivering ”last mile” fixed, nomadic, portable and mobile wireless connections on a metropolitan scale.

The Physical (PHY) and MAC-layers has been developed by the *IEEE 802.16 Working Group for Broadband Wireless Access* to enable standardized development and deployment of broadband wireless networks in metropolitan areas.

This standard has been evolved beyond the PHY and MAC-layer to encompass an end-to-end network architecture, created by the WiMAX Forum, an organization promoting global interoperability and use of wireless broadband[1].

This chapter introduce the technology of WiMAX and its network architecture for comparison in the study of mobility in the chapters on homogeneous and heterogeneous mobility.

4.1 802.16

Forming the base of the WiMAX technology is the IEEE Std. 802.16. Though originally intended as a BFWA system serving as backhaul in a PMP network architecture, it is undergoing revisions for the standard to support different forms of mobility.

The current active standard 802.16-2004[2], for fixed installations, promote bandwidths of 70 Mbps or 2-10 Mbps/user covering up to 10 Km^2[10][2].

The upcoming mobility amendment, 802.16e[3], adds support for nomadic roaming at vehicular speeds. Data rates are envisioned at 2-3 Mbps/user for portable and 1-2 Mbps/user when mobile, covering an area of 5 Km[10].

Figure 4.1 on the next page shows the layer coverage [34] of the 802.16 Standard and appendix B contains a more detailed study of the MAC-layer.

802.16 was originally set to operate between a frequency range of 10-66 GHz, thus sporting high capacity but requiring LOS. Amendments to this has come in the form of support for frequencies below between 2-11 GHz and different physical layers.

The Orthogonal Frequency Division Multiplexing (OFDM)[59, p. 53] physical layer is one of the physical formats supported in the standard and it enables better NLOS performance compared to single-carrier and Code Division Multiple Access (CDMA). The

[1] WiMAX is thus not a standard, but an industry forum managing certification of the different 802.16 flavors.

[2] There are claims of up to 50 Km coverage, but these are considered highly theoretical.

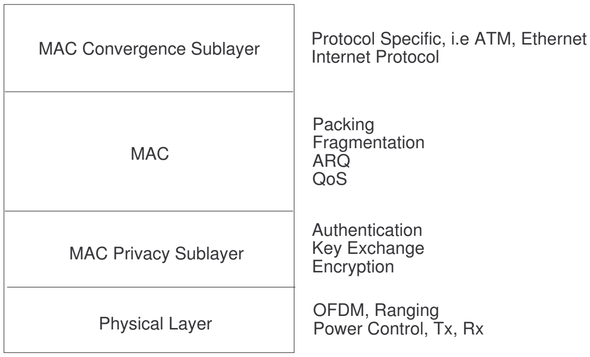


Figure 4.1: Layering of IEEE Std. 802.16

256 carrier ODFM format was chosen by WiMAX Forum for the 802.16-2004 revision of the standard.

The upcoming 802.16e standard goes a step further with Orthogonal Frequency Division Multiple Access (OFDMA)[63], a variant of OFDM, which has the ability to assign a subset of the carriers to specific users. This part of the standard is expected to be ratified during the fourth quarter of 2005, with commercial networks available 2007[18].

The MAC-layer is built to support IP, ATM[3][59, p. 269] and Ethernet traffic through its convergence layer, with five levels of QoS at MAC level in the form of constant bit rate grant, real time polling, enhanced real-time polling, non-real-time polling, and best effort.

Up and downlink can be allocated resources dynamically and has support for TDD and full/half FDD.

MAC Packet Data Units (PDU) can be of variable length, with support for concatenation and fragmentation, which is also available to internal MAC-layer Service Data Units (SDU). This helps save overhead in headers as well as utilize bandwidth and meet QoS demands through packet resizing. Redundant header information can also be omitted with use of Package Header Suppression (PHS)[2, p. 23].

Terminals request QoS according to their needs through a request/grant scheme, with the base station allocating bandwidth dynamically for either user or services. Data traffic over the MAC-layer is scheduled, rather than contention based such as in the case of 802.11. However, ranging and certain types of bandwidth requests can be transmitted over contention periods for increased flexibility and reduced latency.

The MAC-layer also has a security sublayer, performing encryption of MAC PDUs, authentication and encryption key exchange.

The standard comes in different flavors that resembles the 802.11 alphabet soup. Table 4.1 on the facing page contains some of the variations of the 802.16 standard.

[3] Asynchronous Transfer Mode

802.16	Ratified 2001. Core system in 10-66 GHz
802.16a	January 2003. For systems using 11 GHz and below
802.16-2004	June 2003. Also known as 802.16d and contains 802.16a and WiMAX Forum procedures/profiles
802.16e	Q4 2005. Support for nomadic roaming and contains hand off procedures between 802.16 base stations
802.16f/g	Addresses network management, efficient handovers and improved QoS

Table 4.1: The 802.16 Alphabet Soup

4.2 802.16e

The official date for standard ratification of the 802.16e Std. was set to October 2005 but this has been delayed and expected to take place sometime before the end of 2005. This standard specifies roaming of subscriber stations between WiMAX base stations.

The current specification draft[3] adds a new scheduling service called extended real time Polling Service (ertPS), which combines the efficiency of Unsolicited Grant Service (UGS) and real time Polling Service (rtPS). It allows unsolicited bandwidth grants like UGS, but with dynamic size like rtPS. This yields a services class supporting real-time service flows with variable size data packets, suitable for Voice over IP (VoIP) with silence suppression.

The draft defines some previously reserved values in the 802.16-2004 Std. and adds MAC-support for sleep/idle-mode for mobile subscriber stations, paging, locating, power saving classes and defines messages for handover procedures.

Besides a specified handover process the draft also adds two optional handover modes, soft handover and fast base station switching, and defines procedures such as neighbor scanning and topology advertisements.

An issue troubling the 802.16e Std. is the discussion of physical profiles. If the standard will use Scalable OFDMA instead of the OFDM 256 FFT[4] of the fixed version, there is a possibility that hardware updates will be needed for base stations to handle both mobile and stationary subscribers. However some equipment vendors do claim to support software upgrades from fixed to 802.16e[71, p. 1].

The possibility of incompatibility could lead to caution of fixed WiMAX deployment by operators to avoid the risk of costly upgrades between fixed and mobile WiMAX deployments[71, p. 2].

4.3 Network Architecture

With the IEEE Std. 802.16 limiting itself to PHY and MAC-layer, the WiMAX Forum is developing a end-to-end network architecture[19, p. 33], specifying the access/core systems and its functionalities. It contains procedures and protocols for how the network

[4] Fast Fourier Transform

will support e.g mobility, security, internetworking and authentication to a WiMAX subscriber station.

A depiction of the network architecture is presented in the network reference model in figure 4.2. It contains entities such as (Mobile) Subscriber Stations ((M)SS), Access Aervice Aetwork (ASN) and Connectivity Service Network (CSN).

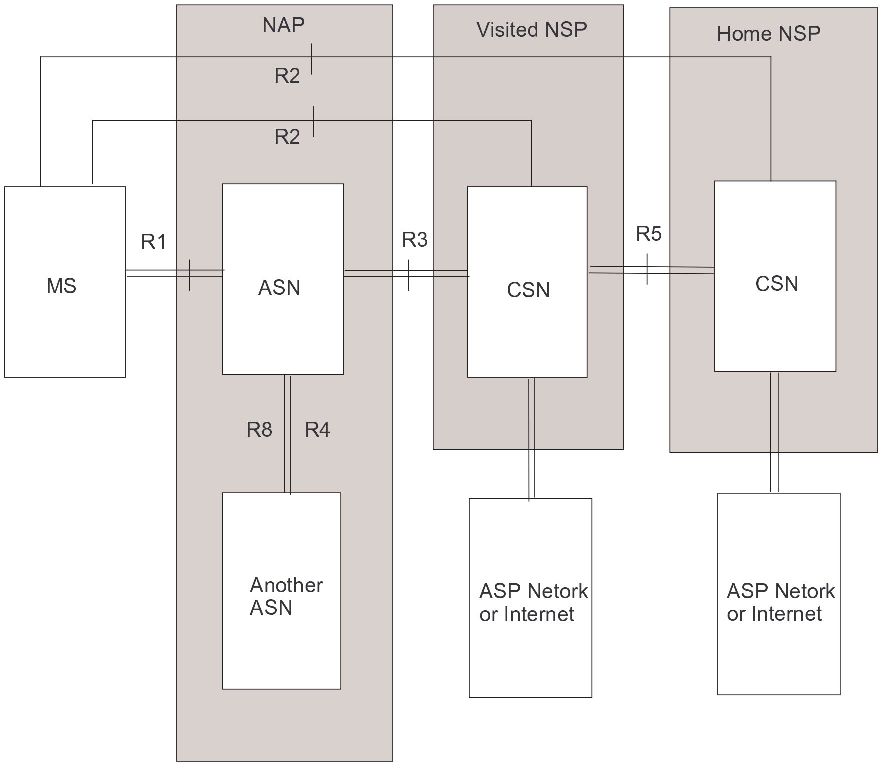


Figure 4.2: The WiMAX Network Reference Model

This reference model also contains interfaces between the different entities. These interfaces define procedures and protocols and serve as logical, rather than physical, links across the entities.

4.3.1 Access Service Network

The ASN consist of one or several ASN Gateways and base stations, supplying WiMAX radio coverage to a geographical area. A ASN manages MAC access functionality such as paging, locating, Radio Resource Management (RRM) and mobility between base stations.

The ASN thus serves as management of the WiMAX radio links only, leaving much of the high level management to the CSN. The ASN can also be used as a proxy, as in the case of proxy Mobile IP (MIP).

The ASN is deployed by a business entity called Network Access Provider (NAP) which provides a SS/MSS with L2 connectivity to a WiMAX radio network and connect users to Network Service Providers (NSP) managing a CSN. The ASN Gateway serves as the interconnection between ASN and CSN.

This logical partition of the access network from the service network enables individual access networks to be deployed, e.g in the case of where several NAP can form cooperation or contractual roaming agreements with each other or one or several NSP.

4.3.2 Connectivity Service Network

A CSN is a set of network functions that provide IP connectivity to WiMAX subscriber stations. The CSN contains gateways for Internet access, routers, servers or proxies for AAA, IP-allocation, user databases, and internetworking devices. It also handles admission and policy control, mobility between ASN and specific WiMAX services such as Location Based Services or Law Enforcement Services.

The CSN is deployed by a business entity called the NSP. WiMAX subscribers enter contractual agreements on e.g services, QoS, bandwidth etc with the NSP and access these services through the ASN it is currently situated in.

The user can then use the service providers network or roam to networks deployed by other companies as long as the home network has a roaming agreement with the visitor network. The foreign ASN uses either its own management functions of the foreign CSN, and proxies them to the home network, or communicates directly with the home network CSN.

4.3.3 Topology Support

The WiMAX technology supports both 2-way PMP networks and a form of decentralized network topology called mesh[2].

Mesh mode differs from PMP as in PMP mode the SSs only talk to the BS and all traffic goes through the BS while in mesh mode all nodes can communicate with each other either directly or by multi-hop routing through other SSs.

A system with access to a backhaul connection is called Mesh BS, while the remaining systems are called Mesh SS. Even though the mesh has a system entitled the Mesh BS, this systems also has to coordinate broadcasts with other nodes.

A mesh can utilize two types of broadcast scheduling. With distributed scheduling, systems within 2-hop radius of each other share schedules and cooperate to ensure collision avoidance and resource grants.

A centralized scheduling mesh relies on the Mesh BS to gather resource requests from Mesh SSs within a certain range and allocate these with individual capacity. This capacity is shared with other Mesh SSs who's data is relayed through the Mesh SSs corresponding with the Mesh BS.

In mesh mode, QoS classification is done on a packet-by-packet basis rather than associated with links as in the case of PMP mode. There exists thus only one link between two communicating mesh nodes.

4.4 Deployment

Choosing frequencies to enable world wide compatible equipment is not easy since every country has its own regulations on the frequency spectrum. The WiMAX Forum group has chosen three frequency bands for the technology, 3.3-3.8GHz, 2.3-2.7GHz and 5.75-5.85GHz [18] where the first two bands are licensed and the last is license exempt. As it seems today most of the continents have at least some available frequencies within the intended ranges, see figure 4.3.

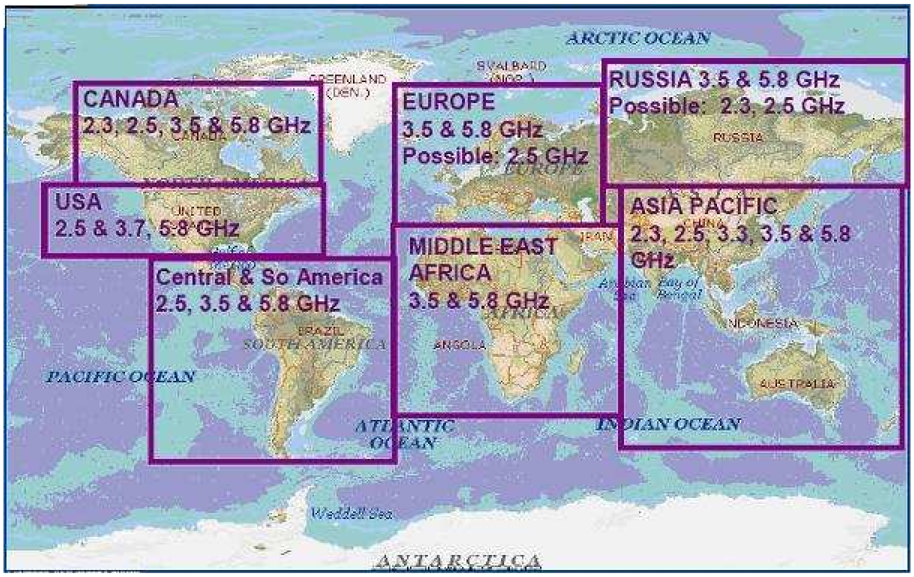


Figure 4.3: Available frequencies, world map [18]

The work on WiMAX is under progress and hence all intended features are not yet ready, see table 4.4 for an estimated timeline. WiMAX Forums first plan is to finish the work with fixed networks and services such as backhauls to Wi-Fi hotspots and outdoor transmissions between base stations and buildings. The next aim of the WiMAX Forum is to complete the nomadic area of use. This includes transmissions to fixed indoor antennas and piconets on enterprise campuses. The mobility market, with features of vehicular speed mobility, will approximately be reached in 2007 or 2008 [18].

Fixed	Outdoor and Backhaul	2005
Nomadic	Metorozone, Fixed indoor and Enterprise campus piconet	2006
Mobile	Mobile	2007/8+

Table 4.2: Estimated timeline for WiMAX

Planned application areas for WiMAX are urban and suburban deployment in dense populated areas and developing countries in need of more bandwidth. Rural areas harder to reach with DSL and cable will easier be connected with WiMAX [21]. Hopes are that with WiMAX it will no longer be necessary to make new and expensive constructions/extensions of/to wired networks in cities.

The planned topology for WiMAX networks using the 802.16-2004 air interface are built up by base stations and both indoor and outdoor antennas. Wi-Fi hot spots in malls, plazas and campus environments can be supported by small outdoor antennas providing a backhaul through the WiMAX technology. Outdoor antennas will also be used to connect rural and sub urban buildings and thus enabling maximum range and coverage. When connecting residential and company buildings in urban centers with a high density of subscribers indoor antennas will be used.

4.5 WiMAX Forum

WiMAX Forum is a non-profit organization that was founded in 2001 and has more than 290 members [18]. The members of WIMAX Forum are a mixture of big and small companies all eager to take part of the standardization and certification attempt.

The work of WiMAX Forum is to assure the interoperability and conformance among broadband wireless access equipment based on the IEEE 802.16 and ETSI[5][16] HiperMAN standards. They strive towards combining the two standards and create a certification process for products that are interoperable with the WiMAX technology. See picture 4.4 for a map over layer harmonization and WiMAX's role[23].

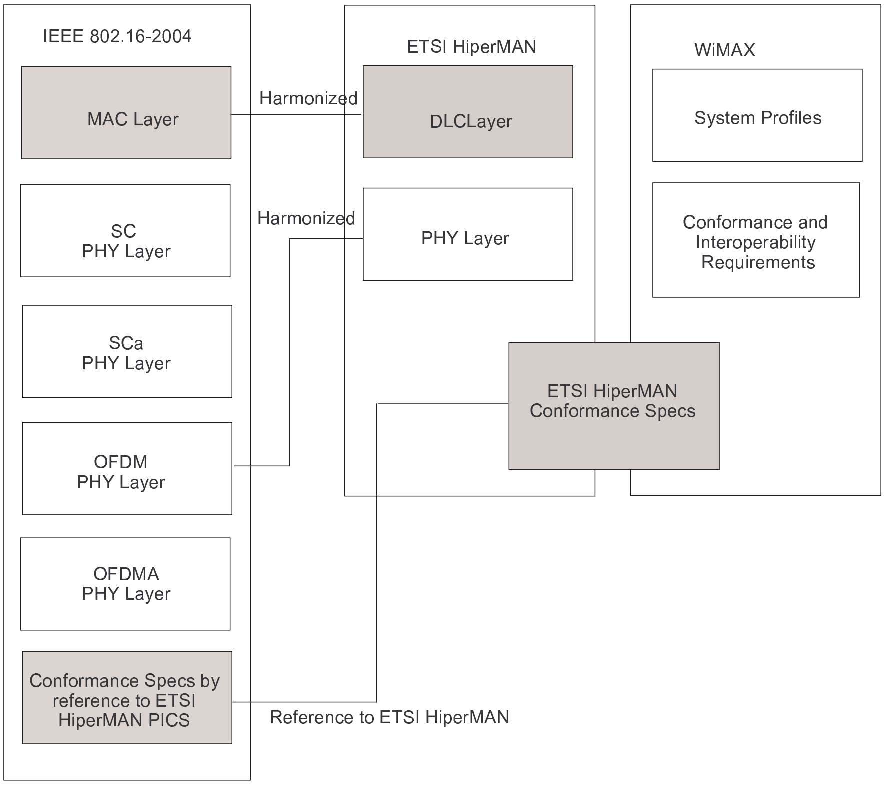

IEEE 802.16-2004
MAC Layer
SC PHY Layer
SCa PHY Layer
OFDM PHY Layer
OFDMA PHY Layer
Conformance Specs by reference to ETSI HiperMAN PICS
Harmonized
Harmonized
Reference to ETSI HiperMAN
ETSI HiperMAN
DLCLayer
PHY Layer
ETSI HiperMAN Conformance Specs
WiMAX
System Profiles
Conformance and Interoperability Requirements

Figure 4.4: Layer Harmonization and WiMAX's role.

To accomplish this the forum works closely with the development team of 802.16 and people from companies within areas such as system integration, equipment manufacturing and applications [21].

[5] European Telecommunications Standards Institute

WiMAX Forum completed their work on test scripts used for examining the interoperability with the WIMAX technology in May 2005 and their certification lab opened in July 2005. During August 2005 the first product tests with hardware were performed and they count on having the first certified product at the end of 2005 [18].

The IEEE 802.16e standard is assumed to be approved in the last quarter of 2005 and the WiMAX Forum aims for certification testing to start in the third quarter of 2006 [18].

Chapter 5

Homogeneous Mobility

Homogeneous mobility is the moving of a MS between networks of the same technology. In this chapter homogeneous mobility will be investigated with focus on horizontal handovers in 802.16e and WiMAX.

In section 5.2 the basic handover mechanisms in 802.16e are described, this includes both methods for gathering network information before a handover and the required functionality to handle the actual handover. Section 5.3 explains how WiMAX is managing the horizontal handovers. The network structure used is described and terms such as anchoring and intra/inter ASN handover are explained. Interesting studies on similar matters are presented in 5.4.

5.1 Background

Homogeneous mobility is less complex then heterogeneous mobility since no other technologies or systems need to be considered when moving inside a homogeneous network. This mobility only requires support in the own network and this lead to less complicated solutions for roaming mechanisms since measurements like e.g. Signal to Noise Ratio (SNR) can be utilized. The demand for mobile systems have grown out of peoples need to move around or travel while still connected. To fulfill this need a great deal of effort have been and still are put in to the development of new standards.

Mobility within connection oriented speech systems like Global System for Mobile Communications (GSM) and UMTS is widely used in the world today. Almost every person that own and use a cell phone uses the homogeneous mobility in those systems daily. When the user moves with the cell phone it will most likely get out of range from the current BS and the system will perform an intra system(horizontal) handover. This transition between BSs is most often unnoticed by the user and this is where well developed horizontal handover procedures are important. If the system do not have well developed procedures for the handover the user will notice the interruption in the ongoing conversation.

