

Mobile WiMAX

by Jarno Pinola and Kostas Pentikousis, VTT Technical Research Centre of Finland

One of the technologies that can lay the foundation for the next generation (fourth generation [4G]) of mobile broadband networks is popularly known as “WiMAX.” WiMAX, *Worldwide Interoperability for Microwave Access*, is designed to deliver wireless broadband bitrates, with *Quality of Service* (QoS) guarantees for different traffic classes, robust security, and mobility. This article provides an overview of mobile WiMAX, which is based on the wireless local and *Metropolitan-Area Network* (MAN) standards IEEE 802.16-2004^[1] and 802.16e-2005^[2]. We introduce WiMAX and focus on its mobile system profile and briefly review the role of the WiMAX Forum. We summarize the critical points of the WiMAX network reference model and present the salient characteristics of the PHY and MAC layers as specified in [1] and [2]. Then we address how mobile nodes enter a WiMAX network and explain the fundamentals of mobility support in WiMAX. Finally, we briefly compare WiMAX with *High-Speed Packet Access* (HSPA), another contender for 4G.

The Role of the WiMAX Forum

The WiMAX Forum is a nonprofit organization formed in 2001 to enhance the compatibility and interoperability of equipment based on the IEEE 802.16 family of standards. The IEEE 802.16 standards provide a large set of fundamentally different options for designing a wireless broadband system, including, for example, multiple options for *Physical* (PHY) layer implementation, *Media Access Control* (MAC) architecture, frequency bands, and duplexing. So many options lead to several possible system variants, which are all compatible with the IEEE standards. Although such multiplicity allows for deployment in very diverse environments, it may spell either solely vertical, single-vendor deployments or no deployment at all, because operators do not want to be locked in with any particular implementation. Thus, a major motivation for establishing the WiMAX Forum was to develop predefined system profiles for equipment manufacturers, which include a subset of the features included in the IEEE 802.16 standards. WiMAX Forum-certified products are guaranteed to be interoperable and to support wireless broadband services from fixed to fully mobile scenarios. The aim is to enable rapid market introduction of new standard-compliant WiMAX equipment and to promote the use of the technology in different sectors.

From IEEE 802.16 to Mobile WiMAX

The IEEE 802.16 standard was originally meant to specify a fixed wireless broadband access technique for point-to-point and point-to-multipoint links. During its development, however, it was decided that mobility support should also be considered.

The WiMAX Forum defines two system profiles based on [1] and [2], called *fixed* and *mobile* system profiles, respectively. Both include mandatory and optional PHY and MAC layer features that are required from all corresponding WiMAX-certified products. Because [1] and [2] specify only the PHY and MAC layers, an end-to-end architecture specification was deemed necessary in order to enable fast growth in manufactured quantities, market share, and interoperability. In response, the WiMAX Forum established the *Network Working Group* (NWG) with the aim of developing an end-to-end network reference model architecture based on IP supporting both fixed and mobile WiMAX (refer to [3] and [4]).

In short, according to the NWG reference model, a WiMAX network is partitioned into three independent architectural components: the user equipment (also referred to as *Customer Premises Equipment* [CPE]), the *Radio Access Network* (RAN, based on IEEE 802.16), and the network providing IP connectivity with the rest of the Internet. Clearly, this model allows a single operator to freely mix and match offerings from different manufacturers for these three parts, at least after interoperable equipment becomes readily available. Furthermore, in principle, each of these components of an operational network can be deployed and managed by different service providers. This scenario makes the network architecture flexible, eases network operation and maintenance, can increase competition under certain conditions, and is conducive to new business models. For example, municipalities can venture jointly with local or national network operators to deploy WiMAX in suburban and rural areas.