When it comes to packet based networks the mobility aspect is a rather new area and there is still a lot of work needed to make the services used as efficient as possible. IEEE[28] are trying to address this area with development of wireless MAN air interfaces such as 802.16e and 802.20. Non profit organizations such as WiMAX Forum[21] takes this work to a higher extent with suggestions to higher layer architecture and handover support.

There are high expectations on WiMAX and its ability to support mobility and it is an interesting area to study. This chapter will look into the roaming world of WiMAX and this naturally lead to a closer look at its different handover possibilities as well as the handover support in 802.16e.

5.2 802.16e handover

The 802.16e standard is the base for mobility in WiMAX. It supports handovers through procedures and functions at BS/MS level. The standard defines the means for gathering information and performing a handover but the decision whether to perform a handover or not is left out. Example situations on when to perform a handover is when the MS need to switch BS to receive higher signal quality or when the MS can obtain improved QoS from another BS. More detailed information about 802.16e handovers can be found in the IEEE Std. 802.16e[3]

5.2.1 Network Topology Acquisition

To be able to perform a handover the MS need to acquire information about the network. This can be done with network topology advertisements or scanning of neighbor BSs with or without the optional association procedure[3, p. 170-174].

Network topology advertisements are broadcast messages sent out by all BSs. These messages contains information about neighboring BSs and their channels. This information will simplify the MS synchronization with a new BS since there is no need for the MS to listen to the target BS's DCD/UCD[1] messages. The serving BS may receive the required information about its neighbors through the backbone.

Additionally the MS may use time to scan its environment for potential target BSs. In this scanning the MS localize BSs and it may investigate the quality of their channels. The MS requests scanning intervals from the serving BS and may start scanning when permitted by the serving BS. During the scanning interval the serving BS assume the MS to be in scanning mode and it may buffer data incoming to the MS. When the MS exits the scanning mode the serving BS start to send the buffered data. The MS can at any time terminate the scanning by starting to send PDUs again. When a BS receives PDUs from a MS supposed to be in scanning mode it will assume that the MS exited scanning mode and resumed normal operation. To reduce overhead due to many scanning requests the MS can ask for a group of scanning intervals. The intervals will be interleaved with periods of normal operation.

The association procedure is an optional feature in the standard and it can occur during scanning.There are three levels of association, scan/association without coordination, association with coordination and network assisted association reporting. The goal of the association is to enable the MS to collect and store information about BSs. The gathered information is saved during a reasonable period of time and it can help the MS further in decisions regarding handovers.

5.2.2 Handover Process

The handover process consists of six different stages; cell reselection, handover decision and initiation, synchronization to target BS downlink, ranging, termination of service

[1] Downlink/Uplink Channel Descriptor. Messages containing information about the downlink and uplink characteristics

and handover cancellation.

The cell reselection is the stage where the MS acquire information about BSs in the network. The information is used in evaluation of the possibility to perform a handover. This can be done by using the information in the network topology advertisements or require a scanning interval to obtain the needed information. The cell reselection phase does not need to occur in relation to a handover decision.

The initiation of a handover is the decision to migrate the MS from the serving BS to a target BS. This decision can be triggered in the MS as well as in the BS. To commence the actual handover the requesting party sends a handover request which will trigger a sequence of handover specific messages to be sent between MS and BS[3, p. 178].

To establish communication with the target BS the MS need to synchronize to its downlink channel. During this phase the MS receives downlink and uplink transmission parameters. If the MS previously received information about this BS (through the network topology acquisition) the length of this process can be shortened.

When the MS is synchronized to the channel it need to perform initial ranging or handover ranging. Ranging is a procedure where the MS receives the correct transmission parameters, e.g time offset and power level. The target BS may obtain information about the MS through the backbone and depending on the target BSs knowledge about the MS some parts of the ranging process may be omitted.

Termination of services at the serving BS is the last step in the handover process. The serving BS will terminate all connections associated with the MS and remove all information in queues, counters etc.

During the handover the MS have the right to cancel the handover and resume normal communication with the serving BS. The only condition is that the MS do not try to cancel after a specified time have elapsed.

5.2.3 Soft Handover and Fast BS Switching

During a soft handover the MS will listen/transmit to several BSs at the same time. This is done by BSs sending the same PDUs to the MS and they all listens to the MS transmissions. The MS will perform diversity combining on the signals received from the BSs and the BSs will in turn perform diversity combining among them to get the uplink PDUs.

In the fast BS switching the MS will listen/transmit to only one BS within an active set of BSs. The BS the MS is listening to at the moment is called the anchor BS. The MS can change the anchor BS on a per frame basis and it will choose the new BS from the active set.

In both types of handover the MS will maintain a list of BSs (the active set) that are involved in the handover. The BSs in the active set are BSs that could be suitable as target BSs.

All the BSs in the active set need to be synchronized to each other to be able to support soft handover and fast BS switching. The reason for this in the soft handover case is that the MS will listen to all the BSs at the same time. If they send different PDUs or send the same PDU but with a difference in time the MS can not interpret the received data. The same apply to the fast BS switching case even though the MS only listen to one BS since the switching can occur from frame to frame. If the BSs send the frames at different points in time the MS can loose/receive duplicate frames when switching.

What type of handover to use is agreed upon in the handover request/response

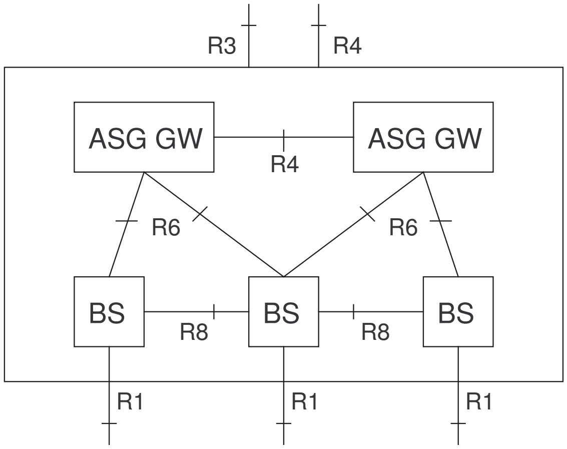


Figure 5.1: ASN Decomposition

messages. Both the soft handover and the fast BS switching are optional in the standard and in the case where they are implemented they can be disabled.

5.3 WiMAX Handover

The WiMAX architecture extends the 802.16 standard and that also includes the mechanisms for handovers. While the 802.16 standard provides support for handover between base stations WiMAX offer protocols for handover higher up in the network structure. The WiMAX architecture shall support mechanisms such as intra/inter ASN handover, roaming between NSPs, seamless handover at vehicular speed and micro/macro mobility. This section will study the architecture and its handover procedures more thoroughly with the focus on intra/inter ASN handovers. For more depth in the subject see the WiMAX draft[19, chapter 7.5].

5.3.1 Access Service Network

Inside an ASN network entity there are at least one ASN Gateway (ASN GW) and a base station. The BS handles the connection to the MS while the ASN GW takes care of the contact with the CSN. An ASN GW can be associated with one or more BSs and a BS can have relations to one or more ASN GWs see figure 5.1. This segmentation of the ASN enables multi vendor systems where different vendors can produce different parts of the ASN and they still function together.

Depending on which role a BS or ASN GW take on in a handover they get different names, see table 5.3.1 on the next page. The BS in charge of the MS before the handover is called the serving BS and the ASN GW the serving BS forwards the data to is the serving ASN GW. The BS and ASN GW associated with the MS after the handover are the target BS and target ASN GW respectively. The term anchoring ASN GW is used when an ASN GW relays MS data to the serving ASN GW.

Serving BS	The BS related to a MS before handover.
Target BS	The BS associated to a MS after handover.
Serving ASN GW	The ASN GW corresponding to the serving BS.
Target ASN GW	The ASN GW connected to the target BS.
Anchor ASN GW	The ASN GW receiving CSN data addressed to the MS.

Table 5.1: BS and ASN GW roles

Anchoring

The anchoring ASN GW is the network's or CSN's attachment to the MS. Incoming data will be sent to the anchoring ASN GW and the CSN does not need to know at which ASN GW the MS's current BS is located. The forwarding of data to the serving ASN GW is performed by the anchoring ASN GW. This makes the mobility of the MS transparent to the CSN and the need to change IP-address becomes less frequent. In the case where the serving ASN GW is receiving the data directly from the network the serving ASN GW is also the anchor. The anchoring ASN GW does not need to be any of the serving or target ASN GWs.

ASN Reference Points

To identify the different interfaces used to communicate within an ASN, with the MS and the rest of the network a number of reference points are introduced [19], see figure 5.1 on the facing page. These reference points define the set of protocols and procedures needed in the communication. Most of the reference points are logical mappings but when, as in the case of R1, the functional entities are in different physical devices the reference point refers to a physical interface.

R1 and R3 are the reference points used in communication with entities outside of the ASN while R6 and R8 are used inside an ASN. The R4 interface is used both inside and outside of the ASN since it is the logical link between ASN GWs regardless of whether they are within the same ASN or in different ASNs. R1 is the physical interface between the MS and the serving BS and R3 is the logical link between ASN GW and CSN. The communication among BSs is handled through R8 while the BS-ASN GW interaction goes via R6.

5.3.2 Functional Decomposition

Three functions have been defined to help with the management of moving MSs. These functions deal with the handover, MS context and data to/from the MS. The Data Path function takes care of path setups and the actual data transmissions. The information and context regarding a MS and the exchange of this in the backbone is handled by the Context function. The most interesting function from a handover point of view is the Handoff function. This function deals with the signaling and decisions associated with the handover. All these functions work in a peer to peer manner and the peers act either as a serving, relaying or target function. This is a general model and the WiMAX document[19] does not state where these functions reside.

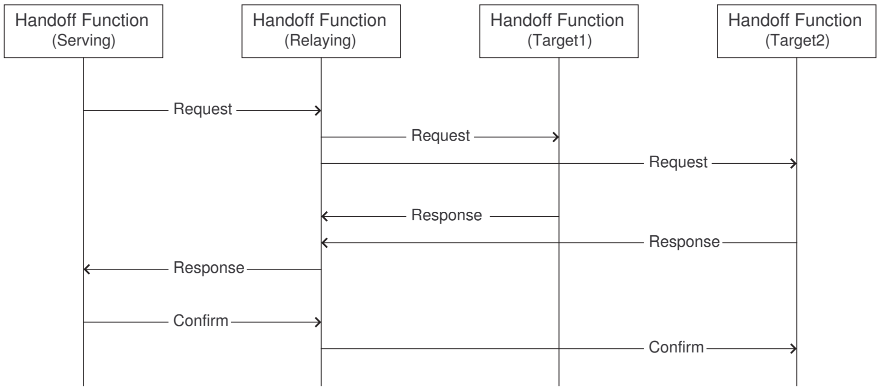

Handoff Function (Serving)
Handoff Function (Relaying)
Handoff Function (Target1)
Handoff Function (Target2)
Request
Request
Request
Response
Response
Response
Confirm
Confirm

Figure 5.2: Handoff function network transaction

In the case of the Handoff function the peer to peer mechanism imply that the serving Handoff function will request a handover. The target Handoff function will take care of the request and send a reply. If it is needed relaying Handoff functions will act in between the two. A possible scenario with the Handoff function is pictured in 5.2.

The Handoff function support mobile initiated handover, network initiated handover, fast BS switching and soft handover according to the 802.16e standard.

5.3.3 Intra ASN Handover

The intra ASN handover is performed between BSs (or sectors within one BS) belonging to the same ASN, see figure 5.3 on the facing page. The BSs can be connected to the same ASN GW or different ASN GWs (within the same ASN) it will still be an intra ASN handover. If there is only one BS within an ASN an intra ASN handover can not be performed unless the BS has several antenna sectors.

The purpose of the intra ASN handover is to minimize the delay and data loss during the MS's transition between BSs. If the MS are using services such as IP or MIP there will be no need for a change of IP-address after the handover since the movement of the MS is not visible from outside the ASN.

The reference points involved in an intra ASN handover is R6, R8 and in some cases R4. R4 is only involved when the target BS is connected to another ASN GW than the serving BS. Exactly what messages that will be sent over which interface during an intra ASN handover is not specified in the WiMAX draft[19].

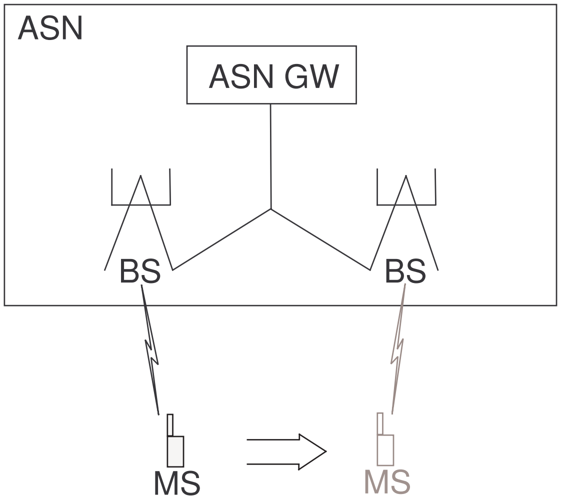


Figure 5.3: Intra ASN Handover

Target ASN GW	Anchor ASN GW	Outcome
No	No	No re-anchoring
No	Yes	Re-anchoring
Yes	-	Re-anchoring

Table 5.2: Re-anchor decisions and outcome

5.3.4 Inter ASN Handover

An inter ASN handover is a handover between BSs not part of the same ASN, see figure 5.4 on the next page. During an inter ASN handover ASN GWs in separate ASNs need to coordinate their actions to make the handover smooth to the MS. There are two possible ways of dealing with the data flow during an inter ASN handover, anchoring and re-anchoring. The purpose of anchoring is to avoid an path update and hence a redirection of the data path, where in the re-anchoring case an update will be performed.

The decision to anchor or re-anchor the data path is made by the target or anchor ASN GW and there are three different decision procedures with two possible outcomes 5.3.4. Either both parties can decide that a re-anchoring is not needed or one of the ASN GW decides that it wants a re-anchoring. If the target ASN GW wants a re-anchoring the anchor ASN GW will follow that decision and vice versa. It is always the target ASN GW who will make its decision first. What this decision is based upon is implementation dependent and not included in the scope of the WiMAX document[19].

5.3.5 Changes

The WiMAX draft has changed during the time span of this thesis work and this section address some of the changes regarding horizontal handovers.In the WiMAX draft[20] some minor changes/additions to the handover section have been made in comparison

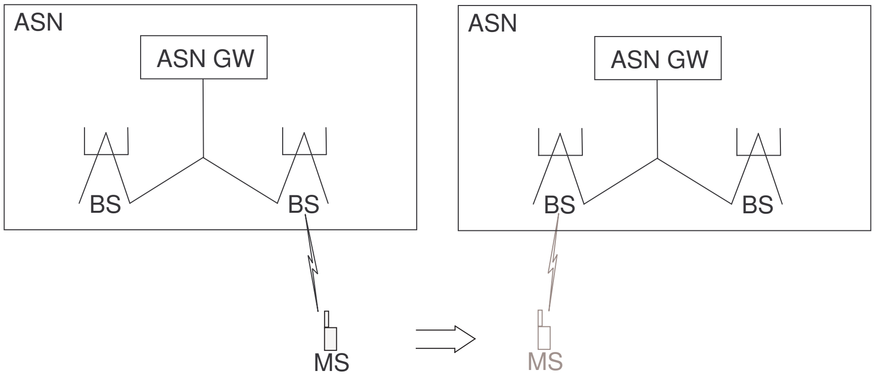


Figure 5.4: Inter ASN Handover

to the older draft[19]. The main ideas are the same but some concepts have changed.

In the earlier draft[19] the focus of handovers have been on intra/inter ASN handovers and thus from and ASN point of view. In the new draft[20] this has been put into the shadows a bit and the new focus is instead on handover with or without Care of Address[2][59, p. 308] update. Intra ASN handover falls under the category handover without CoA update since there is no need to receive a new IP address during those handovers. The inter ASN handover is in the case where anchoring is used a handover without CoA update but otherwise it is a handover with CoA update.

A new statement is that a handover can, from a QoS point of view, be either controlled or uncontrolled. A controlled handover has to follow a set of conditions and if any of these are broken the handover is considered to be uncontrolled.

It is also explained how WiMAX will support the 802.16e association procedures possibly needed during scanning in the network topology acquisition phase.

The section about functional decomposition have been increased with sequence diagrams showing the cooperation between the functions. They show how the functions work together and trigger eachother.

5.4 Related Work

Research in mobility is a hyped area right now and a lot of work is done by many different persons and companies. Associations like IEEE refine and evolve existing standards or produce new ones. Companies put together working groups with each other to assure interoperability among equipment and a lot of students make thesis works with studies of and suggestions to existing/coming standards.

This section brings forth some of the interesting projects that have been found during the course of this thesis.

[2] A temporay IP-address for a mobile node

5.4.1 802.16d with mobility support

The 802.16d (now called 802.16-2004) only covers fixed networks which have lead to IEEE conducting work in a mobility version, 802.16e. Another approach to solve the lack of mobility in 802.16-2004 have been made in "Mobility Support for IEEE 802.16d Wireless Networks"[42]. Their approach has the goal to enable mobility in 802.16-2004 without modifying the standard. To accomplish this mobility, tools such as hierarchical MIP, selected parts of the 802.12-2004 initialization process and a in the standard predefined message have been utilized.

An existing message have been chosen to serve as handover request and acknowledgment. The MS will send this message to the BS when it wish to perform a handover and the BS will respond with the same message. The message chosen does not affect either of the MS or BS if they do not have the mobility functionality, it simply tells the receiver to carry on as usual.

In the initialization phase during the handover the target BS skip parts such as authentication and exchange of encryption keys and this information is instead sent through the backhaul. By reusing only necessary parts of the initialization process the delay during the handover is minimized.

In this solution the restriction that only the MS can request a handover have been made. This due to the fact that a MS can stay silent during a period of time when it has nothing to send. The problem lies in the BS interpretation of this silence; is the MS not sending or is it out of transmission range? If the MS takes care of the handover request the BS do not need to figure out the answer to this question.

Further in this article they show that the PHY-layer of 802.16-2004 is suitable for mobility. Their calculations show that such a system should be able to support moving terminals with limited speed. Combining this information with the suggested mechanism for managing handovers show that it is possible to get mobility in 802.16-2004.

5.4.2 802.16e with seamless mobility

Even though IEEE just completed the work on the 802.16e standard there have already been suggested a mechanism for enabling seamless handover in networks based on the standard. It is described in "A Seamless Handover Mechanism for IEEE 802.16e Broadband Wireless Access"[7]. The mechanism is called Last Packet Marking (LPM) and integrates MAC-layer handover with the Network layer handover to decrease the handover effects on TCP service performance.

LPM mainly consists of the handover support in 802.16e, a few new messages[3] and buffering of packets at BSs. The messages added contains information about routing. The network model used consists of BSs and a hierarchy of routers connecting BSs, see figure 5.5 on the next page. The main idea of LPM is to send incoming MS packets to both serving BS and target BS from the point in time when the MS is thinking of performing a handover. The target BS will buffer incoming data and forward it to the MS when the handover is complete.

LPM simulations was performed on a 802.11 WLAN since the selected network simulator does not have the 802.16e implemented. The authors claims that during the circumstances the chosen alternative works as good as a 802.16e implementation would have. To evaluate the LPM mechanism the throughput of TCP packets per second was

[3] New messages would require a change in the 802.16e standard.

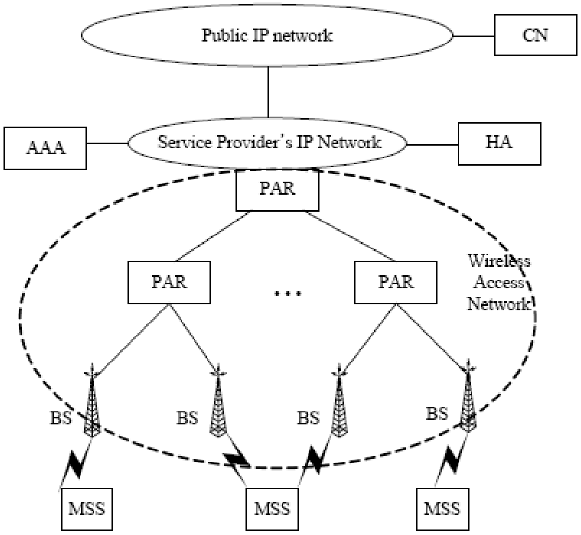


Figure 5.5: Wireless Access Network Model[7, p. 4]

measured. The results shows that the system suffered from large throughput drops during handover when not using LPM. With LPM the handover affects on the throughput was minimal. This shows that LPM is an effective and useful mechanism.

5.4.3 802.20

802.20 is a new standard developed by IEEE especially for use in mobile networks[4]. Due to this the standard can have some advantages over 802.16e (where the mobility is added upon a standard for fixed wireless networks) since problems can arise when making extensions to an existing standard. 802.16-2004 is build for fixed wireless networks and 802.16e has focus on mobility. Flaws from the original standard can tag along and some design choices might not optimal for the new target area. But on the other hand it can sometimes be faster to rework something already present then to start from scratch. This can lead to 802.16e products reaching the market before 802.20 compliant equipment does[44].

The 802.16e and 802.20 standards have many similarities and both are aiming to fill the gap between high mobility cellular networks and high data rate WLANs but there are some slight differences[3, p. 1][4].

802.20 operates in frequencies below 3.5 GHz and the standard specifies the PHY and MAC-layers of the air interface. It is constructed to provide peak data rates higher than 1 Mbps/user in cells with up to 15 km coverage[68][44]. This will be done in speeds up to 250 km/h, that means slightly higher speeds than the 802.16e will be able to handle. This high speed makes it possible to e.g. deploy 802.20 in high speed trains[44].

An additional thing that pleads for 802.16 is the fact that the WiMAX Forum working group prolongs the standard with its WiMAX End-to-End Network Systems Architecture and fights for enabling interoperability between network equipment from different vendors[21]. The vendor products can be certified and this can help 802.16 on its road to success.

802.20 do not have such a pronounced organization that brings the standard in to the bigger picture but it might be on its way. Telcordia have invited companies to a working group that will process the 802.20 standard. In a broad outline they will study coexistence between 802.20 and other systems in licensed bands below 3.5 GHz. The working group will produce conformance and interoperability specifications as a basis for certification testing. They will try to find procedures and services to enable cross-vendor interoperability and interoperability to other systems[30]. This group could be 802.20's WiMAX Forum.

5.5 Conclusions

Mobility in networks is an evolving area and it is interesting to take part of the results produced by organizations such as IEEE and WiMAX Forum.

The support for horizontal handovers seem well developed in both 802.16e and WiMAX. The basic hard handover is supported as well as a soft handover and fast BS switching. This allows for using the most appropriate handover both from MS and network point of view. If the MS have connections that are e.g. sensitive to lost packets or delay it can be appropriate to use the soft handover or fast BS switching. How ever if the performance is not that important the hard handover can be used and network resources will be spared.

The 802.16e standard aims to provide mobility through roaming which is shown through the defined mechanisms for handover. A MS have the possibilities to gather the information needed to make a handover decision and when the decision is made the MS can chose the type of handover best suited for its needs. The mechanisms in the MAC-layer is further augmented by support of handover and handover information gathering in the core network. The 802.16e MAC-layer is complex with a lot of information mentioned but not explained. Many details is called to be "out of scope" and left to the higher layers.

WiMAX has 802.16e as its air interface and the architecture further refine the mobility mechanisms in the standard and support them in its core network.

In comparison to other technologies WiMAX have both already well used handover mechanisms and some new features. The hard handover is something that is used by most technologies that support mobility since it is simple and do not use additional network resources. The fast BS switching is a newer procedure and seems promising when supporting a more reliable connection QoS guarantees.

Some standards used today (Bluetooth, certain implementations of WLAN etc.) support mobility, or rather roaming, on a smaller scale i.e. the MS is wireless but there is no extensive support for handovers. The MS does not have any means to gather information about potential handover targets or to get support from the network in the transition. In these systems the MS simply disconnects from a BS and reconnect to another when the signal is lost. The handover is in this case somewhat reactive while WiMAX provides mechanisms for proactive handovers.

Systems like GSM and UMTS have been supporting mobility in their networks from the start and they have well functioning mechanisms for handovers. GSM have intra/inter cell/BSC[4]/MSC[5] handovers[59, p. 102-103] similar to the WiMAX inter/intra ASN handovers. The notion of a soft handover can be found in both UMTS and WiMAX and is a good way to provide a more robust handover.

WiMAX with its base in the 802.16e standard have developed promising support for homogeneous mobility.

[4] Base Station Controller

[5] Mobile Services Switching Center

Chapter 6

Heterogeneous Mobility

As a second study of the mobility capabilities of WiMAX, the goal of this section is to present an evaluation of the network architecture and its foundation in the IEEE Std. 802.16 with respect to mobility and inter-networking with other types of networks.

After an introduction to the UMTS and Wi-Fi network structure, focused on the mobility aspect, WiMAX is discussed and issues of interest are highlighted.

6.1 Background

With the emergence of the third generation cellular networks, offering greater coverage and mobility but lower data rates compared to fixed wireless broadband networks such as Wi-Fi, the inter-networking and mobility between these two network technologies has spawned new research areas.

The differences in the two technologies regarding QoS, range and data rates has formed a void in the topology of network technology, a void possibly filled by BFWA systems and in this paper the case of WiMAX.

As an example, a user can enjoy the superior range of a BFWA network compared to Wi-Fi hot spots when mobile, and through policy based decisions such as mobility, cost or performance requirements be transfered to a Wi-Fi hotspot when one is available. BFWA inter-networking with UMTS could be used for almost global continuous connection with reduced data rate when outside the range of BFWA base stations.

These three different network technologies could complement each other in the borderline of a MAN network where users are mobile but still require either a persistent connection or high data rates, and all of this with ubiquitous movement between the different technologies.

The consumer increase of 3G devices and Wi-Fi market development[1] is reason enough to investigate the inter-communication of WiMAX with these technologies, as telecommunication providers aim to protect their investments and extend customer services through new types of media.

This chapter will look at such heterogeneous mobility and what support is offered in the IEEE Std. 802.16 and the WiMAX network architecture.

[1] ZDNet Research predicts WLAN market to reach $5 bln by 2006 and 70 mln 3G users by year-end 2005[58].

6.2 UMTS

The Universal Mobile Telecommunication System (UMTS) is the IETF standard for the third generation of cellular networks, intended to provide global coverage of telephony, message and data services. The Third Generation Partnership Project (3GPP) was formed from a number of telecommunication standardization organs to further the technical development of 3G networks.

According to UMTS Forum[61] there are currently more than 30 UMTS networks deployed and operational in 25 countries, with the first commercial network launched in Japan 2001.

6.2.1 Features

UMTS is the telecom industry's move towards adding bearer services for data transfers together with tele services. This was proposed by ETSI to be a development of the GSM systems, providing more efficient deployment of the UMTS network on top of an existing infrastructure[59, p. 136].

UMTS utilizes WCDMA as radio access technology with ETSI basic requirements on data raters found in table 6.1.

144 kbits/s	Rural outdoor access and satellite
384 kbits/s	Urban outdoor access at max speeds of 120 km/h
2 Mbits/s	Indoor and short range outdoor at speeds of 10km/h

Table 6.1: UMTS Data rates

As new services are added there is a need to prioritize. Voice services are sensitive to delays but tolerate losses, while data services such as e-mail require lossless transfers, but with less requirements on latency and data rate. The network services classify traffic according to four different classes, found in table 6.2[72].

Conversational	Voice, video telephony
Streaming	Multimedia, webcast, streaming video
Interactive	Web browsing, network gaming
Background	E-mail, SMS, downloads

Table 6.2: UMTS QoS Classes

The security features of UMTS are built on the GSM system, but with some extensions that take the previous weaknesses of GSM in consideration[38].

The security is present through authentication, confidentiality and anonymity to provide user security and mitigate the issues associated with wireless telephony. Mutual authentication between User Equipment (UE) and base stations ensures that both parties know who they are communicating with.

Contrary to GSM, UMTS handles signaling and data integrity which protects against attacks by the use of false base stations. The data can be encrypted according to negotiated cryptography suits and keys, enabling traffic confidentiality. This protection extends to the Radio Network Controller (RNC) and not just between UE and BS[38].

UMTS also hides user identities, location and service patterns. Universal Subscriber Identity Module (USIM) enables user specific authentication, while International Mobile Equipment Identity (IMEI) is used to authenticate specific devices[59, p. 666].

The architecture also makes use of IETF protocols for IP traffic, such as IP Security (IPsec), between networks[27].

Development

The standardization of future UMTS developments has been split into several substandards, called Releases.

Release 5, finalized March 2002, contains changes of the network core to an all-IP-core, IP-multimedia services (IMS) supported by the IETS session initiation protocol (SIP). The release also adds High Speed Downlink Packet Access (HSDPA) with speeds up to 8-10 Mbit/s[59, p. 141].

In April 2005 the Global mobile Suppliers Association (GSA), a forum representing GSM/3G suppliers world-wide, announced[48] the finalization of Release 6. This release will evolve IMS services, adding Multimedia Broadcast Multicast Service (MBMS) for multimedia distribution to multiple receivers. The release also adds inter-networking features for WLAN.

To complement the downlink HSDPA, Release 6 supplies High Speed Uplink Packet Access (HSUPA), adding high speed symmetrical data communication.

A future extension to 3G, called Super 3G is said to support data rates of up to 100 Mbps[2]. Several articles associate Super 3G with the HSDPA technology, which is faulty[41] as Super 3G is a new technology taking speeds beyond HSDPA.

6.2.2 Network Architecture

The UMTS network reference architecture consists of three major components, depicted in figure 6.1[59, p. 142].

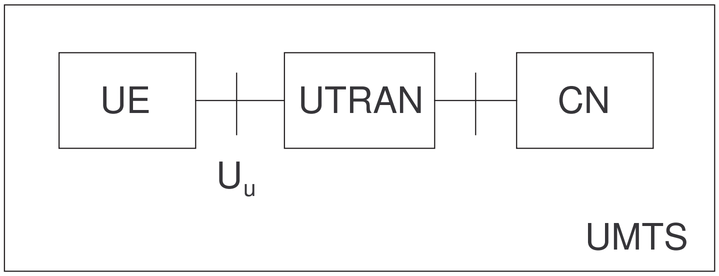


Figure 6.1: UMTS Reference Architecture Main Components

The Universal Terrestrial Radio Access Network (UTRAN) handles radio resource management, cell mobility and radio access for the UE situated in the radio cell covered by the UTRAN. UTRAN consists of antennas, controlled by a component called Node B, which in turn is managed by a RNC.

[2] Such data rates are purely theoretical

The radio access network, UTRAN, is connected to the Core Network (CN), which consists of functions for user management such as AAA, databases and location registers. Furthermore, the CN consists of gateways for inter-networking and serves as infrastructure for packet and circuit switched traffic.

6.2.3 UMTS and WiMAX

In a general observation, the network reference architecture of both WiMAX and UMTS are very similar. They both contain well-defined interfaces between user equipment, base stations and the core infrastructure. A direct mapping of ASN to the UTRAN and CSN to the core network is possible since the functionality basically are distributed in the same fashion. Figure 6.2 depicts both the UMTS and WiMAX architecture.

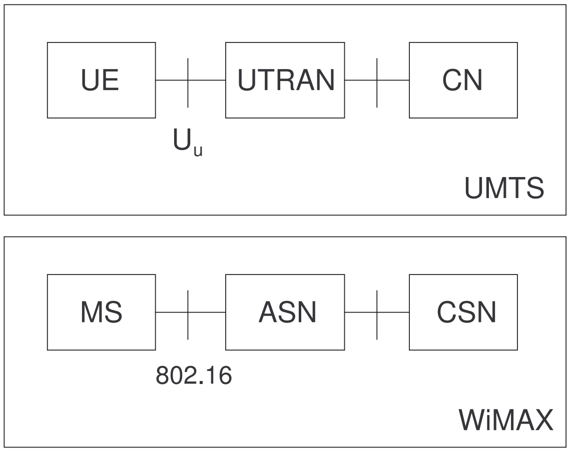


Figure 6.2: UMTS And WiMAX Architecture Components

This support for management in the CN of both WiMAX and UMTS is of importance for service portability for roaming users from either network[74] and this network intelligence help facilitate the heterogeneous mobility.

When transferring a user between the two networks there is also the factors of diversity in data rates and end-to-end delay to account for. WiMAX technology supports five classes of QoS that, all similar to UMTS QoS parameters and thus lend to mapping of traffic flows moved from one network to another or between the networks.

A possible mapping is displayed in table 6.3.

UMTS	WiMAX
Conversational	Unsolicited Grant Services
Streaming	non real time Polling Services (nrtPS)
Interactive	(extended) real time Polling Services
Background	Best Effort (BE)

Table 6.3: UMTS/WiMAX QoS Mappings

With the HSDPA and HSUPA amendments to the original UMTS networks and the coming Super 3G networks, data rates between the two networks will not be as diverse. Some degradation of e.g video quality and regular download speeds could however be expected, when moving from WiMAX to UMTS, though still maintained by the UMTS system[15][54].

Security has been included from the start in WiMAX with the MAC security sublayer, supporting both DES3 and AES encryption standards[69]. Much like UMTS WiMAX also supports mutual authentication of a client and the WiMAX network.

In the WiMAX network reference document, the following is stated[19, p. 26]:

> ”The architecture shall support data integrity, replay protection, confidentiality and non-repudiation using maximum key lengths permissible under global export regulations.”

There is thus spoken support for security measures to protect traffic in much the same extension as that of the cellular networks.

When considering mobility between a WiMAX and UMTS network the coupling between the two also needs to be studied. The WiMAX network architecture has been made to fit the 3GPP model of inter-networking with alternate access networks[19, p. 39]. Since UMTS has undergone standardization to internetwork with Wi-Fi access networks this provides finalized requirements and standards for WiMAX to utilize and increases interoperability.

The difference with the WiMAX access network compared to the Wi-Fi analogy, is that the WiMAX ASN does not supply IP-connectivity, which adds the requirement of a logical CSN inter-networking unit with DHCP and DNS to be located inside the ASN.

There are two types of coupling between networks, loose and tight, depending on how integrated the networks are with each other. The couplings introduce different issues in both networks[19, p. 39][49].

Loose Coupling

The WiMAX network architecture suggests a loosely coupled integration of the two technologies[19, p. 41]. Figure 6.3 on the next page shows a model of such a coupling.

In this scenario access to 3G AAA services is granted through PDG edge routers connected to the WiMAX network or routed through the Internet. The traffic of the WiMAX network is separated from the UMTS network and it can supply its own mechanisms for mobility, authentication and billing.

This could be the case where separate WiMAX business entities and network providers enter roaming agreements with 3G providers. Loosely coupled networks can thus be used to over time extend the coverage offered by a 3G network in a less restricted fashion and can result in less management and deployment costs for the 3G providers extending user data services.

It also means that mobility needs to be managed at a higher level by protocols such as the IEEE Std. 802.21[5] for Media Independent Handover or MIP. It is thus more complicated to add network intelligence in managing vertical handovers between the networks, possibly increased latency in the handover phase.

Many loosely coupled networks would require capacity overhead in management traffic to monitor status and capacity when supporting network initiated vertical handovers. In order to facilitate network initiated handovers, each network must share and monitor admission control information and resource management in order to make the right decision while sustaining user QoS.

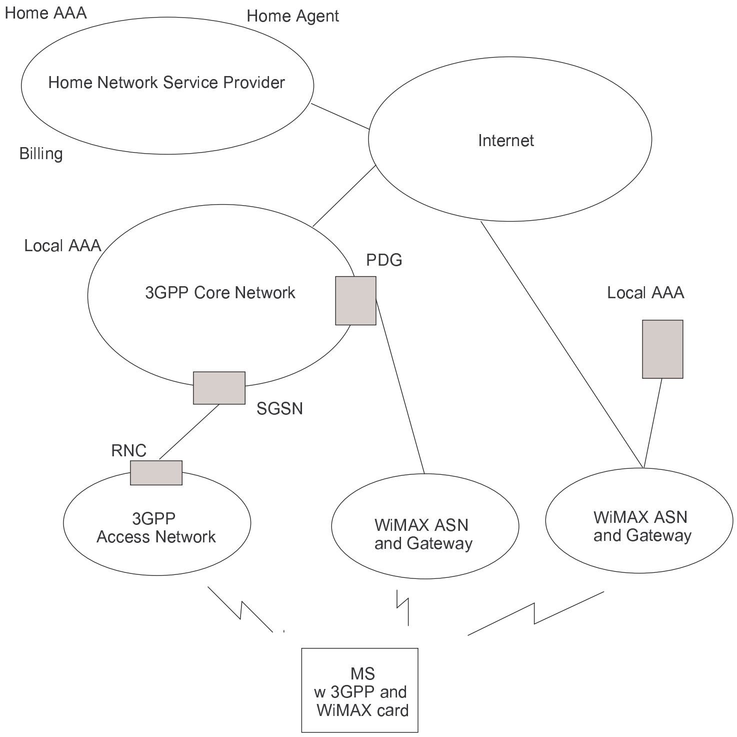


Figure 6.3: WiMAX/UMTS Loose Coupling

Tight Coupling

In a tightly coupled network, the WiMAX and UMTS would share the core network components, such as gateways, AAA and infrastructure. The WiMAX network is connected to the UMTS Gateway GPRS Support Node (GGSN) and the packet switched domain of the UMTS core network. The WiMAX network appears as a RNC to the UMTS network, and a gateway emulates RNC behavior. This requires the user equipment to run UMTS protocol stacks.

Tight coupling has the advantage that the same billing, authentication and mobility protocols can be reused for the WiMAX network. There is also the possibility to use the tight coupling when addressing vertical mobility, as the core network has first hand knowledge of resources in both networks and can support network initiated vertical handovers.

Tight coupling suggested[47] for UMTS and Wi-Fi networks involve a mixed solution of mobile IP for handling mobility together with a mobility gateway. The mobility gateway intercepts all traffic and ensure it is directed to the right network (UMTS or Wi-Fi). This is only necessary when the addressing scheme is not mutual.

An important consequence to tight coupling is that WiMAX traffic is introduced to the UMTS network. The characteristics of cellular and IP traffic is quite different and some parts of the core network might have to be adapted to handle new types of load and traffic patterns to avoid capacity conflicts.

Some[49] argue that this approach is only practical when the networks combined are owned by the same operator and the integration is done as a form of patch to the existing system to reuse old network components.

This approach suffers from significantly less flexibility in coverage extensions compared to the loose coupling, but allows the WiMAX network to be tailored to the operators needs when added to the UMTS infrastructure.

6.3 Wi-Fi

Much like the WiMAX Forum the Wi-Fi Alliance[8] is a non-profit organization of wireless manufacturers and service providers, working for certification and use of WLAN equipment.

The Wi-Fi technology is a network architecture using the radio access technologies of IEEE Stds. 802.11a, 802.11b and 802.11g. Analysts of the Wi-Fi market predicts a compounded annual growth of revenues at a rate of 44-66% by 2008[57][17], and already the WLAN technology has become market pervasive even down to private consumers.

6.3.1 Features

Wi-Fi extends 802.11 with different suits of security, speeds and topology configurations. A selection of IEEE add ons for the original 802.11 standard is shown in table 6.4[33].

802.11a	5GHz OFDM 54Mbps
802.11b	2.4GHz DSSS[3] 11Mbps
802.11e	QoS features
802.11f	Inter Access Point Protocol, mobility support
802.11g	High rate extension for 2.4GHz band to 54Mbps
802.11i	MAC security extension, AES

Table 6.4: 802.11 Flavors

6.3.2 Network Architecture

The simplicity of the network architecture of 802.11 WLAN is a strength as well as a weakness. The lack of core network management simplifies installation and integration, but limits network control, service and user management. An example of a WLAN network configuration[59, p. 208] is shown in figure 6.4 on the next page.

Since the WLAN operates much like a regular LAN, the provider is free to deploy own DHCP or DNS servers, or relay such services from a second party. The small scale of WLAN networks makes it appropriate for small business types access networks, where users remain stationary for a while and later move on, rather than supporting constant mobility through subnets.

6.3.3 Wi-Fi and WiMAX

All though Wi-Fi specifications have grown to include Extensible Authentication Protocol (EAP), Temporal Key Integrity Protocol (TKIP) and 802.11i functionality, the

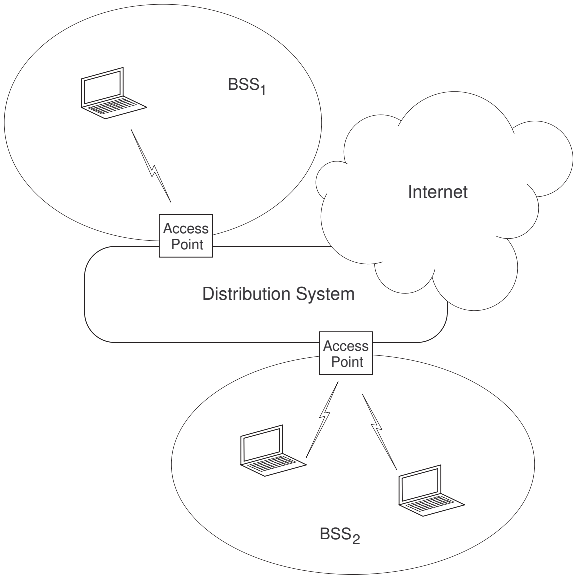


Figure 6.4: A Common WLAN Configuration

standard has been criticized for its security flaws. The original Wired Equivalent Privacy (WEP) has several weaknesses such as the relatively short 24-bit initialization vector (IV) and static keys.