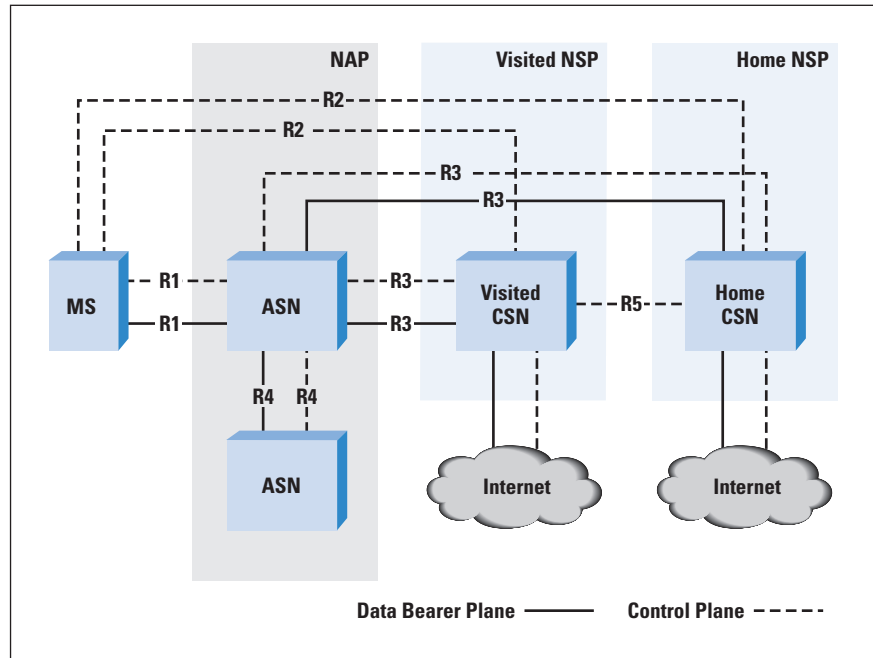
In contrast with earlier wireless data networks^[5], IP is fundamental in a WiMAX network. Indeed, IP currently plays a dominant role in the present state of the telecommunications industry. The premise is that by embracing IP, service providers and equipment manufacturers will face fewer problems when introducing WiMAX into their networks and product portfolios. Moreover, protocols standardized by the *Internet Engineering Task Force* (IETF) are preferred over proprietary solutions and are adopted as extensively as possible in the reference model.

Mobile WiMAX Network Reference Model

The WiMAX Forum NWG network reference model defines three basic architectural entities: the *Mobile Station* (MS), the *Access Service Network* (ASN), and the *Connectivity Service Network* (CSN). The role of the MS is to provide user access to the WiMAX network. The ASN is the Radio Access Network and is formed by numerous *Base Stations* (BSs) and *ASN Gateways* (ASN-GWs), managed by a *Network Access Provider* (NAP). CSN is the network entity providing IP connectivity to the WiMAX radio equipment, including all the IP core network functions required for internetworking with the rest of the world. CSNs are maintained by *Network Service Providers* (NSPs).

The ASN and CSN are further broken up into smaller functional entities, which communicate with each other using standardized interfaces called *reference points*. These reference points guarantee that a certain set of protocols and procedures are always supported and can function irrespective of the underlying hardware. The currently defined reference points are used for different control and management purposes, as well as for data bearing between the network entities. Figure 1 illustrates the network reference model and the main reference points.

Figure 1: WiMAX Forum NWG Network Reference Model



The reference points are defined as follows in [3]: Reference point R1 consists of protocols and procedures compliant to [1], [2], and [6]. R1 implements the specifications of the air interface between the MS and the BS. R2, an interface between the MS and a CSN, is used solely for management purposes, including mobility management. R3 serves the same purpose between an ASN and a CSN, and R4 is used for micromobility management between two ASNs. R5 enables interworking between two CSNs for macromobility management.

In addition to reference points R1–R5, another three intra-ASN reference points are defined (not illustrated in Figure 1). R6, which consists of a set of control- and bearer-plane protocols for BS and ASN-GW communication, controls the data path and MS mobility events between these two ASN entities. R7 is an optional set of protocols used for coordinating R6 functions. Finally, R8 consists of bearer-plane protocols that enable data transfer between the base stations involved in a handover (also called *handoff*).