Wi-Fi is trying to plug these holes but fact remains that wireless IP-technology does not have the heads up on security as the cellular industry when it comes to mobility. This might limit the number of services that can be transfered and protected when migrating from WiMAX to WLAN.

The radio technology itself has some inherent weaknesses, such as the 2.4 GHz band sensitivity to interference from kitchen appliances and other electronic devices. Although there are amendments for QoS support and traffic on other frequency bands, there are limits to how far the 802.11 standard can be ”patched”.

Transport of bulk data and ”cheaper” access through Wi-Fi hotspots seem to be the interesting aspects of WLAN/WiMAX mobility. It is this type of integration with Wi-Fi hotspots that is in focus, where a WiMAX network serves as a backhaul for WLAN access points.

802.11f, supplying inter access point mobility is a step towards WLAN roaming, but the scale of WLAN networks makes fast mobility impractical as handovers would be performed very rapidly due to the relatively small cell area. There has been proprietary solutions for handling roaming between access points in a 802.11 network, but the 802.11f standard will allow roaming interoperability between access points of different manufacturers.

Faster mobility is something that WiMAX, with its MAN coverage, might be able to handle better with the 802.16e standard for roaming. The performance of this WiMAX flavor remains to be seen in practice, but tests have already begun[7].

The need for handovers between the different systems come when mobile devices change IP-subnets. This can be managed at a higher level in a WiMAX network where layer 3 mobility[4] is not performed unless the mobile device moves between ASNs. In 802.11 this type of L3 handover is performed when a device moves between access points on different subnets. MIP is extensively used to handle this type of mobility.

QoS in WLAN has been limited to the selection of either the Point Coordination Function (PCF) or Distributed Coordination Function (DCF), limiting the QoS capabilities, but with 802.11e two new functions for QoS is introduced[22]. These two functions can offer eight user priorities, but cover only one connection. WiMAX on the other hand can apply its QoS parameters to services, users, traffic types and cover several connections with traffic differentiation on each connection[50].

The ability to differentiate both users and traffic speaks for the integration of WLAN as an hotspot extension to WiMAX access networks, since the traffic backhauled from the WLAN access point can be served with respect to the attached users.

As in the case of UMTS and WiMAX in section 6.2.3 on page 38, mobility across Wi-Fi and WiMAX networks can be integrated according to two models. Since both networks are IP-based, integration does not require translations in protocols when bridging the infrastructure.

Loose Coupling

The lack of network management features in WLAN requires additional components by the WLAN operator to obtain a higher degree of network intelligence when performing vertical handovers. Network intelligence can be used to facilitate context transfers of AAA and services prior to handovers. This will reduce the latency of the handover and make the transition seamless[74].

As in the case of UMTS, loose coupling enables flexibility for operators as roaming agreements can be made to extend WiMAX data service coverage without expanding or integrating WLAN in a WiMAX network.

Crucial to global roaming between Wi-Fi hotspots and WiMAX networks is better service tracking in WLAN where accounting and billing is supported[31] as WLAN currently lacks such infrastructure.

Tight Coupling

In this type of integration, components of the WiMAX network are already present to support user mobility and AAA, and, similar to UMTS, integration proxies can be inserted into the core network to support seamless connection transfers while avoiding dual traffic flows. Care needs to be taken with such a solution so end-to-end semantics are preserved[59, p. 357]. Proxies also add a point of failure which requires some redundancy, but managing mobility at a higher level in the core infrastructure could mean better scalability.

Location management would also be made easier with the knowledge of network structure when moving between the networks. With location management mobile devices can refrain from searching for other network technologies on other interfaces, thus saving power.

A WLAN attached to a WiMAX network could use the CSN to supply users with IP-connectivity and IP-address allocation.

[4] Also termed Macro Mobility.

6.4 Related Work

With 4G networks on the horizon, integration and mobility between heterogeneous networks has spawned many research projects, such as Release 6 of the UMTS standard, including inter-networking with WLAN.

Some of these project address the issues of Layer 3 handovers. This is caused by the duality of the IP address, where an IP-address both functions as an indicator for identity and point of attachment.

6.4.1 802.21

IEEE is currently working on the 802.21 standard for Media Independent Handover (MIH), a cross layer standard for managing handovers across different IEEE 802 and non-802 technologies[5] as a form of middleware. This is shown in figure 6.5.

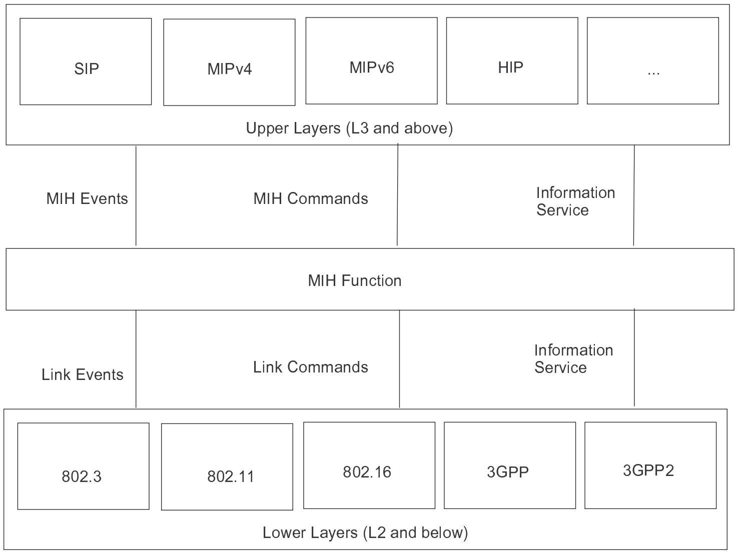


Figure 6.5: IEEE Std. 802.21 Layering

802.21 contains three main components: an architecture for protocols between network elements, handover functions with technology independent Media Independent Handovers[5] (SAP) and primitives, MAC-layer SAPs and primitives for specific access technologies.

The technology specific additions is proposed to be ammendments to the specification of each standard. MIH would be placed like a middleware in network components managing mobility and the mobile device, supplying functions to higher layers.

MIH extends MIP in that MIH can process information from Layer 2 regarding the network status and utilize multiple technological interfaces to do so.

The standard utilizes triggers on multiple layers and proposes a Layer 2.5 with optimized handovers, soft and hard, for mobility in heterogeneous networks[9].

[5] The point in the protocol stack where the services of a lower layer are available to its next higher layer.

The latest draft of the standard was published in July 2005 and time of completion is set to May 2006, but this is an estimation left to be confirmed. Current similar proprietary solutions exist in e.g NetMotion Mobility XE[70] and PCTEL[26], both supplying the means for heterogeneous mobility.

6.4.2 Host Identity Protocol

IETF is working on the Host Identity Protocol (HIP)[35], dealing with the security issues of IP traffic and in the process addressing the issues of the IP packet routing. Since an IP address is used to identify a host connected to an IP-network as well as the topological attachment, mobility is severely hindered but can be mitigated using MIP[59, p. 304].

This however requires multiple IP addresses, redirection of flows and signaling protocols on top of the IP layer.

HIP allows host identities to be independent of packet flows, thus making it possible to route packet via host unique identity rather than IP address. It is placed as a layer within the IP-layer, between the routing and end-to-end functions, enabling separation of the location and identification characteristics of IP-addresses. This allows users to be identified by a Host Identifier (HI), introducing a new address space.

The HI is represented by a 128-bit long identifier, called the Host Identity Tag (HIT). This tag is generated by applying a cryptographic hash function over the HI.

The design allows it to function with the addition of the rendezvous server and changes in the DNS directories.

HIP facilitates peer-to-peer security and mobility, and through the elimination of IP-address duality it can greatly facilitate L3 mobility across subnets and also heterogeneous networks. Jokela and colleagues[36] describes mobility and multi-homing/multi-addressing for HIP over IPv4 and IPv6. Liang and colleagues[43] discusses a comparison of MIP, Migrate and HIP for host mobility.

HIP does however not support multicast and protection against flooding Denial of Service (DOS) attacks. Nikander and others[52] explores the possibility of combining HIP with Secure-i^3 and contains a description of both protocols.

6.4.3 Cellular Assisted Heterogeneous Networking

Cellular Assisted Heterogeneous Networking(CAHN)[13] is a project working to utilize a signaling plane separate from actual data connections in Ad hoc and peer-to-peer networks. It proposes to use the global, always on, characteristics of cellular networks such as GSM/GPRS.

Cellular networks already contain functions for authentication, identification and low bandwidth signaling. Heterogeneous networks tend to be very different in such characteristics as AAA, power consumption, coverage and bit rates.

CAHN makes use of the Mobile Subscriber Integrated Services Digital Network Number (MSISDN) to identify users rather than IP, MAC or other addressing schemes. Users already have security mechanisms and billing arrangements with their cellular operator, which could be reused when setting up connections between users.

This separates users from the pervasive requirement for topology correctness of many current addressing schemes that typically also describe a point of attachment along with identity.

Experiments have been performed with SMS as signaling between two hosts to request a connection set up, but SMS as control bearer proved to be too delay sensitive. The Fast Associated Control Channel (FACCH) was therefore suggested, instead of the Slow Associated Control Channel (SACCH) used by SMS, to promote CAHN protocol signaling[45].

Users send each other session join requests via the CAHN protocol, utilizing cellular authentication and identifying each other via the MSISDN. When a join request for a session arrives the devices can scan for available technologies such as WLAN, Bluetooth etc and setup a connection through a desired medium.

This has the advantage that signaling is separate on a low power, always connected, interface and high bandwidth interfaces which consume power are only activated when used as a bearer service. These high bandwidth interfaces do not need to continously activate or stay connected as connection signaling is separate, waking the appropriate interface only for actual data transfer.

This does how ever require that the user have a cell phone with service agreement to a cellular provider and that there is an infrastructure, separate from the heterogeneous access technologies, supporting the CAHN architecture.

6.5 Conclusions

The WiMAX network reference architecture has a stated support for mobility between heterogeneous networks, but precisely how this will function is somewhat unclear as full mobility is not expected until 2007-2008 and the possible tell-tales of the 802.16e roaming standard is not yet finalized.

Despite this, the standard can be evaluated at a higher level, considering the internetworking structure and the technologies at hand.

The general IP-convergence of both cellular and computer networks is a factor considered by WiMAX. MIP is the de facto standard for L3 mobility and is supported in the network reference architecture. The goal for mobility within NAPs is to specifically facilitate L3 handover protocols.

Handling these types of macro mobility handovers with L3 mobility protocols like 802.21 and MIP can bring 802.11 to support the same mobility level as WiMAX and UMTS. This would give 802.11 the infrastructure to support access points rather than using access points to support the infrastructure. This is a difference from cellular networks and WiMAX where the base stations are a central piece.

Besides L3 mobility there is also options of L2 and cross layer mobility support to consider. MIH is one standard utilizing cross layer information.

The WiMAX architecture also considers the transition from IPv4 to IPv6, supporting MIP for both formats. There is however inherent weaknesses in MIP when it comes to mobility characterized with frequent vertical and horizontal handovers[64] causing L3 handovers, but the hierarchical structure of the WiMAX network supports moving this type of mobility to a higher level by using internal Data Paths for routing within the access network.

UMTS is moving towards the IP-traffic characteristics of the Internet while WLAN is moving towards better QoS, mobility and cellular characteristics, with all these characteristics shared by WiMAX.

Even though WiMAX originally was designed for fixed wireless broadband, the OFDMA physical layer with its NLOS and data rate capabilities, MAC-layer with security and QoS, and network architecture with management entities show potential for mobility support. It does however remain to practically test and evaluate the actual capabilities of WiMAX in a live environment to validate the hype surrounding this new architecture.

How this integration is made between the different technologies will affect the mobility performance, and much of the research done today mainly covers horizontal handovers, with signal power as the measured characteristics as base for handover decision.

More research needs to be made where policies defining metrics such as QoS and cost is a key factor when switching networks. Simple signal strength measurements can not capture the diversity and application of the different access technologies offered by loosely coupled networks served by a multitude of service providers.

Chapter 7

Simulation Study

This chapter describes simulation scenarios of the WiMAX MAC-layer, the environment they were conducted in and an analysis of the results obtained. The three main scenarios included are handover, fast base station switching and network entry. The process of implementing the test scenarios described below will also help with evaluating the simulation environment. This evaluation is done in chapter 8 on page 57.

7.1 Environment

The scenarios are implemented in GloMoSim, a software supporting scalable simulations of numerous types of wireless protocols. It is built on PARSEC[39], a parallel discrete event simulator.

GloMoSim uses a layered structure, much like the Open Systems Interconnection (OSI)[60] seven layer network stack. Table 7.1 shows some of the models available for the major layers in GloMoSim.

Layer	*Model*
Propagation Model	Two-Ray, Free Space
Data Link	802.11, CSMA, MACA, TSMA
Network(Routing)	Static Routing Table, OSPF, AODV
Transport	TCP, UDP
Application	Telnet, FTP, CBR

Table 7.1: GloMoSim Protocol Layers

These models can be mixed independently of each other. Besides selecting model for each layer, there is also several general or model specific options to configure. The output is presented in a statistics file after the simulation has completed.

For more information, see the GloMoSim tutorial[53] or the more technical description of how GloMoSim works, by Zeng et al.[75].

7.1.1 Network Topology

The network topology used for the simulations consists of:

- MS - a mobile station standing at a fixed location, in range of both base stations BS1 and BS2.
- BS1 - base station in range of MS, but out of range of BS2.
- BS2 - base station in range of MS, but out of range of BS1.
- Wired Recipient - connected by a wired network to both BS1 and BS2.

Figure 7.1 shows the placement of the nodes serving as base stations, mobile station and recipient nodes.

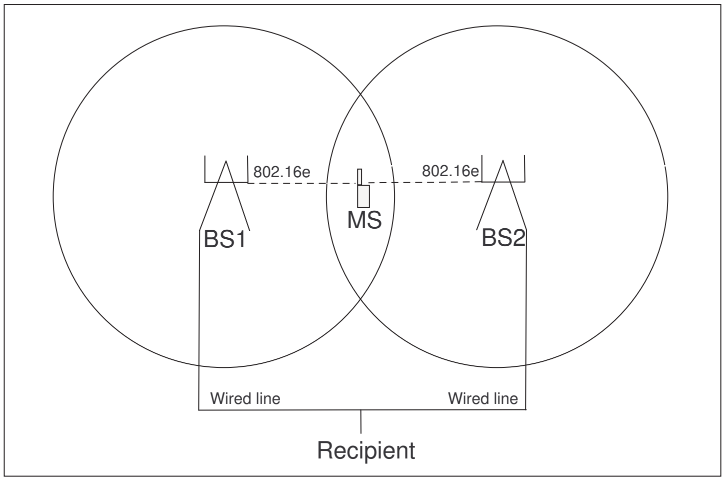


Figure 7.1: Network Topology

The base stations BS1 and BS2 use the implemented features of 802.16e for managing the air interface towards the MS. The base stations use a default GloMoSim wired line to communicate with the wired recipient. The wired recipient can not communicate directly with the MS or vice verse.

7.1.2 Simulation Parameters

As there is no protocol model for any version of 802.16 for GloMoSim, the simulation configuration uses the PHY-layer of 802.11. The 802.16e features implemented for this thesis was inserted at the MAC-layer. At the top layer there is a constant bit generator which can send packets addressed to application layers of other nodes. These packets are sent with a known size, interval and duration through the protocol stacks.

Table 7.1.2 on the next page contains the selected layer models used in the simulation.

The backoff contention windows is set to 10 transmission opportunities. This means that a MS will defer between 0-10 transmissions to send initial ranging messages to the BS. Each initial ranging allocation contains 3 transmission opportunities.

Layer	*Model*
Propagation Model	Two-Ray
Physical	802.11
Data Link(MAC)	802.16e
Network(Routing)	Static Routing Table
Application/Transport	Constant Bit Generator

Table 7.2: Selected GloMoSim Layers

7.2 Scenarios

The three scenarios implemented are; initial network entry, hard handover and Fast Base Station Switching (FBSS). A more detailed description of FBSS and soft handover[1] can be found in [25].

The scenarios were developed for two purposes. The first was to facilitate measurements of WiMAX processes like network entry and handover. The second was to facilitate the evaluation of GloMoSim as environment for future development as a WiMAX simulator.

The measurements done in these scenarios are in no way intended for use as performance evaluation of WiMAX and its air interface, but rather as a general MAC-layer performance comparison and analysis of the procedures used within WiMAX and their differences.

7.2.1 Network Entry

This scenario serves as a pilot study of coding a MAC-layer in GloMoSim and is also required by a node entering a WiMAX network. During this phase the MS performs ranging and negotiates capabilities with a BS. The end result is a setup of management and transport connections between the BS and MS.

The network entry phase can be seen in figure A.1 on page 73 according to the 802.16-2004 standard[2, p. 167].

The scenario starts with the MS operating on the same frequency as BS1. At a random time[2] after the simulation start the MS initiates the network entry sequence until it has one transport channel to BS1. As network entry is complete, the MAC-layer in the MS will relay data packets from the application layer to BS1. Data packets sent before network entry is complete will be lost.

The scenario makes the following assumptions.

- Initial Ranging is always successful. To avoid cross layering the 802.11 PHY into the 802.16 MAC, there is no actual measurement of signal quality.
- No computational overhead delays the response messages. This is the largest contributor to the idealization of measured data. The BS will need to interact with other base stations, gateways or upper layer protocols in practice, but such a architecture is not available for this simulation.

[1] Soft Handover is not covered in this thesis

[2] This random time is within an interval such that all periodic broadcast messages get transmitted at least once.

– MS does not need to perform authentication, key exchange, DHCP negotiation or time synchronization. Such protocols are currently unavailable

Initial analysis hinted that the selected frame rate would be the deciding factory for how fast network entry can be completed, but this is not the case. Since the MS needs to obtain the UCD[3] message in order to send on the uplink, the interval in which this message is sent is crucial for the network entry time[4].

The 802.16-2004 specification draft[2, p. 637] has this interval set at a maximum of 10 seconds, but does not recommend a minimum or default value. Worst case scenario would then mean that network entry can take slightly longer than 10s, as the MS needs to wait the full interval and then continue with the process.

The backoff performed during initial, contention based, ranging also affects the network entry time but this impact is, in order of millisecond, a lot smaller compared to the length of the UCD interval which is in order of seconds.

For this scenario the time for network entry is measured with varying frame rates. Frame duration is varied according to table 7.3. These frame durations are from the OFDM PHY specification of the 802.16-2004 standard[2, p. 460].

Frame Duration(ms)	*Frames per second*
2.5	400
4	250
5	200
8	125
10	100
12.5	80
20	50

Table 7.3: Tested Frame Rates

Results

Figure 7.2 on the facing page show how the network entry time varies with the frame rate for a UCD interval of 2s.

Even though network entry delays could be neglected in certain scenarios when first entering a WiMAX network, it can seriously impact the time it takes to perform handover. This can be seen in the next scenario.

7.2.2 Hard Handover

The WiMAX mandatory handover case requires communication with both BSs and thus a frequency change for the MS[5]. Before and after the handover is completed the data traffic shall continue uninterrupted, with the flow of data now moved from one BS to another after a completed handover.

This type of handover is basically performed by releasing the MS from the associated BS and then performing network entry at another BS. Figure A.2 on page 74 shows

[3] This message contains parameters for decoding the uplink.

[4] Provided the MS has no prior knowledge of up/downlink parameters.

[5] Since the 802.11 PHY-layer is used, the only way to differentiate individual base stations is to listen to their respective frequencies.

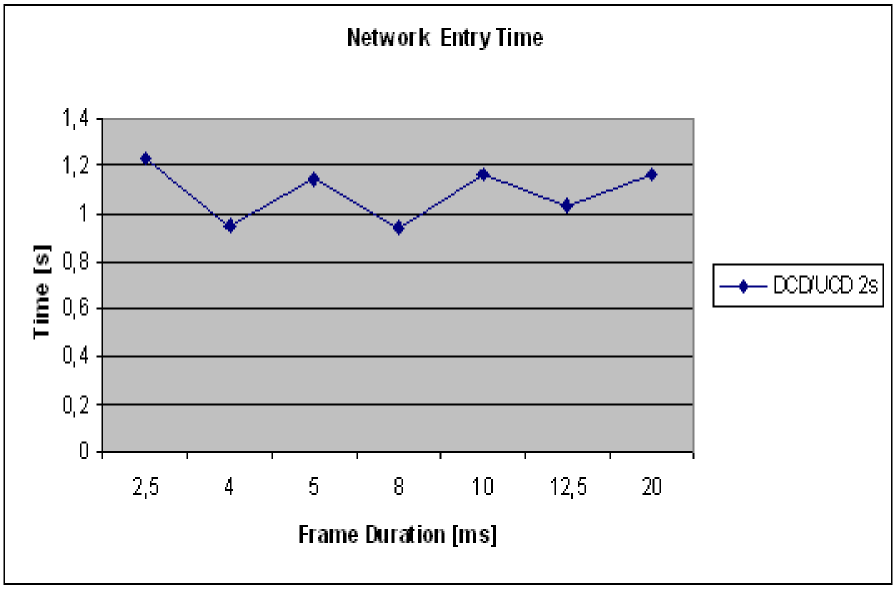


Figure 7.2: Measured Network Entry Times

the transition stages of a handover according to the 802.16e standard[3, p. 176]. The handover used in this case is MS initiated at a random interval between two broadcasted UCD messages[6]. Since the network entry procedure from the first scenario is used for hard handover, the same assumptions still hold.

It is also assumed that the target BS always has resources available for accommodating the MS connections.

Measured in this scenario is the time it takes to perform a handover, from the decision to make the handover to the time when application data is received at the second base station. It is also of interest to pin-point the parameters that govern how fast a handover can be completed.

This scenario will vary frame rates and measure the different network entry times for a fixed UCD interval. Another measurement will for a selected frame rate vary the UCD interval.

This case will also look at how throughput of application data is affected by the handover and its associated overhead.

Results

As this case is similar to network entry, with addition to exchange of handover request, response and indication messages, the resulting time measured will be affected by the same parameters, i.e. the interval of periodic messages needed for the process to complete.

Figure 7.3 on the following page depicts how the handover time is affected by varying frame rates for an UCD interval of 2 seconds.

[6] This broadcast message defines, like the network entry case, when the MS can reach uplink synchronization.

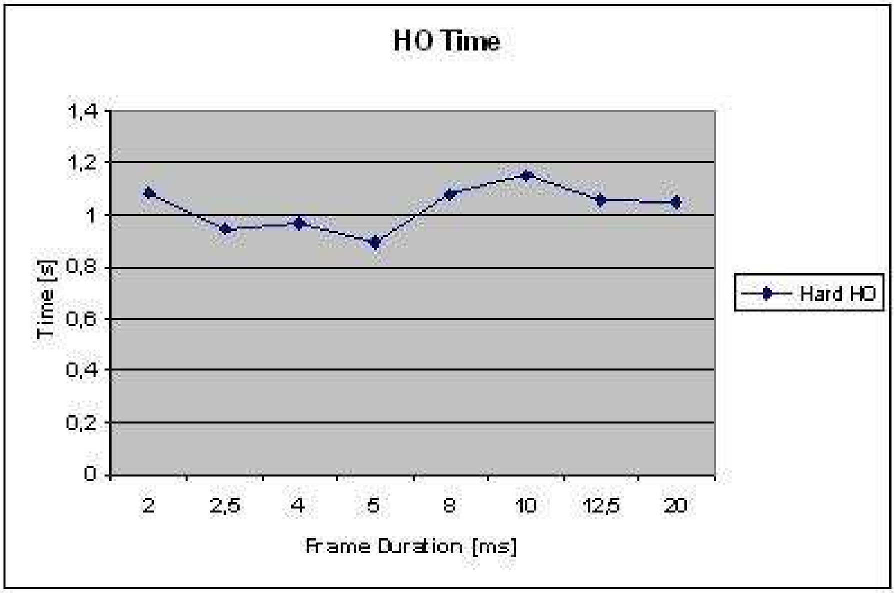


Figure 7.3: Measured Handover Times (Varying Frame Rate)

An increase in frame rate does not speed up the handover procedure through more frequent request/response opportunities, as this procedure is dependent on broadcast messages that use a fixed time interval independently of the frame rate.

If the information needed for uplink synchronization is sent via management messages or through periodic neighbor scanning, the handover could be completed faster. The MOB_NBR-ADV[3, p. 171] management message has the purpose of distributing this information and is broadcasted at regular intervals by a BS, sharing channel information of neighbor base stations without the need for a MS to gather this information from each individual adjacent BSs. This does however introduce more management overhead[24, p. 54].

The impact on handover time by the UCD interval is shown in figure 7.4 on the next page. Frame rate is constant at 5 ms.

If the decision to make a handover has a uniform distribution over the time interval, the results are as expected with the average handover time being almost half the UCD-interval.

Figure 7.5 on page 56 shows the throughput at BS1 and BS2 during the simulation. Throughput is reduced when the traffic of management messages reduced the portion of uplink available for application data. As the MS de-registers with BS1, the throughput for BS1 drops to zero. Maximum throughput at BS2 is reached after network entry is complete.

7.2.3 Fast Base Station Switching

FBSS is used for diversity combining[67] of signals from several base stations. When measured signal characteristics fall beneath a threshold value, the MS initiates an anchor change of BS. The MS can immediately switch between base stations on a frame by frame

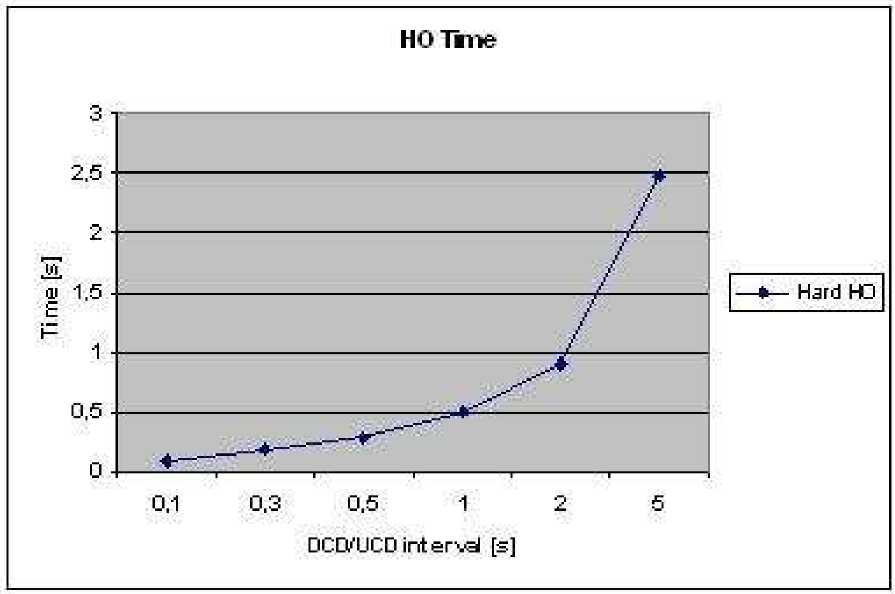


Figure 7.4: Measured Handover Times (Varying UCD)

basis to quickly adapt to deteriorating channel environments. This is made possible by distributing the information needed for such a switch before hand.

A message sequence diagram of the FBSS phase can be seen in figure A.3 on page 74 according to the 802.16e standard[3, p. 186].

Results

Figure 7.6 on the following page shows the throughput at BS1 and BS2 during the simulation. Throughput is reduced when the traffic of management messages reduced the portion of uplink available for application data. These management messages consummate the anchor update and when this is procedure is completed the MS can send on the next frame of BS2.

The throughput is never reduced to zero and the MS does not need to stop transmitting application layer data.

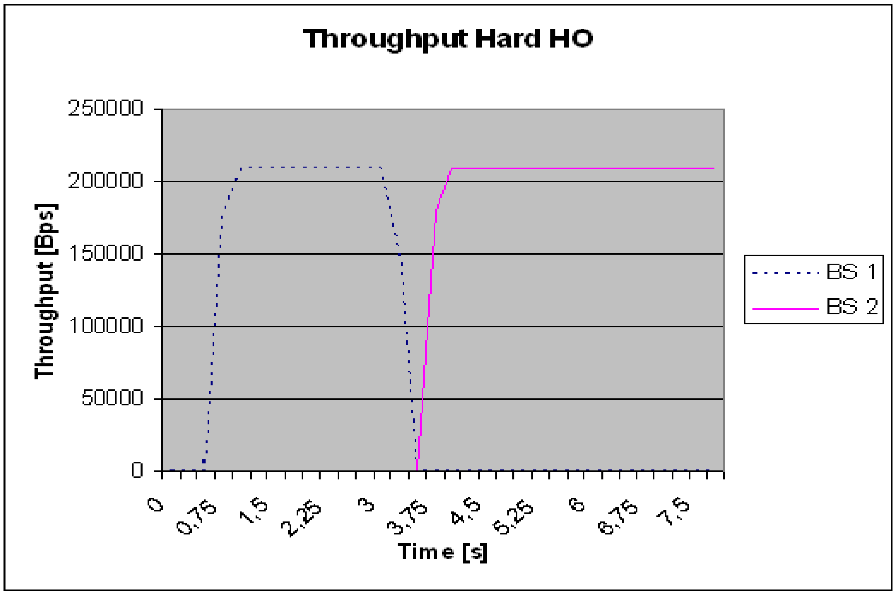


Figure 7.5: Handover: Throughput of BS1 and BS2

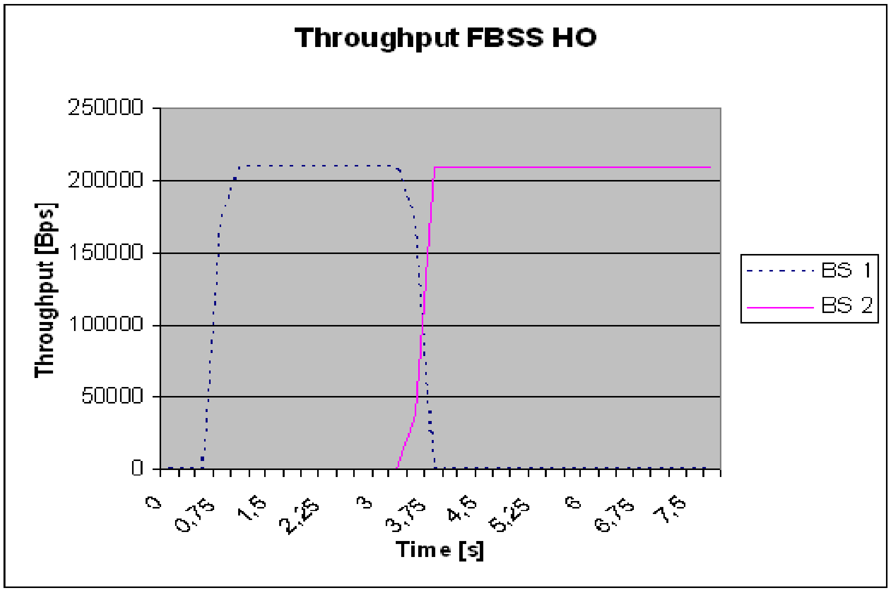


Figure 7.6: FBSS: Throughput of BS1 and BS2

Chapter 8

Conclusion

The WiMAX technology and architecture form a complex, but feature-rich environment for supplying end user mobility. It shares many characteristics of cellular networks, such as architecture support of billing, mobility, and QoS, but also scales down with the technology being used for bridging other networks.

The future of WiMAX seems to be something similar to figure 8.1. WiMAX working in conjunction with other access technologies, each complementing each other.

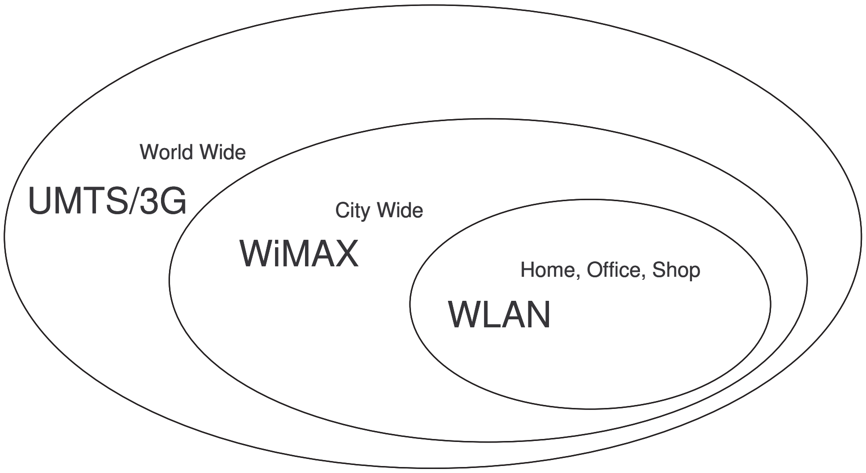


Figure 8.1: Heterogeneous Technologies Enveloping Each Other

The practical capabilities does however remain to be evaluated and in the end the question remains if WiMAX can generate enough revenues to motivate deployment and on what scale this deployment will be.

Table 8 on the following page contains the problems defined in the Problem Statement, Chapter 2. Each has a reference to the chapter or section where it is discussed.

This thesis has touched varying depths of WiMAX and the mobility aspects. Even though it is not exhaustive in its contents, the authors hope it still will serve a purpose for those interested in WiMAX and its future development.

Problem	*Reference*
Mobility capabilities of WiMAX in homogeneous networks	Chapter 5
Mobility capabilities of WiMAX in heterogeneous networks	Chapter 6
GloMoSim evaluation	Chapter 8.1
Efficiency of mobility support	Chapter 7 Chapter 8.3

Table 8.1: References to the Problem Statement

8.1 GloMoSim

GloMoSim has strength in its simplicity and the fact that measurements can be completely tailored, limited only by the complexity of the implementation. This freedom also makes the work challenging when implementing a very complex layer in the protocol stack.

The layer model of GloMoSim also makes it easy to insert new protocols and the interface between the protocols is very plain. If needed these interfaces can be bypassed and cross layer information made available.

When working with a dynamic system like WiMAX the output format is somewhat crude. A single statistics file at the end of the simulation is enough when performing e.g additive measurements of messages passed or total average values, but does not support interaction with nodes during runtime unless they are predefined in the protocol.

For this implementation some MS behavior is even defined in the source code, e.g when to perform network entry or handover, which really limits the use to someone with knowledge of the protocol implementation.

A more advanced version of the mobility feature would be suitable to control the behavior of mobile stations over time. This feature lets GloMoSim read an input file containing movement coordinates for a node. Output adapted to runtime events could be useful. A suggestion could be to define a number of node actions in an external file, possibly with timing instructions, such as a MS powering up at a certain time interval.

8.2 Implementation

The original plan was to take parts of the already available 802.11 protocol for GloMoSim and customize it to suit 802.16e. However few pieces of the 802.11 protocol proved useful for code reuse, but it was very useful for studying use of internal GloMoSim methods and data structures. As the available documentation of source code and design was limited, trial and error formed the corner stone for actual implementation. In the work of prototyping a complex technical standard, the challenge has been to identify both crucial functionality and scenarios interesting to perform measurements on.

8.2.1 Future Work

The section on future work covers much of the limitations of the current implementation. This implementation only covers specific parts of the 802.16e standard and much remains to be designed and built. Besides the MAC-layer, ASN Gateways and other core network components also needs to be defined for a complete WiMAX architecture, probably spanning many different layers of the protocol stack.

Scalability

Currently there is only support for one mobile station which is limited to five connections. Traffic is classified on basis of the message destination address and should be moved to a more versatile classification unit.

Base stations needs to be evolved to handle multiple MS through admission control.

Scheduling

QoS and scalability are fundamental parts of WiMAX and the MAC-layer, but has not been covered in this thesis. All nodes have a fixed, hard coded bandwidth and connection configuration. Such settings should be manageable through some type of interface, such as the config.in file already available for configuring GloMoSim.

No management of the QoS classes defined in the 802.16e specification has been implemented and an entity managing these needs to be constructed. All traffic is scheduled according to best effort. Section 8.4 on the next page contain a reference to another thesis work on QoS in 802.16.

There is also no support for the procedures of fragmentation and aggregation of packets over the air interface. They are stated as mandatory by the 802.16-2004 specification.

Ranging

Timing adjustments and monitoring of radio signal is crucial for implementing the ranging and scanning phases. As there are several methods for scanning base stations, this can introduce several interesting scenarios for simulation. This functionality would also make the scenarios studied in this thesis more complete.

The current implementation does not support actual node mobility since there are no timing measurements that can synchronize the MS with the BS in accordance with propagation delays. This is needed as the BS is controlling the timing of all nodes in its cell area and nodes on different distances from the BS will suffer deviation from burst scheduling as intercepted transmission times vary with the distance.

8.3 Simulation Results

Even though the designed prototype is lacking in some fundamental areas, there are some detectable characteristics derived from the simulated scenarios and initial analysis.

A system supporting FBSS will require bigger overhead than support of only hard handover to the traffic of management messages if it is to manage FBSS clusters of base stations. It does however gain from the advantages of diversity combining. FBSS basically performs L2 combining, while the soft handover suggested in the 802.16e specification works as a L1 combining[25]. This overhead is however spread over time and not introduced when handover is performed as in the case of regular, hard handover.

It is hard to compare, what the draft calls different handover procedures or modes, as they are performed on different levels. Hard handover, FBSS and soft handover are similar but still very different. They have different application areas and performance is thus hard to compare. It is however possible to evaluate what affects the different methods.

If the information needed for association with a new base station is sent during MS association with a BS and the FBSS cluster does not change too frequently, gains in transmission quality for a dynamic radio environment can be made. Support of the mandatory hard handover is however much less complex, as support of FBSS adds requirements to the MS and BS, such as frame synchronization[25, p. 2].

The convergence time for handovers also rely on the MS performing regular ranging with neighboring base stations continously during periodic handover scanning[3, p. 588].

The scanning and ranging processes were not considered in the scenarios due to a limited time plan, but these events would make the measurements done in the scenarios more complete. This is due to handovers affecting the system not only when it is being performed, but also when it is being maintained or prepared through events spread over a longer time. A comparison of the scanning methods[3, p. 170] together with an study of overhead in the context of handover decisions would be the next natural step in the development.

The study is considered inconclusive as there are more parameters, than those covered here, that affect the handover performance of the 802.16e MAC-layer.

8.4 Related Work

There are many other projects related to investigating the capabilities of WiMAX and its air interface. Listed here is a summary and references to some of the more extensive WiMAX related software developments.

8.4.1 QoS Scheduling

In the thesis *An Efficient QoS Scheduling Architecture for IEEE 802.16 Wireless MANs*[46] the authors presents an architecture for supplying guaranteed bandwidth and delays for the various QoS flows in 802.16.

The architecture uses Grant Per Subscriber Station(GPSS) for the BS scheduling bandwidth to stations. The stations then allocate this bandwidth to their respective service flows, based on priority and bandwidth requirements.

The scheduling of service flows in each subscriber station uses a combination of strict priority scheduling and Weighted Fair Queuing(WFQ) to obtain the guaranteed delays and bandwidth. The algorithm for uplink scheduling is presented in the thesis. They proceed to implement and analyze the architecture in the QualNet 3.6 network simulator[62].

The thesis has an easy to follow disposition and also explains key components and functionality of the 802.16 MAC-layer that is of importance for supporting an QoS architecture. The analysis contains various measurements on traffic loads on the flows, bandwidth utilization and delays as a proof of concept for several different scenarios where the load on the different service flows are varied.

As this is a QoS architecture, and a QoS architecture only, there is no mentioning of other MAC-layer functionality or procedures.

8.4.2 Radiowave Propagation Simulator

Radioplan[55], a German RF simulation and optimizations specialist, offers their Radiowave Propagation Simulator(RPS) free to students.

It contains tools to simulate small indoor and outdoor radio environments for technologies such as WLAN, WiMAX and UMTS. The simulator enables testing of different physical characteristics of these technologies, such as penetration and reflection.

It supports plug-ins for custom models, graphic surface plots and import of environment data from different Computer-Aided Design (CAD) formats.

The free software support for 750 graphic environment polygons, two reflections, two penetrations and one diffraction for each ray. These restrictions can be removed by upgrading to the Professional or Enterprise edition of RPS.

Figure 8.2 shows an example of the interface displaying a 3D-view of outdoor radio coverage.