With respect to mobility, the reference model considers two different scenarios called *ASN-anchored mobility* and *CSN-anchored mobility*. ASN-anchored mobility (or intra-ASN mobility, or micromobility) management is employed when MS handovers occur from one BS to another, and both are controlled by the same ASN-GW. On the other hand, CSN-anchored mobility (or inter-ASN mobility, or macromobility) management is employed when MS movement dictates a handover from the currently serving BS to another one that is in a different subnetwork, controlled by a different ASN-GW. In the ASN-anchored case, handovers are managed solely by the MS and the ASN. In the CSN-anchored case, both ASN and CSN entities are engaged in mobility management.

Typically, ASN-anchored mobility procedures take precedence and CSN-anchored mobility management is employed only if necessary. Because ASN-anchored mobility takes place inside a single ASN, it does not change the MS network layer (IP) configuration. Three different functions are specified for ASN-anchored mobility management, all considered peer-to-peer interactions between different architectural entities:

- The *handoff* (HO) function controls the handover decision operation and handover signaling. The HO function supports mobile- and network-initiated handovers and, additionally, it may support *Fast Base Station Switching* (FBSS) or *Macro Diversity Handover* (MDHO)^[2].
- The *Data Path* (DP) function manages the data path setup and data packet transmission between two functional entities.
- The context function addresses the exchanges required in order to retrieve or set up any state in the network elements.

On the other hand, when MS movement necessitates CSN-anchored mobility management, the MS IP layer configuration changes as a result of the handover. In this case, mobility management is based on *Mobile IPv4* (MIPv4)^[7] or *Mobile IPv6* (MIPv6)^[8], if the MS supports it. Alternatively, the reference model adopts *Proxy MIP* (PMIP)^[9] to handle the handover. In PMIP, the MIP function is moved from the MS to a network instance called a *PMIPv4 client*, which takes care of all MIP signaling on behalf of the MS. Support for PMIP is specified only for MIPv4 in [3] and [4]. Note that in a handover from one ASN to another, MIP is used to complement ASN-anchored mobility management. The latter is still necessary to control the link-layer handover procedures. That is, after the micromobility handover is successfully completed, MIP independently takes care of the macromobility handover, that is, establishes communication paths between the new ASN-GW and the CSN. CSN-anchored mobility handovers are always network-initiated.

By embracing IETF protocols and providing an end-to-end architecture with independent functional entities, the WiMAX Forum NWG network reference model provides a clear framework for the application developers to work in. The model provides only operational requirements and does not prescribe particular technical solutions to realize them, allowing for proprietary yet standards-compliant implementations and enabling technical competition between different manufacturers.

Before examining mobility support in WiMAX, we review the basics of the IEEE 802.16 PHY and MAC layers.

OFDM and OFDMA

IEEE 802.16 and thus WiMAX adopted *Orthogonal Frequency Division Multiplexing* (OFDM), a multicarrier modulation scheme, as its PHY layer. In OFDM, the available bandwidth is divided into several parallel orthogonal subcarriers with lower bandwidth. A wideband channel is defined as a group of adjacent narrowband channels: a high-bitrate data stream is divided into these subcarriers and multiple narrowband data streams are transmitted over the air. Because the data symbol duration is inversely proportional to bitrate, the transmitted symbol duration is increased and the level of *Inter-Symbol Interference* (ISI) can be reduced. ISI is caused by multipath propagation in the wireless communication medium, where the transmitted data symbols can arrive at the receiver through different propagation routes because of reflections from buildings in urban areas and from hills and trees in rural areas. OFDM also uses guard intervals between successive data symbols and cyclic prefixes in order to decrease the effect of ISI even more.