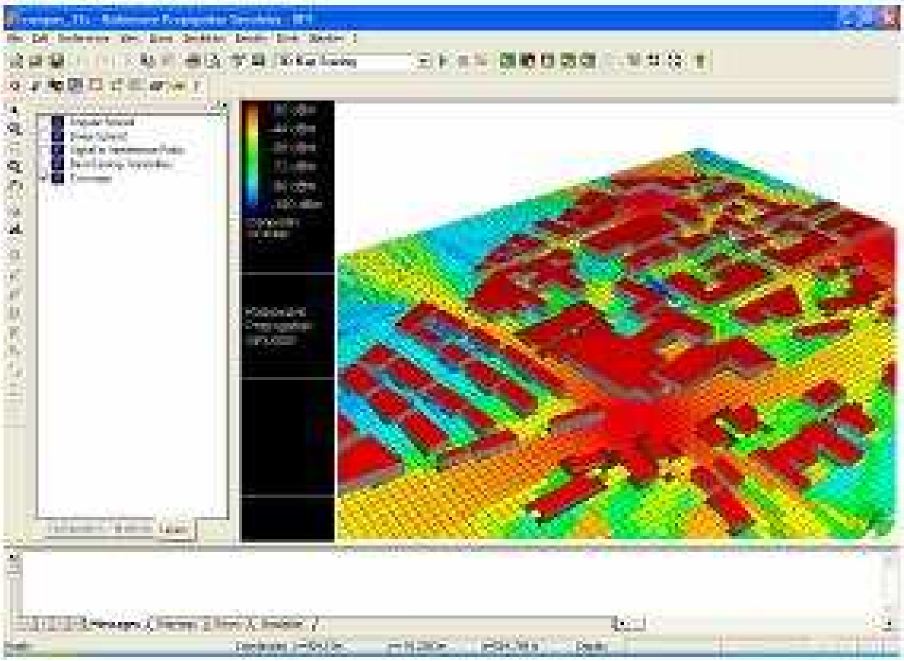


Figure 8.2: RPS Outdoor Coverage Distribution[14, p. 179]

The simulator can be downloaded at their web page[56] free of charge for educational or personal use. The accompanying user manual seems detailed with many screen shots and examples for easy learning of the software.

Chapter 9

Abbreviations and Acronyms

3G	Third Generation Network
3GPP	3rd Generation Partnership Project
AAA	Authentication, Authorization and Accounting
AES	Advanced Encryption Standard
AODV	Ad hoc on Demand Distance Vector routing
ARQ	Automatic Repeat Request
ASN	Access Service Network
ASN GW	ASN Gateway
ATM	Asynchronous Transfer Mode
BE	Best Effort
BS	Base Station
BSC	Base Station Controller
BFWA	Broadband Fixed Wireless Access
CAD	Computer-Aided Design
CAHN	Cellular Assisted Heterogeneous Networking
CBR	Constant Bit Rate
CDMA	Code Division Multiple Access
CN	Core Network
CoA	Care of Address
CRC	Cyclic Redundancy Check
CSMA	Carrier Sense Multiple Access
CSN	Connectivity Service Network
DCD	Downlink Channel Descriptor
DCF	Distributed Coordination Function
DES3	Tripple Data Encryption Standard
DHCP	Dynamic Host Configuration Protocol
DIUC	Downlink Interval Usage Code
DL	Downlink

DNS	Domain Name System
DOS	Denial of Service
DSA	Dynamic Service Addition
DSC	Dynamic Service Change
DSL	Digital Subscriber Line
DSSS	Direct Sequence Spread Spectrum

EAP	Extensible Authentication Protocol
ertPS	extended real time Polling Service
ETSI	European Telecommunications Standards Institute

FACCH	Fast Associated Control Channel
FBSS	Fast Base Station Switching
FDD	Frequency Division Duplex
FEC	Forward Error Correction
FFT	Fast Fourier Transform
FTP	File Transfer Protocol
FWA	Fixed Wireless Access

GloMoSim	Global Mobile Information System Simulator
GPRS	General Packet Radio Service
GPSS	Grant Per Subscriber Station
GSA	Global Mobile Suppliers Association
GSM	Global System for Mobile Communications

HI	Host Identifier
HIP	Host Identity Protocol
HIT	Host Identity Tag
HSDPA	High Speed Downlink Packet Access
HSUPA	High Speed Uplink Packet Access

IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IMEI	International Mobile Equipment Identity
IMS	IP-Multimedia Services
IP	Internet Protocol
IPsec	IP Security
ISP	Internet Service Provider

L2	Layer 2, MAC Layer
L3	Layer 3, Network Layer
LAN	Local Area Network
LOS	Line of Sight
LPM	Last Packet Marketing

MAC	Medium Access Control
MACA	Multiple Access with Collision Avoidance
MAN	Metropolitan Area Network
MBMS	Multimedia Broadcast Multicast Service
MIH	Media Independent Handover
MIP	Mobile IP
MPEG	Moving Picture Experts Group
MS	Mobile Station
MSC	Mobile Services Switching Center
MSISDN	Mobile Subscriber Integrated Services Digital Network Number
MSS	Mobile SS
NAP	Network Access Provider
NGN	Next Generation Network
NLOS	Non Line of Sight
nrtPS	Non Real Time Polling Services
NSP	Network Service Provider
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
OSI	Open Systems Interconnection
OSPF	Open Shortest Path First
PCF	Point Coordination Function
PDG	Packet Data Gateway
PDU	Packet Data Unit
PHS	Package Header Suppression
PHSF	PHS Field
PHSI	PHS Index
PHY	Physical
PMP	Point to Multipoint
PS	Physical Slots
PSTN	Public Switched Telephone Network
PTP	Point to Point
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
RNC	Radio Network Controller
RPS	Radiowave Propagation Simulator
RRM	Radio Resource Management
RTG	Receive/transmit Transition Gap
rtPS	Real Time Polling Service

SACCH	Slow Associated Control Channel
SAP	Service Access Point
SDU	Service Data Unit
SGSN	Serving GPRS Support Node
SMS	Short Message Service
SNR	Signal to Noise Ratio
SS	Subscriber Station
TCP	Transmission Control Protocol
TDD	Time Division Duplex
TFTP	Trivial File Transfer Protocol
TKIP	Temporal Key Integrity Protocol
TLV	Type/Length/Value
TSMA	Time Spread Multiple Access
TTG	Transmit/receive Transition Gap
UCD	Uplink Channel Descriptor
UDP	User Datagram Protocol
UE	User Equipment
UGS	Unsolicited Grant Service
UIUC	Uplink Interval Usage Code
UL	Uplink
UMTS	Universal Mobile Telecommunications
USIM	Universal Subscriber Identity Module
UTRAN	UTRA Network
VoIP	Voice over IP
WCDMA	Wideband Code Division Multiple Access
WEP	Wired Equivalent Privacy
WFQ	Weighted Fair Queuing
Wi-Fi	Wireless Fidelity
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network

References

[1] IEEE Std 802.11. Wireless Local Area Networks. Webpage, February 2006. http://www.ieee802.org/11/.

[2] IEEE Std 802.16-2004. Part 16: Air Interface for Fixed Broadband Wireless Access Systems. Technical report, June 2004. IEEE Standard for Local and metropolitan area networks. All images used with authorization from IEEE.

[3] IEEE Std 802.16e/D9. Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems - Amendment for Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands. Technical report, June 2005. Draft IEEE Standard for Local and metropolitan area networks.

[4] IEEE Std 802.20. IEEE 802.20 Mission and Project Scope. Webpage, February 2006. http://grouper.ieee.org/groups/802/20/.

[5] IEEE Std. 802.21. Media Independent Handover. Webpage, February 2006. http://www.ieee802.org/21/.

[6] IEEE Std 802.3. CSMA/CD (ETHERNET). Webpage, February 2006. http://www.ieee802.org/3/.

[7] Kyung ah Kim, Chong-Kwon Kim, and Tongsok Kim. A Seamless Handover Mechanism for IEEE 802.16e Broadband Wireless Access. September 26 2005. International Conference on Computational Science (2) 2005: 527-534 http://popeye.snu.ac.kr/publication/file/05_ij_01.pdf,April 2006, Image used with authorization from authors.

[8] Wi-Fi Alliance. Wi-Fi Official Homepage. Webpage, February 2006. http://www.wi-fi.org.

[9] Yana Bi, Jianwen Huang, Xiao Hu, Bo Fang, Mei Song, and Junde Song. IEEE P802.21 Proposal on Layer 2.5 Framework and Handover Optimization. Technical report, November 2004.

[10] Jean-Pierre Bienaimé. IMT-2000 vs. Fixed Wireless Access (FWA) systems. Nairobi, Kenya, May 2005. UMTS Forum. ITU/BDT Regional Seminar on Fixed Mobile Convergence and Guidelines on the smooth transition of existing mobile networks to IMT-2000 for Developing Countries for Africa.

[11] Howstuffworks.com Curt Franklin. How DSL Works. Webpage, February. http://computer.howstuffworks.com/dsl.htm.

[12] Erik Dahlman, Hannes Ekström, Anders Furuskär, Jonas Karlsson, Michael Meyer, Stefan Parkvall, Johan Torsner, and Mattias Wahlqvist. The long-term evolution of 3G. Technical report, 2005. http://www.ericsson.com/ericsson/corpinfo/publications/review/2005_02/files/200506.pdf, April 2006.

[13] Marc Danzeisen, Torsten Braun, Isabel Steiner, and Marc Heissenbüttel. Simulations on Heterogeneous Networking with CAHN. Technical report. http://www.iam.unibe.ch/ braun/intern/ccnc.pdf, April 2006.

[14] Dr. J. Deißner, J. Hübner, D. Hunold, and Dr. J. Voigt. *Radiowave Propagation Simulator User Manual Version 5.3*, March 2004. Images used with authorization from Radioplan.

[15] Ericsson. WCDMA Evolved. The first step - HSDPA. Technical report, May 2004. http://www.ericsson.com/technology/whitepapers/wcdma_evolved.pdf, April 2006.

[16] ETSI. Welcome to ETSI! Webpage, February 2006. http://www.etsi.org/.

[17] eWeek. Public Wireless LAN Trends. Webpage, February 2006. http://www.fiberlink.com/release/en-US/Home/KnowledgeBase/Resources/Stats/.

[18] Lutz Fielbrandt. WiMAX - Demystified, September 2005. Funkschau Workshop WLAN & WiMAX, Image used with authorization from author.

[19] WiMAX Forum. WiMAX End-to-End Network Systems Architecture. Technical report, August 15 2005. Draft. Stage 2: Architecture Tenets, Reference Model and Reference Points.

[20] WiMAX Forum. WiMAX End-to-End Network Systems Architecture. Technical report, December 14 2005. Draft. Stage 2: Architecture Tenets, Reference Model and Reference Points.

[21] WiMAX Forum. WiMAX Home. Webpage, September 2005. http://www.wimaxforum.org/home.

[22] Priyank Garg, Rushabh Doshi, Russel Greene, Mary Baker, Majid Malek, and Xiaoyan Cheng. Using IEEE 802.11e MAC for QoS over Wireless. February 2006. In Proceedings of the 22nd IEEE International Performance Computing and Communications Conference (IPCCC 2003), Phoenix, Arizona, April 2003. IEEE Computer Society. http://mosquitonet.stanford.edu/software/802.11e/ipc84.pdf.

[23] Wi-LAN Inc. Gordon Antonello. WiMAX Standards - a "IEEE 802.16 Standards Update. Webpage, February 2006. http://www.wimax.com/commentary/spotlight/news10-30-05mw1.

[24] IEEE 802.16 Broadband Wireless Access Working Group. IEEE 802.16 Documents and contributions. Technical report, November 2004. http://0-grouper.ieee.org.csulib.ctstateu.edu/groups/802/16/docs/05/80216-05_010.pdf, April 2006.

[25] IEEE 802.16 Broadband Wireless Access Working Group. Soft Handover and Fast BS Switching Procedure. Technical report, June 2004. http://www.ieee802.org/16/tge/contrib/C80216e-04_171.pdf, April 2006.

[26] PCTEL's Mobility Solutions Group. Mobility Management Suite. Company Webpage, February 2006. http://mobilitysolutions.pctel.com.

[27] International Engineering Consortium (IEC). Universal Mobile Telecommunications System (UMTS) Protocols and Protocol Testing. Webpage, February 2006. http://www.iec.org/online/tutorials/umts/.

[28] IEEE. IEEE. Webpage, February 2006. http://www.ieee.com/portal/site/iportals/.

[29] IETF. IETF Home Page nss. Webpage, February 2006. http://www.ietf.org/.

[30] Telcordia Technologies Inc. Telcordia Digest of Technical Information - Invitations to Participate - Generic Requirements for Mobile Broadband Wireless Access. Webpage, December 2005. http://www.telcordia.com/services/genericreq/digest/invitations/newgrmbwa.html.

[31] Webmethods Inc. WiMAX and WiFi Hotspot Usage Settlements Using XML based IPDR and RADIUS. Technical report, February 2004. http://www1.webmethods.com/PDF/webMethods_wp-WiFi.pdf, April 2006.

[32] NetworkWorld.com Jim Duffy. Wireless group studies 'Super 3G' standard. Webpage, February 2006. http://www.networkworld.com/edge/news/2005/01053gpp.html.

[33] WiFi Planet Jim Geier. 802.11 Alphabet Soup. Webpage, February 2006. http://www.wi-fiplanet.com/tutorials/article.php/1439551.

[34] David Johnston and Hassan Yaghoobi. Peering Into the WiMAX Spec: Part 1. Webpage, February 2006. http://www.commsdesign.com/design_corner/?articleID=17500156.

[35] Petri Jokela, Pekka Nikander, Jan Melen, Jukka Ylitalo, and Jorma Wall. Host Identity Protocol - Extended Abstract. Wireless World Research Forum (WWRF8bis), Beijing, China, February 26-27, 2004.
http://www.jokela.org/publications/wwrf8bis.pdf, April 2006.

[36] Petri Jokela, Pekka Nikander, Jan Melen, Jukka Ylitalo, and Jorma Wall. Host Identity Protocol: Achieving IPv4 - IPv6 handovers without tunneling. Evolute workshop 2003: "Beyond 3G Evolution of Systems and Services" , pp. A-2/1-5, University of Surrey, Guildford, UK, November 10, 2003.
http://www.cs.hut.fi/ pmrg/publications/VHO/2003/Jokela_Nikander _Melen_Ylitalo_Wall_HIPAHWT.pdf, April 2006.

[37] Declan O. Sullivan Julien Oberlé, Finola Bourke. Fixed Wireless Access. No date specified.

[38] P. Howard K. Boman, G. Horn and V. Niemi. UMTS Security. February 2006. Electronics & Communication Engineering Journal October 2002 http://www.c7.com/ss7/whitepapers/cellular/umts_security.pdf.

[39] UCLA Parallel Computing Laboratory. PARallel Simulation Environment for Complex systems. Webpage, February 2006. http://pcl.cs.ucla.edu/projects/parsec/.

[40] UCLA Parallel Comupting Laboratory. GloMoSim - Global Mobile Information Systems Simulation Library. Webpage, February 2006. http://pcl.cs.ucla.edu/projects/glomosim/.

[41] NyTeknik Lars Anders Karlberg. Ericsson hjälper 3 bli först med turbo-3G. February 2006. NyTeknik 050207 http://nyteknik.se/art/38771.

[42] Kin K. Leung, Sayandev Mukherjee, and Gerorge E. Rittenhouse. Mobility Support for IEEE 802.16d Wireless Networks. Technical report, 2005. IEEE Communications Society.

[43] Ye Liang, Xuehai Wang, and Congzhi Wang. Host Mobility Support in IP Network. Technical report, March 2004. http://me.cs.dal.ca/ xwang/courses/cs6704/report.pdf, April 2006.

[44] Vikki Lipset. 802.16e vs. 802.20. Webpage, February 2006. http://www.wi-fiplanet.com/columns/article.php/3072471.

[45] Ehsan Maghsoodi. Design and implementation of WLAN support for cellular assisted heterogeneous networking. Technical report, November 2004. Slides, Diploma Thesis. http://www.iam.unibe.ch/ rvs/teaching/ws04_seminar/Maghsoodi.pdf, April 2006.

[46] Supriya Maheshwari. An Efficient QoS Scheduling Architecture for IEEE 802.16 Wireless MANs. Master's thesis, Indian Institute of Technology, Bombay, India, 2005.

[47] Parmod Mehta, Gaurish M S Khandeparkar, and M C Swamy. Inter-operability between Heterogeneous Networks. Technical report, Wipro Technologies, 2002. http://wireless.ittoolbox.com/pub/PM040102.pdf.pdf, April 2006.

[48] Global mobile Suppliers Association. Specifications for 3GPP Release 6 finalized. Webpage, February 2006. http://www.gsacom.com/news/gsa_176.php4.

[49] Ronan Morrissey, Julian Leonard, and Neil O.Driscoll. UMTS (3G)/WLAN Integration.

[50] Govindan Nair, Joey Chou, Tomasz Madejski, Krzysztof Perycz, David Putzolu, and Jerry Sydir. IEEE 802.16 Medium Access Control and Service Provisioning. Webpage, February 2006. Intel® Technology Journal http://download.intel.com/technology/itj/2004/volume08issue03 /art04_ieee80216mac/vol8_art04.pdf.

[51] Computer Networks and Internets. Q & a on homogeneous and heterogeneous networks. Webpage, February 2006. http://www.netbook.cs.purdue.edu/othrpags/qanda97.htm.

[52] Pekka Nikander, Jari Arkko, and Börje Ohlman. Host Identity Indirection Infrastructure. November 2004. The Second Swedish National Computer Networking Workshop 2004 (SNCNW2004), Karlstad University, Karlstad, Sweden, Nov 23-24, 2004
http://www.tml.tkk.fi/ pnr/publications/sncnw2004.pdf, April 2006.

[53] Jorge Nuevo. *A Comprehensible GloMoSim Tutorial*. Québec, Canada, March 2004. http://externe.inrs-emt.uquebec.ca/users/nuevo/glomoman.pdf, April 2006.

[54] SPG Media PLC. HSUPA (High Speed Uplink Packet Access). Webpage, February 2006. http://www.mobilecomms-technology.com/projects/hsupa/.

[55] Radioplan. Radioplan Gives Free Software to Students. Webpage, February 2006. http://www.radioplan.com/news/news_23.html.

[56] Radioplan. Student Edition Free to Download for Educational, Personal and Evaluation Use. Webpage, February 2006. http://www.radioplan.com/download.html.

[57] Insight Research. Europe Grows WiFi Wireless Faster Than North America. Webpage, February 2006. http://www.3g.co.uk/PR/Nov2003/6060.htm.

[58] ZDNet Research. WLAN market to reach $5 bln by 2006. Webpage, December 2005. http://blogs.zdnet.com/ITFacts/wp-trackback.php?p=8280.

[59] Jochen Schiller. *Mobile Communications*. Addison-Wesley, London, 2003. ISBN: 0 321 12381 6.

[60] Susning.nu. OSI-referensmodellen. Webpage, February 2006. http://susning.nu/OSI-referensmodellen.

[61] Universal Mobile Telecommunications System. UMTS Forum. Webpage, February 2006. http://www.umts-forum.org.

[62] Scalable Network Technologies. QualNet 3.9. Webpage, February 2006. http://www.scalable-networks.com.

[63] WiFi-Planet Tim Sanders. Tutorial: The Many Flavors of OFDMA. Webpage, February 2006. http://www.wi-fiplanet.com/tutorials/article.php/3557416.

[64] Pablo Vidales, Leo Patanapongpibul, Glenford Mapp, and Andy Hopper. Experiences with Heterogeneous Wireless Networks, Unveiling the Challenges. HET-NET's 04. Second International Working Conference, Ilkley, West Yorkshire, U.K, July 26-28, 2004.
http://www.cl.cam.ac.uk/Research/DTG/ pav25/publications/HetNets04-Vidales.pdf, April 2006.

[65] the free encyclopedia Wikipedia. Handoff. Webpage, February 2006. http://en.wikipedia.org/wiki/Handoff.

[66] the free encyclopedia Wikipedia. Public Switched Telephone Network. Webpage, February 2006. http://en.wikipedia.org/wiki/PSTN.

[67] wikipedia.com. diversity Combining. Webpage, February 2006. http://en.wikipedia.org/wiki/Diversity_combining.

[68] Sarah Kate Wilson, Joanne Wilson, John Chen, and Reza Arefi. Broadband mobile systems. Technical report, February 2006. Contribution to 802.20 working group meeting July 8-12 2002,http://grouper.ieee.org/groups/802/20/Dot16_Ar/C80216sgm-02_19r1.pdf, April 2006.