One reason for the wide adoption of OFDM in modern broadband communication systems is its hardware implementation simplicity. OFDM signals can be formed and processed using *Inverse Fast Fourier Transform* (IFFT) and *Fast Fourier Transform* (FFT), at the transmitter and receiver, respectively, and both transforms can be implemented directly in hardware for higher performance. OFDM bodes well for mobile broadband systems through frequency diversity and adaptivity in both modulation and channel coding. By using *Adaptive Modulation and Coding* (AMC), the end-to-end quality deterioration due to the excess delays and deep fading conditions caused by mobility can be prevented, or at least diminished.

OFDM can also be used as a multiaccess scheme by having subcarriers grouped into subchannels, which can be assigned to different users contending for the data link. Each subchannel can contain a different number of subcarriers, and by altering the subcarrier group sizes and observing the channel conditions, it is possible to use differentiation in the channel allocation for different users.

This technique of using OFDM as a multiaccess scheme is called *Orthogonal Frequency Division Multiple Access* (OFDMA). Mobile WiMAX uses OFDMA as its PHY layer instead of plain OFDM, and subchannelization to both uplink and downlink transmissions is possible.

In OFDMA, the subcarriers assigned to subchannels can be either concurrent or taken from different regions of the total bandwidth. Both of these allocation schemes have advantages. When subcarriers assigned to one subchannel are distributed over the available bandwidth, frequency diversity can be attained. In mobile systems this diversity is advantageous because it can be used to make the transmission link more resistant against fast fading. A subchannelization scheme based on dispersed subcarrier allocation to subchannels, called *Partial Usage of Subcarriers* (PUSC), is mandatory in all mobile WiMAX implementations.

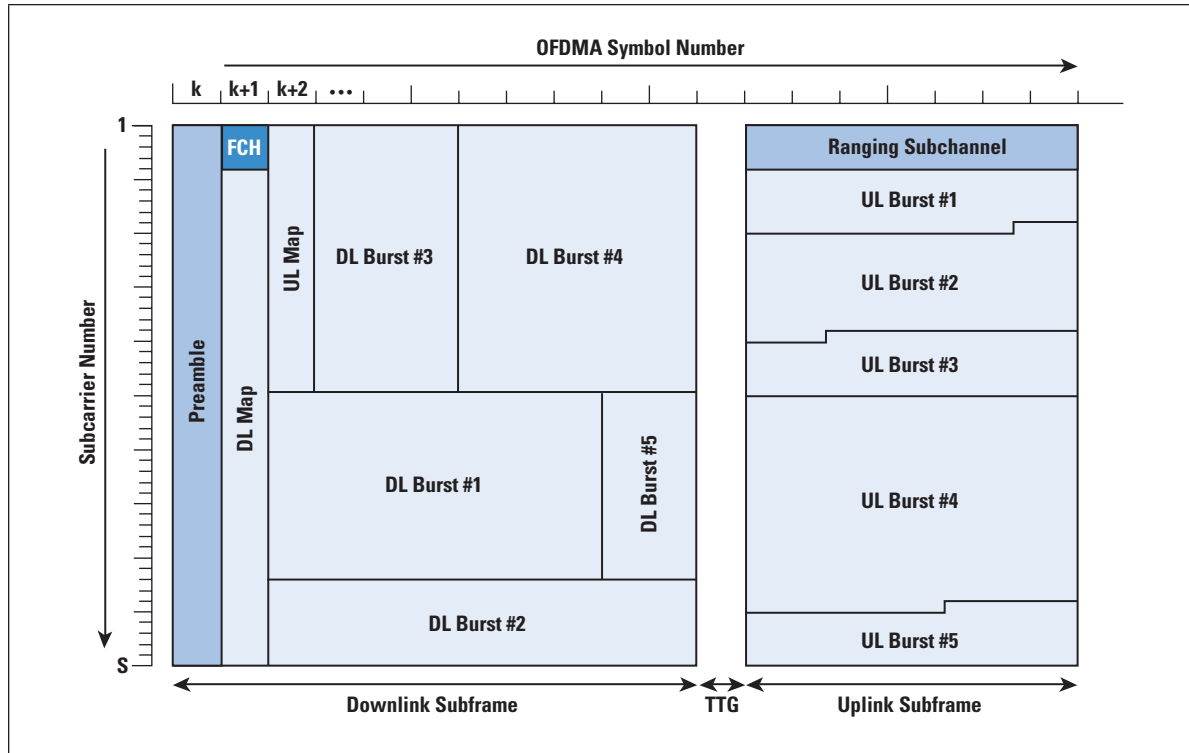
WiMAX systems can use *Time-Division Duplexing* (TDD) or *Frequency-Division Duplexing* (FDD) when allocating air interface resources to users. In TDD, the uplink and downlink transmissions are done over the same carrier frequencies and the separation between the transmission directions is done by assigning time slots, in which the transmission to one direction at a time is scheduled. In FDD, uplink and downlink transmissions are done simultaneously over different carrier frequencies.

Commonly used in mobile WiMAX equipment, TDD allows more flexible sharing of the available bandwidth between the uplink and downlink transmissions. On balance, TDD requires synchronization between multiple adjacent base stations so that transmissions in neighboring cells do not interfere with each other. A TDD frame (Figure 2) is divided into two subframes: first comes a downlink frame and after a short guard interval, called the *Transmit/Receive Transition Gap* (TTG), an uplink frame follows in the same frequency band. Each downlink subframe starts with a preamble, which is used for synchronization and channel estimation. To enhance tolerance against mobility-inflicted channel impairments, WiMAX allows optional support for a more frequent preamble repetition during transmission. In the uplink, short preambles, also called *midambles*, can be used after 8, 16, or 32 OFDM symbols, and in the downlink, short preambles in front of every data burst can be used. After the preamble comes a *Frame Control Header* (FCH), which consists of uplink and downlink *Media Access Protocol* (MAP) messages, which inform users about their transmission parameters.

Flexible data multiplexing from different users into one OFDM or OFDMA frame is also supported, as illustrated in Figure 2. Both uplink and downlink subframes can include data bursts of different types from multiple users, and they can be of variable length.

A small portion of the uplink subframe is reserved for transmission parameter adjustment and bandwidth request purposes. Moreover, small amounts of user data can be sent in this portion of the uplink subframe. The total OFDM frame size can range between 2.5 and 20 ms, but the initially supported frame size in present WiMAX equipment is 5 ms.