[69] WiMAX.com. What is the WiMAX Security scheme/protocol? Webpage, February 2006. http://www.wimax.com/education/faq/faq29.

[70] NetMotion Wireless. NetMotion Mobility XE. Company Webpage, February 2006. http://www.netmotionwireless.com/product/mobility_keys.asp.

[71] The Register Wireless Watch. WiMAX summit: Standards-plus could harm 802.16 roadmap. Webpage, February 2006. http://www.theregister.co.uk/2005/04/11/wimax_summit/.

[72] UMTS World. Overview of The Universal Mobile Telecommunication System. Webpage, February 2006. http://www.umtsworld.com/technology/overview.htm.

[73] UMTS World. UMTS Handover. Webpage, February 2006. http://www.umtsworld.com/technology/handover.htm.

[74] Mika Ylianttila. *Vertical Handoff and Mobility - System Architecture and Transition Analysis.* PhD thesis, Faculty of Technology, University of Oulu, 2005.

[75] Xiang Zeng, Rajive Bagrodia, and Mario Gerla. GloMoSim: A library for Parallel Simulation of large-scale wireless networks. Technical report, University of California, California, Los Angeles. http://www.scalable-networks.com/pdf/glomosim.pdf, April 2006.

Appendix A

Sequence Charts

A.1 Network Entry

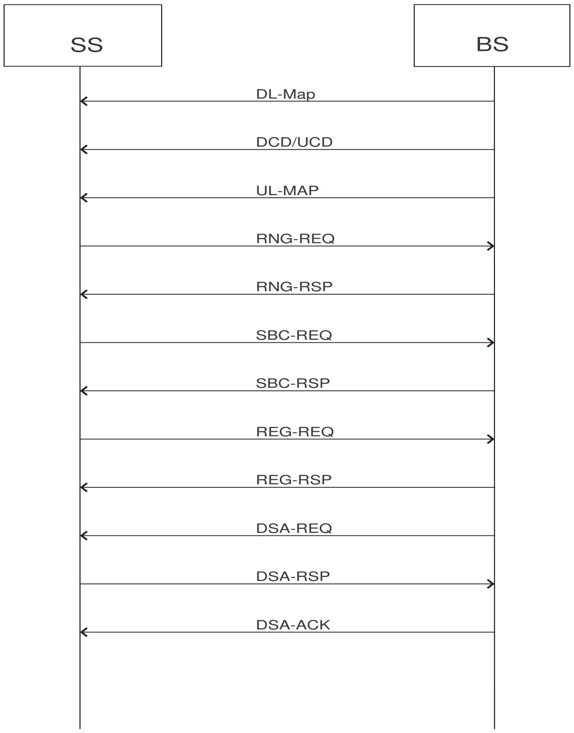


Figure A.1: Initial Network Entry Message Exchange

A.2 Hard Handover

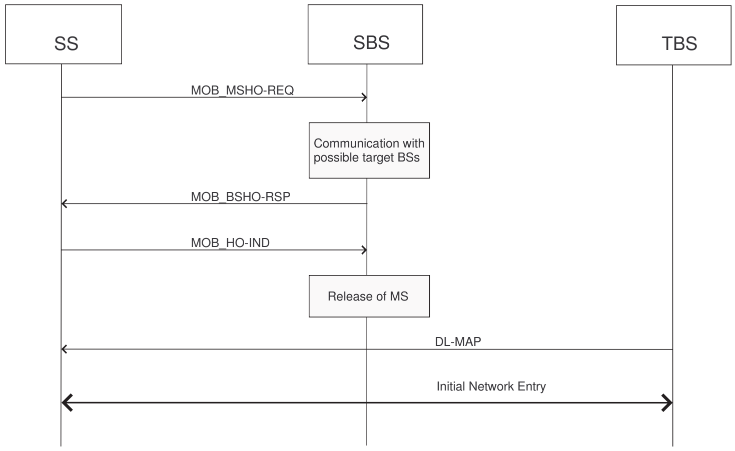


Figure A.2: Hard Handover Message Exchange

A.3 Fast Base Station Switching

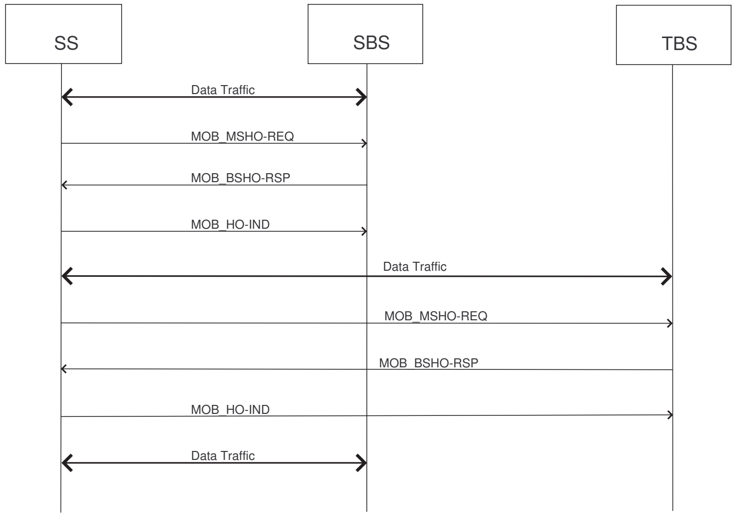


Figure A.3: Fase Base Station Switching Message Exchange

Appendix B

802.16 MAC-summary

Throughout this thesis a lot of effort has been put into reading and understanding the technological specifications of the 802.16 standard. To summarize this work for those interested in the MAC-layer, but wanting to avoid the complete technical specifications, this appendix present key aspects of the 802.16-2004 and 802.16e air interface.

Figure B.1 shows the layering of the 802.16 standard. As 802.16e is an ammendment to the 802.16-2004 standard, the term 802.16 is used when describing characteristics that are similar in both.

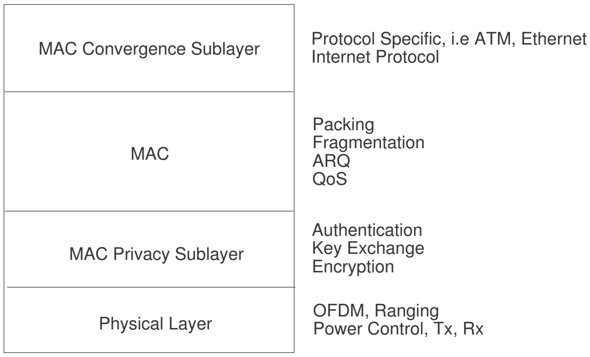


Figure B.1: Layering of IEEE Std. 802.16

This appendix covers the two upper most layers, the convergence- and MAC-layer. They form the fundamental parts of the air interface which governs how the limited radio resources are shared by base stations and subscriber stations.

B.1 Convergence Sublayer

The Convergence Sublayer(CS) performs the following two main tasks[2, p. 17].

– Packet Classification

– Payload Header Suppression

The draft contains two specifications for convergence layer, ATM and packet CS. This section covers the packet CS, but much of the theory applies to ATM CS as well.

The process of Packet Classification is where the CS receives a higher layer packet and maps this packet to a service flow[1]. Since each service flow is associated with specific QoS parameters, the classification of a packet to a flow leads to the delivery of that packet with appropriate QoS constraints.

The classification is made based on different criterion, such as destination or source IP-adders. If a packet matches a criteria it is delivered to a MAC connection that has been matched to that criteria, i.e the classification results in an appropriate connection identifier(CID) of a connection, as seen in Figure B.2

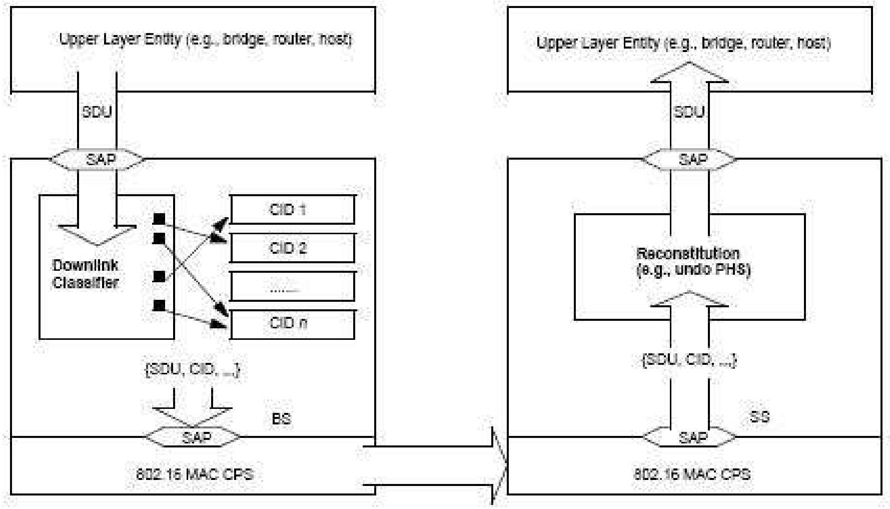

Upper Layer Entity (e.g., bridge, router, host)
SDU
SAP
Downlink Classifier
CID 1
CID 2
....
CID n
{SDU, CID, ...}
BS
SAP
802.16 MAC CPS
Upper Layer Entity (e.g., bridge, router, host)
SDU
SAP
Reconstitution (e.g., undo PHS)
{SDU, CID, ...}
SS
SAP
802.16 MAC CPS

Figure B.2: Classification and CID mapping[2, p. 22]

Several classifiers may exist for the same service flow, and since they can overlap they are explicitly ordered.

The capability of Payload Header Suppression(PHS) is optional and is used to remove repetitive or redundant information from higher layer packets headers.

When a packet is classified it can also be mapped to a Package Header Suppression Rule. The information in the packet header is compared with a Package Header Suppression Field(PHSF). If the header bits match the PHSF, some of these bits can be masked. Such bits desirable to mask can be higher layer static fields, such as IP-adders. Dynamic fields can be left intact by using the Package Header Suppression Mask, which specifies which bits that are not to be suppressed.

When receiving a packet applied to PHS, the receiver unmasks the appropriate bits and reassembles the packet headers before delivery to higher layers. The information needed for PHS is thus needed on both receiving and sending entity. A Package Header Suppression Index(PHSI) is added to the packet as a reference to the appropriate PHSF on the receiving side.

[1] A service flow is a connection with a set of QoS parameters

PHS Rules can be created dynamically through management messages like Dynamic Service Addition(DSA)[2, p. 62] or Dynamic Service Change(DSC)[2, p. 65]. These rules can also be created over time, e.g where some fields are unknown at creation but made available and added to the rule later.

After passing through the CS, the packet delivered to the appropriate service flow in the MAC-layer has the format shown in Figure B.3. The standard calls this a MAC SDU, and it contains the PHSI[2] and the higher layer PDU.

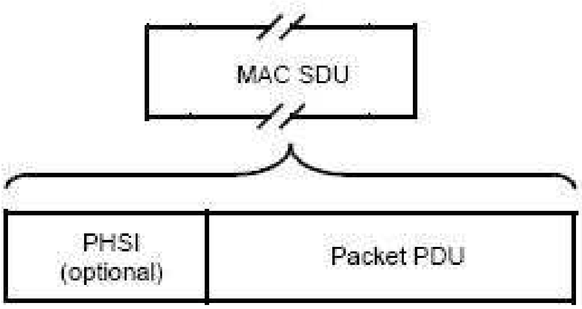


Figure B.3: MAC SDU Format[2, p. 21]

The MAC SDU is now in the hands of the MAC-layer which is responsible for delivering the packet to the receiver over the air interface.

B.2 Medium Access Control

The MAC-layer is what controls and distributes the limited radio spectrum resources between the base stations and subscriber stations, while the physical(PHY) layer handles radio specific functions like modulation, encoding and physical frequency allocations.

Since the medium is limited by the radio bandwidth, it is important that this resource distribution is done as efficiently as possible.

The following are some functions performed by the MAC-layer.

– Scheduling of data

– QoS setup and maintenance

– Connection management

– Handovers, Idle/sleep mode

The 802.16-2004 specification supports two modes of operation; PMP mode or Mesh mode. Mesh mode is mainly used for ad hoc networks[59, p. 330] or similar systems that can make use of the capability of subscriber stations to route and send information directly to each other. In PMP mode traffic can only occur between a base stations and subscriber stations.

This chapter covers MAC in PMP mode, but much of the theory is applicable to Mesh mode operation as well.

[2] So that the receiver can select the appropriate PHSF for unmasking the header

The 802.16 MAC-layer is connection oriented and all data communication is associated to a connection. A connection together with QoS parameters make up a service flow, which is a fundamental term in the standard.

QoS is maintained through five different QoS-classes, which basically are different scheduling mechanisms.

B.2.1 Framing

The MAC-layer has support for both TDD and FDD framing, where TDD separates uplink and downlink in time and FDD separates them by frequency. Figure B.4 shows how Physical Slots(PS) make a general TDD frame structure.

The frame size can be varied in accordance to different physical profiles. The partition of the frame between uplink and downlink can also be adjusted.

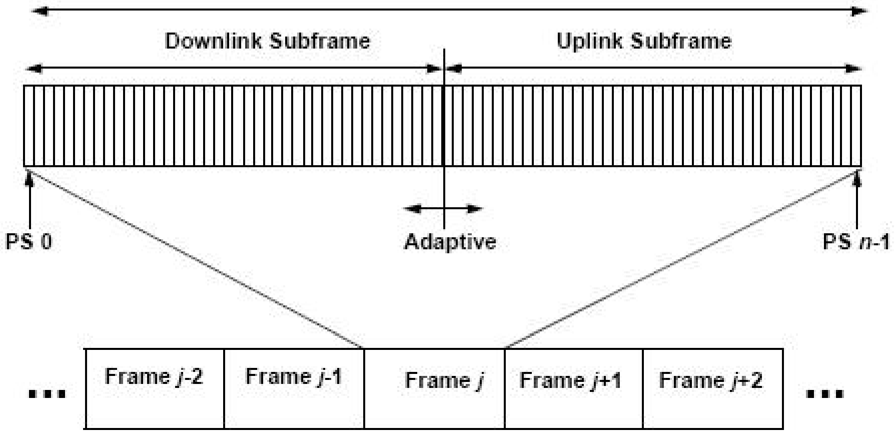


Figure B.4: TDD Frame Structure[2, p. 153]

As an example of the contents of the uplink and downlink parts of each frame, Figure B.5 on the next page shows the OFDM PHY specification frame structure which best illustrates the relations of PDUs and other parts of the up/downlink.

Between the uplink and downlink there are guard spaces to allow for switching the radio between receiving and transmitting mode. These guard spaces are called Transmit/receive Transition Gap(TTG) and Receive/transmit Transition Gap(RTG).

Downlink

The downlink (DL) consist of several physical bursts of different modulation/coding and where the bursts are sent in decreasing robustness. These bursts are addressed to different connections through Connection Identifiers(CID). These identifiers can identify burst addressed to individual subscriber stations, broadcast messages or multicast messages and have specified values, e.g multicast polling bursts always have a CID between 65280-65533[2, p. 643].

After synchronization preambles and other PHY required data comes the broadcast burst. This burst contains important control messages like the DL-MAP, UL[3]-MAP,

[3] Uplink

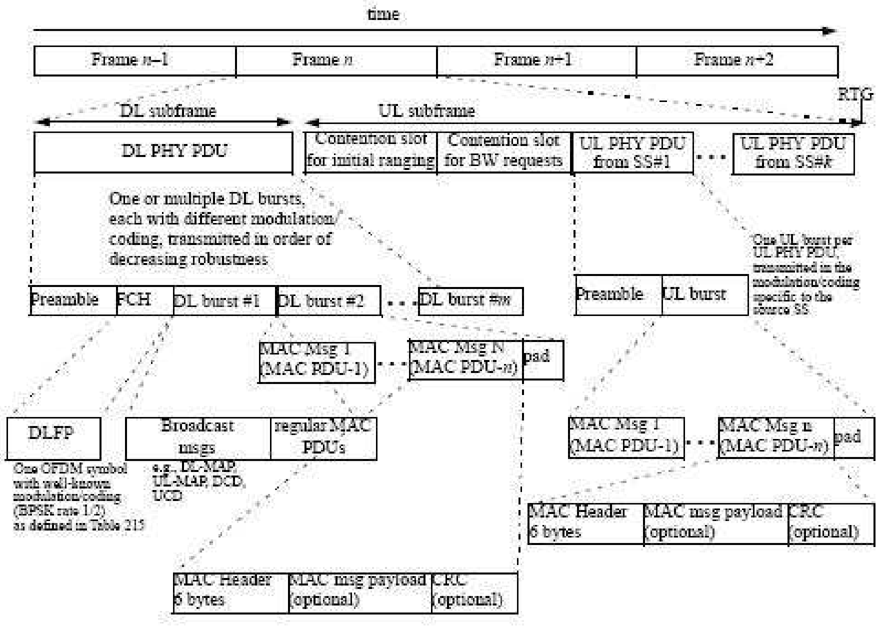

time
Frame n–1
Frame n
Frame n+1
Frame n+2
RTG
DL subframe
UL subframe
DL PHY PDU
Contention slot for initial ranging
Contention slot for BW requests
UL PHY PDU from SS#1
UL PHY PDU from SS#k
One or multiple DL bursts, each with different modulation/ coding, transmitted in order of decreasing robustness
Preamble
FCH
DL burst #1
DL burst #2
DL burst #m
Preamble
UL burst
One UL burst per UL PHY PDU, transmitted in the modulation/coding specific to the source SS
MAC Msg 1 (MAC PDU-1)
MAC Msg N (MAC PDU-n)
pad
DLFP
Broadcast msgs
regular MAC PDUs
One OFDM symbol with well-known modulation/coding (BPSK rate 1/2) as defined in Table 215
e.g., DL-MAP, UL-MAP, DCD, UCD
MAC Msg 1 (MAC PDU-1)
MAC Msg n (MAC PDU-n)
pad
MAC Header 6 bytes
MAC msg payload (optional)
CRC (optional)
MAC Header 6 bytes
MAC msg payload (optional)
CRC (optional)

Figure B.5: OFDM Frame Structure[2, p. 450]

UCD and DCD messages.

The DL-MAP, if scheduled for transmission in the frame, is always first and this describes the content of the downlink. It contains base station and operator identifier and also contain information elements(IE) regarding the downlink data bursts, i.e transmission from the BS to different MSs.

These IEs in the DL-MAP act as a type of pointer to bursts on the downlink. They contain information on what CID the burst is addressed to and what type of burst it is. This is to allow the MS to through the DL-MAP identify which bursts it should listen to and how it should set its radio so it can decode them.

Figure B.6 on the following page shows an example of how the DL-MAP contains information regarding when a particular burst starts. The Downlink Interval Usage Code(DIUC) and its UL-MAP counterpart, Uplink Interval Usage Code(UIUC) is an index to the type of bursts. This is the index who's details is found in the periodic DCD/UCD messages.

After the DL-MAP comes the UL-MAP, which functions in much the same way as the DL-MAP but instead describes the uplink bursts. These burst have been allocated by the BS for MSs to send uplink data. The IEs in the UL-MAP describe what type of burst it is and what time the allocated burst starts.

The UCD and DCD messages define the detailed characteristics of the uplink and downlink bursts. The IEs in the UL- and DL-MAP contains references to what type each burst is, but the UCD and DCD messages contain the details on how each type of burst should be decoded.

These are transmitted at a less frequent interval and changes in the modulation of each burst type is signaled through a DCD/UCD Count. If this changes in the UL- or

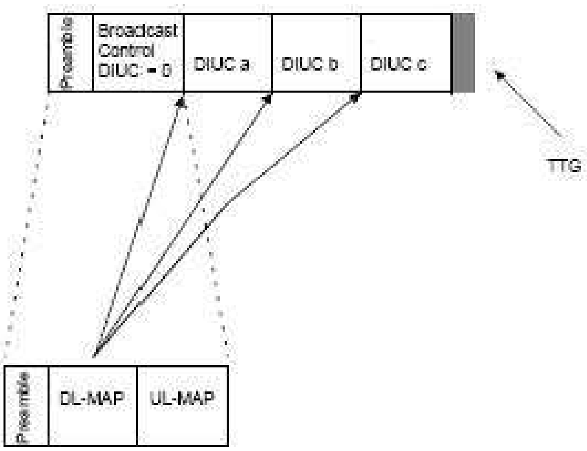


Figure B.6: DL-MAP references to downlink bursts of varying types[2, p. 310]

DL-MAP the information in the UCD or DCD messages has changed and the affected burst has new physical characteristics.

The bursts can thus change physical characteristics to adapt to the environment or requirements on data rates. The DCD/UCD counts lets the stations ignore old information until a change is detected.