Figure 2: An example of a WiMAX OFDMA Frame

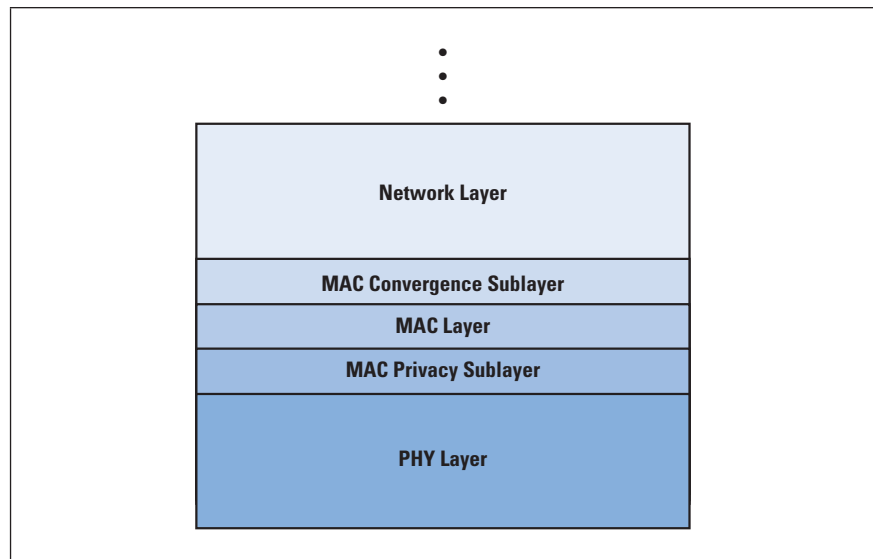


Media Access Control

The MAC layer is primarily an adaptation layer between the PHY layer and the upper layers. Its most important task, when transmitting data, is to receive *MAC Service Data Units* (MSDUs) from the layer above, aggregate and encapsulate them into *MAC Protocol Data Units* (MPDUs), and pass them down to the OFDM or OFDMA PHY layer for transmission. When data is received, the MAC layer takes MPDUs from the PHY layer, decapsulates and reorganizes them into MSDUs, and passes them on to the upper-layer protocols.

An additional layer between the MAC and upper protocol layers called the *Convergence Sublayer* (CS) is also defined in [1] and [2] and illustrated in Figure 3. For the upper layers, CS functions as an interface to the MAC layer. Even though in principle a CS is presented for a variety of different protocols, currently [3] and [4] support CS only for IP and Ethernet. Other protocols can, of course, use these CSs through encapsulation. The CS may also support upper-protocol header compression.

Figure 3: WiMAX Protocol Stack



Similarly with the PHY layer, shown in Figure 3, the MAC layer allows flexible allocation of transmission capacity to different users. Variably sized MPDUs from different flows can be included into one data burst before being handed over to the PHY layer for transmission. Multiple small MSDUs can be aggregated into one MPDU and, conversely, one big MSDU can be fragmented into multiple small ones in order to further enhance system performance. For example, by bundling up several MPDUs or MSDUs, the PHY and MAC layer header overheads, respectively, can be reduced.

It is important to remember that the BS MAC layer manages bandwidth allocation for both uplink and downlink transmissions. The BS assigns bandwidth for the downlink transmission according to incoming network traffic. For the uplink transmission, bandwidth is allocated based on the requests received from the MS. Because basically all connections are controlled by the BS, QoS can be efficiently implemented into WiMAX equipment. Currently, the MAC layer of a mobile WiMAX BS should include support for five different QoS classes, briefly summarized in Table 1.