The bursts following the first, broadcast burst, contains PDUs addressed to subscriber stations.

Uplink

The uplink is shared by the subscriber stations, where each subscriber station uses their own modulation and encoding to transmit to the base station.

The IEs in the UL-MAP sent on the downlink indicates who each burst is addressed to and what type of burst it is. The part of the uplink is reserved for contention based initial ranging, used when subscriber stations need to perform ranging with a base station to achieve synchronization and appropriate signal strengths for further communication.

The next part of the uplink is for contention based bandwidth requests, where subscriber stations can send request messages without waiting for individually allocated bursts.

As these two periods are contention based, collisions may occur between subscriber stations transmitting at the same time during this period. To handle this the UCD message contains information on backoff window for both ranging and bandwidth request intervals.

When a MS wants to transmit in one of these contention periods it sets its contention window according to the UCD message and randomly selects a number within this windows that represents the number of transmission opportunities the MS should defer transmission.

Transmission opportunities are of known size, stated in the UCD message. The MS counts the number of transmission opportunities that is valid for transmission of the message, i.e the countdown does not continue for opportunities in other types of bursts.

After deferring the appropriate number of transmission opportunities the MS can send its message. A lack of reply is interpreted as the contention transmission being lost. The backoff windows is increased by a factor of two and the process repeated until a defined number of maximum retries have been reached.

The BS has through the UCD specified backoff information power to shape the contention process, e.g setting the backoff window to create Ethernet style backoff.

The following bursts are on a subscriber basis, i.e each subscriber station has its own allocated portion of the uplink to send data using one specific burst type.

B.2.2 MAC PDUs

802.16 has a wide array of management messages[2, p. 43]. These management messages and upper layer data packets can be packaged and sent using different methods.

The standard makes use of Type/Length/Value(TLV), a formating scheme that enables more dynamic contents in messages. Parameters are sent using this scheme which enables parsing of a predefined type, length and value. Messages can thus contain variable number of parameters, which can be detected using this scheme.

PDUs can either be concatenated to fill an allocated burst, or fragmented. Fragmentation can allow for more efficient use of allocated bandwidth with regards to QoS requirements.

There is also support for the Automatic Repeat Request(ARQ) protocol for handling retransmission of erroneous messages. This can be enabled for individual connections. Automatic Repeat Request is an error detection and correction technique. The receiver detects errors but cannot correct them and instead sends a retransmission request to the transmitter which then repeats the transmission.

Figure B.7 shows the MAC PDU format. Each PDU consists of a fixed length MAC header. The payload contains zero or more subheaders and zero or more SDUs or fragments thereof. The size of a MAC PDU varies based on the contents of the payload. The PDU may optionally contain a Cyclic Redundancy Check(CRC) at the end.

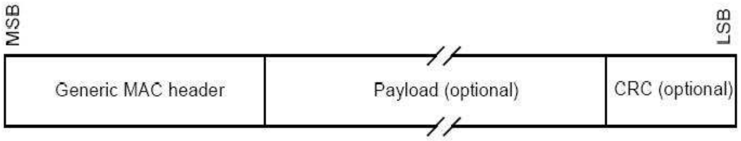


Figure B.7: MAC PDU Format[2, p. 35]

B.2.3 Connections

Each SS has a 48-bit universal MAC address, much like the MAC address of regular network interfaces. It is used when first identifying an SS towards a BS or in the authentication process.

Each connection setup by the BS is identified by a 16-bit CID, which works as an index for looking up type of service or other parameters.

As a subscriber station enters a BS cell and performs the steps necessary to begin data communication it is granted two mandatory and one optional management connections.

- Basic Connection
- Primary Management
- Secondary Management

These three connections are associated to a QoS level, which differentiates management into three priorities.

The Basic Connection is used for time critical, short MAC management messages, while the Primary Management connection delivers longer more delay tolerant MAC messages. The Secondary Management connection is used for standards based protocols like DHCP, TFTP[4] etc. An SS can have a static configuration and thus not need the Secondary Management connection, e.g static IP-address and thus no need for DHCP.

When these three management connections are set up, they each receive a CID that identifies that connection. The CID is the same for both SS and BS and is mediated through the RNG-RSP[2, p. 50] and REG-RSP[2, p. 52] MAC management messages during the network entry or base station association phase.

The standard specifies what type of management messages use which type of connection.

Besides these management connections there is also the possibility to setup transport connections. These connections hold PDUs delivered to the MAC-layer from higher layers and are also associated with a scheduling service based on their QoS requirements.

Transport connections are setup based on what has been provisioned to the SS. The registration of an SS with the network management entities or changes in the services subscribed to by the user initiates setup or change of connections.

Connections are the lower level of data transport services and are associated with a higher level service flow. As mentioned previously, a connection together with QoS parameters constitute a service flow.

A service flow can be anyone of three types: Provisioned, Admitted or Active. A provisioned flow is a flow who's parameters are known by the network management system, e.g such as a template for different flows.

An admitted flow has resources reserved but not committed and Active flows have their required QoS resources committed by the BS, e.g the BS is providing bandwidth request opportunities or unsolicited grants.

The two last states form a two-phase activation model[5] where resources are conserved(admitted) until there is an end-to-end connection established and then activated.

B.2.4 Scheduling

The 802.16e[3, p. 150] draft specifies five different scheduling mechanisms and each connection is associated with one of these types. The services are defined by a set of QoS parameters that quantify its behavior.

The five scheduling classes, also called QoS classes, are:

- UGS - Unsolicited Grant Service
- rtPS - real time Polling Service
- nrtPS - non real time Polling Service

[4] Trivial File Transfer Protocol

[5] Also known in telephony applications

– BE - Best Effort

– ertPS - extended real time Polling Service

The scheduling services can be managed through the DSA and DSC management messages.

By associating a connection with a scheduling service the BS can anticipate the behavior and needs of a connection. The scheduling services defines how data is granted to an SS and what mechanisms are used.

UGS

This services is designed for real-time data streams with fixed size data packets issued at periodic intervals, e.g constant bit rate applications like VoIP without silence suppression.

The BS allocates periodic data bursts for this service based on the size of the Maximum sustained Traffic Rate. The SS reports the status of the connection associated to the service with a Slip indicator, which if set indicates that the queue of packets haven grown too large and the BS needs to allocated larger bursts for this connection.

Since the traffic rate is allocated on a real time need basis, there is no need for an MS to send requests, except through the Slip indicator present in the Grant Management Subheader. This subheader is an optional part of the generic management header[2, p. 35].

Mandatory parameters for this service are Maximum Sustained Traffic Rate, Maximum Latency, Tolerated Jitter and Request/Transmission Policy.

rtPS

rtPS services support real-time service flows with variable data size packets, e.g MPEG[6] video. The BS schedules periodic unicast opportunities for the SS to request bandwidth based on current real-time needs and packet sizes.

Mandatory parameters for this service are Minimum Reserved Traffic Rate, Maximum Sustained Traffic Rate, Maximum Latency and Request/Transmission Policy.

nrtPS

This service support more delay tolerant data streams with variable size data packets where there is a requirement for minimum data rate.

Bandwidth requests for this service can be made both during the BS scheduled request intervals or in contention periods.

Mandatory parameters for nrtPS are Minimum Reserved Traffic Rate, Maximum Sustained Traffic Rate, Traffic Priority and Request/Transmission Policy.

BE

Best Effort services manage data streams with no minimum service level and are reserved bandwidth based on resource availability. This type of service class is useful for e.g background applications where delay has no impact.

Mandatory parameters for this service are Maximum Sustained Traffic Rate, Traffic Priority and Request/Transmission Policy.

[6] Moving Picture Experts Group

ertPS

This scheduling service is introduced in the 802.16e draft and is a combination of UGS and rtPS. Like UGS data traffic is allocated bandwidth without solicitation but, unlike UGS, nrtPS can make this allocation on a dynamic manner. It thus has less overhead than rtPS that requires unicast request opportunities, but retains the flexibility of variable size grants.

An example of ertPS application is VoIP with silence suppression.

Mandatory parameters for Extended rtPS is Maximum Sustained Traffic Rate, the Minimum Reserved Traffic Rate, the Maximum Latency and the Request/Transmission Policy.

B.2.5 Bandwidth Requests

An SS may request bandwidth either by sending an Bandwidth Request message or through an optional Piggyback Request in a Grant Management subheader.

These requests can be sent in any uplink burst, except during an initial ranging interval. The requests are either incremental or aggregate, where an incremental request adds the requested bandwidth to the amount already allocated by the BS. An aggregate request replaces the current value with the value in the request.

Piggyback requests can only be incremental since they have no field to indicate the type of the request and requests done during bursts where collisions can occur should be aggregate.

An SS always request bandwidth with reference to a particular connection, but the BS allocates bandwidth addressed to the SS Basic Management Connection, and not individual connections. This makes the SS unaware of which request the grant refers to.

The BS can assign bandwidth for the explicit purpose of SS bandwidth request messages. This can be done in either groups of SSs or individual SSs and come in the form of specific burst types in the UL-MAP.

Unicast polling is done with bursts directed at the SS's Basic CID and broadcast or multicast polling is done with contention interval bursts that has standard defined CIDs.

Multicast polling is useful when there is insufficient bandwidth to perform unicast polling on a large amount of, perhaps inactive, SSs. Downlink multicast is achieved by assigning the same CID to a connection shared by all members of the multicast group.

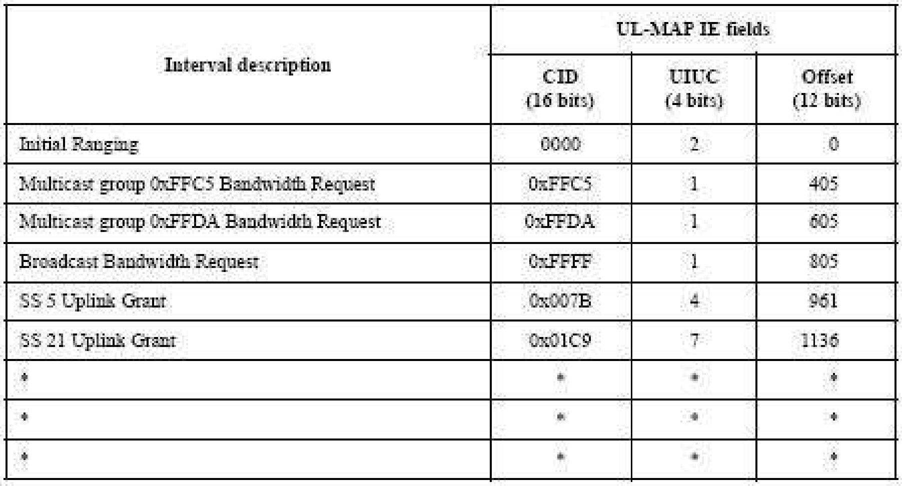
Figure B.8 on the facing page shows the IEs of an UL-MAP. The offset states where the burst starts and the UIUC-field states the encoding and modulation type that the SS should transmit on during this interval. The CID fields show how the bursts are addressed to either two different multicast groups, broadcast or individual SS. The Uplink Grants are intervals where the SSs can send uplink data, while the bandwidth request regions are for sending bandwidth requests only.

B.2.6 Network Entry and Handover

The process of associating a mobile subscriber with a base station is similar to that of the mandatory handover process, stated in the 802.16e draft.

Figure B.9 on page 86 shows initial network entry, followed by a handover to another BS.

The process of network entry[2, p. 167] contains steps such as acquiring physical(bit) and MAC(frame) synchronization, ranging to adjust suitable transmission power, setup



Interval description	UL-MAP IE fields		
	CID (16 bits)	UIUC (4 bits)	Offset (12 bits)
Initial Ranging	0000	2	0
Multicast group 0xFFC5 Bandwidth Request	0xFFC5	1	405
Multicast group 0xFFDA Bandwidth Request	0xFFDA	1	605
Broadcast Bandwidth Request	0xFFFF	1	805
SS 5 Uplink Grant	0x007B	4	961
SS 21 Uplink Grant	0x01C9	7	1136
*	*	*	*
*	*	*	*
*	*	*	*

Figure B.8: Sample UL-MAP with multicast and broadcast IE[2, p. 145]

of management connections, authentication and retrieval of provisioned services and higher layer protocol interaction such as IP-connectivity through DHCP.

In the case of handover[3, p. 170], the MS can during certain intervals perform ranging to synchronize and obtain information about neighboring BSs that could serve as candidates for handover. It can also receive this information from its current BS through a certain broadcast message, MOB_NBR-ADV. This message contains channel information provided by neighboring BSs DCD/UCD messages. This information is distributed to the BSs through the backbone and saves the MS the trouble of scheduling intervals to perform the scanning process.

There are different methods for scanning neighboring base stations[3, p. 172], where the MS can to different degrees associate itself with a possible new anchor BS. Depending on how much information is exchanged during this scanning process, the actual handover can be shortened since required information already is available.

B.2.7 Sleep Mode

A MS can enter sleep mode to save power or to decrease the usage of BS resources[7]. The BS manages Power Saving Classes for each MS. Each MS connection is associated to a particular Power Saving Class and there are three different classes. These three different classes are suitable for different service flows, as each flow has different requirements and methods for requesting bandwidth and sending data.

An example would be one rtPS and one BE service flow belonging to a single class while two UGS connections belong to two different classes, depending on their QoS parameters.

The Power Saving classes can be activated/deactivated repeatedly, with each class having its own sleep/listening window. Periods where this window does not overlap the MS is considered unavailable and the BS should not send packets to the MS. The BS

[7] E.g a sleeping MS does not need to be scheduled bandwidth

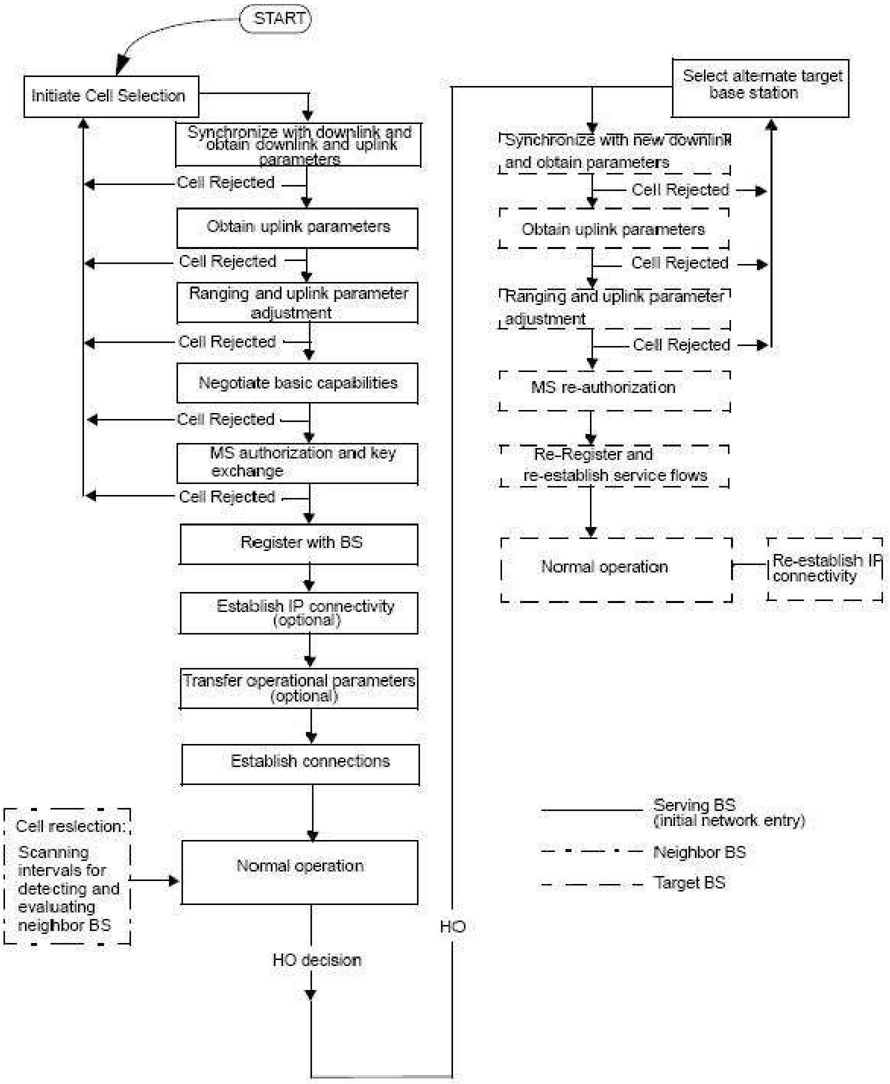

START
Initiate Cell Selection
Synchronize with downlink and obtain downlink and uplink parameters
Cell Rejected
Obtain uplink parameters
Cell Rejected
Ranging and uplink parameter adjustment
Cell Rejected
Negotiate basic capabilities
Cell Rejected
MS authorization and key exchange
Cell Rejected
Register with BS
Establish IP connectivity (optional)
Transfer operational parameters (optional)
Establish connections
Cell reslection: Scanning intervals for detecting and evaluating neighbor BS
Normal operation
HO decision
HO
Select alternate target base station
Synchronize with new downlink and obtain parameters
Cell Rejected
Obtain uplink parameters
Cell Rejected
Ranging and uplink parameter adjustment
Cell Rejected
MS re-authorization
Re-Register and re-establish service flows
Normal operation
Re-establish IP connectivity
Serving BS (initial network entry)
Neighbor BS
Target BS

Figure B.9: Handover and initial network entry[3, p. 176]

can choose to drop packets or buffer them. Multicast packets can be delayed until all members of the multicast group are available.

Figure B.10 on the facing page shows an example of an MS with two Power Saving Classes and how their listening windows overlap to create intervals where the MS is available or unavailable.

B.2.8 Idle Mode

Idle mode is a mechanism used to allow a MS to move around a larger geographical area and possibly many cell boundaries without having to perform handover or send other

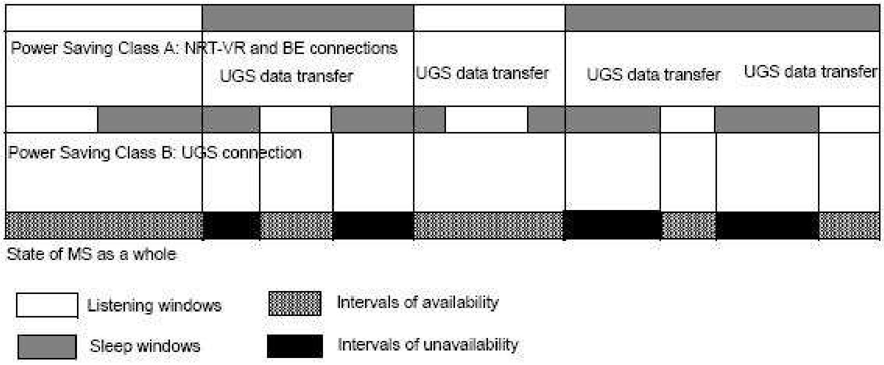

Power Saving Class A: NRT-VR and BE connections
UGS data transfer
UGS data transfer
UGS data transfer
UGS data transfer
Power Saving Class B: UGS connection
State of MS as a whole
Listening windows
Sleep windows
Intervals of availability
Intervals of unavailability

Figure B.10: MS in sleep mode with two Power Saving Classes[3, p. 166]

management traffic. This is primarily intended for MS that are inactive and through this mechanism they can still be reached on the downlink without being registered to a specific BS. It thus removes many of the MAC management requirements of handover and normal operations.

A group of BSs constitute a paging group, where the MS can be contacted on the downlink without the need for the MS to send on the uplink. These paging groups needs to balanced in size so the overhead does not become to big while keeping the group large enough to provide longer mobility range.

Idle mode can be initiated by either the MS or by the BS and this process is controlled by the DREG-group[3, p. 58] of management messages. The BS can together with the management system retain certain MS information for future MS network re-entry.

The paging group has a known cyclic interval of Paging Broadcasts, which the MS synchronizes with after decoding a preferred BS downlink. The BS can during Paging Broadcast send data to the MS or command the MS to perform network entry or perform ranging to establish location and acknowledge messages.

since MSs does not have CIDs to identify messages addressed to them, a Paging Broadcast contains a hash of the MSs MAC address.

Location Updates are done in regular intervals and can be done in either a secure or un-secure manner. These updates are done by the MS performing network entry and informing the BS of location in the ranging phase. If the MS has moved to a new Paging Group it updates its Paging Group ID and the BS informs the old paging group of the MS transition through the backbone.

The paging controller unit and the MS store an Idle Timer which is updated when the MS performs Location Update. If the timer expire the paging controller discards the paging information regarding the MS and the MS assumes that the paging controller has discarded this information.