Table 1: Mobile WiMAX QoS Classes

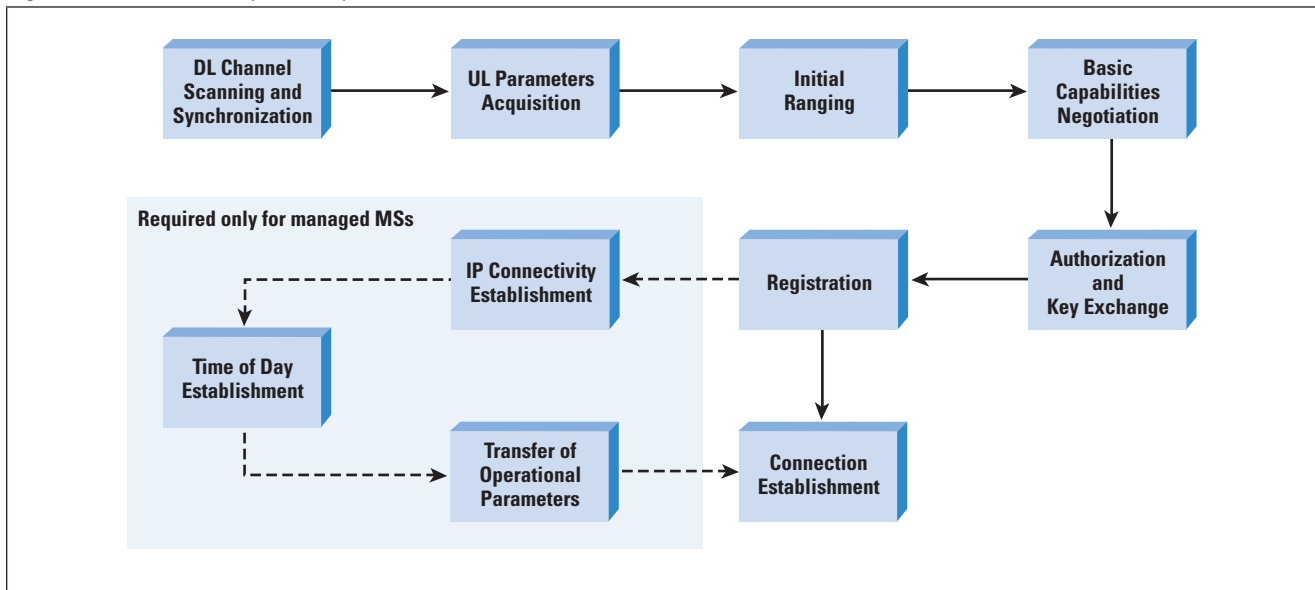
QoS Class	Supported Service	Example Application
Unsolicited Grant Services (UGS)	Latency- and jitter-sensitive applications with fixed-size data packets at Constant Bitrate (CBR)	Voice over IP (VoIP) without silence suppression
Real-Time Variable Rate (RT-VR)	Real-time applications with variable-size data packet bursts	Video and audio streaming
Non-Real-Time Polling Services (nrtPS)	Delay-tolerant applications with variable-size data packets and guaranteed bitrate demands	File transfers
Extended Real-Time Variable Rate (ERT-VR)	Real-time applications with Variable Bitrate (VBR) data streams and guaranteed bitrate and delay demands	VoIP with silence suppression
Best Effort (BE)	Data streams with no minimum service-level demands	Web browsing, instant messaging, and data transfer

Prior to any data transmission over a WiMAX link, the MS and the BS must form a unidirectional connection between their respective MAC layers. A unique identifier, called *Connection Identifier (CID)*, is assigned to each uplink and downlink connection pair. The CID serves as a temporary address for the transmitted data packets over the WiMAX link. Another identifier, called *Service Flow Identifier (SFID)*, is assigned by the BS to unidirectional packet flows with the same QoS parameters, that is, service flows. The BS also handles the mapping of SFIDs to CIDs in the QoS control process. Note that the MAC layer incorporates sophisticated power-management techniques and robust, state-of-the-art security features, but these features are out of scope for this article.

Network Entry and Reentry

Figure 4 illustrates the basic steps that every MS must go through when entering or reentering a WiMAX network. First, a MS scans the downlink channel and synchronizes with the BS, after which the MS acquires the transmit parameters for the uplink transmission from the BS *Uplink Channel Descriptor (UCD)* message and performs initial ranging, hence acquiring the correct timing offset and power adjustments. A MS extracts an initial ranging-interval time slot from an uplink MAP message. If a MS cannot complete the initial ranging successfully, it must start scanning for a new downlink channel.

Figure 4: Network Entry/Reentry Procedure



The basic capabilities negotiation process starts when the MS sends a message containing its capabilities to the BS; the BS responds with a message containing the capabilities it has in common with the MS. If *Privacy Key Management (PKM)* is enabled at both the MS and the BS, the next step is to perform the authorization and key-exchange procedure, so that the MS can register with the network. The BS sends back a registration response message that contains the secondary management CID, if the MS is managed.

After a managed MS obtains this secondary management CID, it becomes “manageable.” The successful reception of the registration response message is a prerequisite for any MS in order to be able to transmit to and receive from the network.

When a managed MS enters the network, the next step is to establish IP connectivity by using the assigned secondary management connection and by either invoking the *Dynamic Host Configuration Protocol* (DHCP)^[10] or DHCPv6^[11], or using the IPv6 stateless address autoconfiguration^[12], depending on the information provided by the BS registration response message. If the MS uses MIPv4 or MIPv6, it can secure its address by using the secondary management connection with MIP. The establishment of IP connectivity and time of day, as well as the transfer of the operational parameters, are needed only for managed MSs. These parameters can be managed with IP management messages through a secondary management connection, for example, by using the DHCP, *Trivial File Transfer Protocol* (TFTP)^[13], or *Simple Network Management Protocol* (SNMP)^[14]. These additional steps during network entry are necessary for the operation of the IP management protocols.

If DHCP is used to establish IP connectivity, a managed MS must also establish the time of day so that the management system can time-stamp certain events. Both the MS and the BS must be set at the same time of day, with an accuracy of the nearest second. The time of day is retrieved using the secondary management connection with the *Time Protocol*^[15]. The current time is formed by combining the time retrieved from the server with the time offset extracted from the DHCP reply message. Although the time of day is not needed for the registration to complete successfully, it is required in order to keep the connection operational. Finally, the managed MS must acquire its operational parameters with TFTP.

After a managed MS has obtained its operational parameters, or after an unmanaged MS has registered with the network, the MS preprovisioned service-flow connections are established.

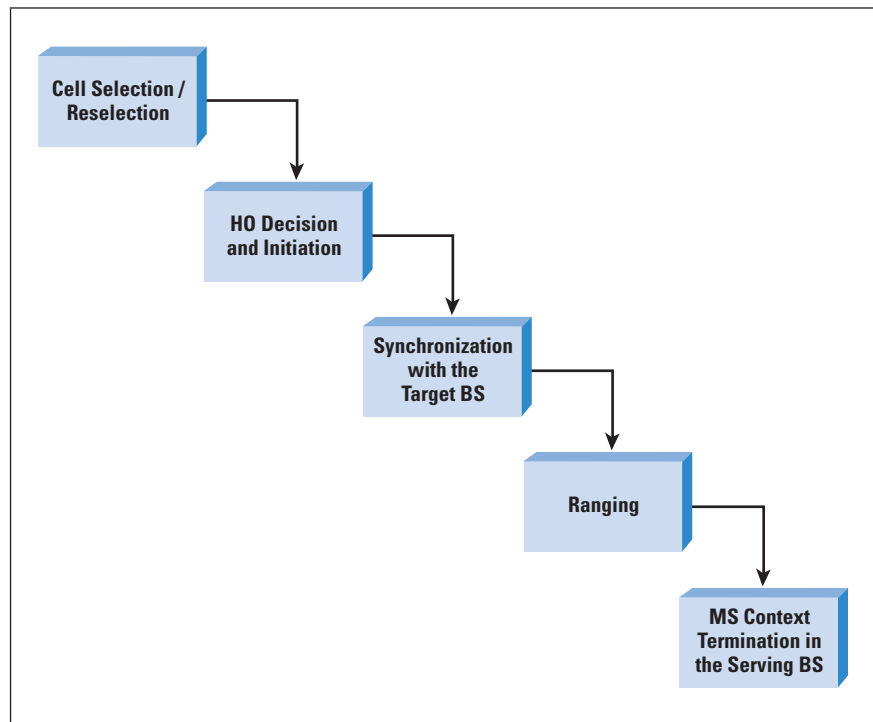
Mobility Support

As discussed previously, IEEE 802.16e introduced mobility support, defining an OFDMA PHY layer and signaling mechanisms to enable location and mobility management, paving the way for mobile WiMAX. The WiMAX Forum details four mobility scenarios in addition to the fixed WiMAX scenario. In the nomadic and portable mobility scenarios, the point of attachment of a fixed *Subscriber Station* (SS) can change. The simple mobility scenario allows MSs to roam within the coverage area with speeds up to 60 km/h, but handovers may cause connection interruptions of up to 1 second. In the so-called *full-mobility scenario*, the MS speed can be as much as 120 km/h, and transparent handovers are supported. This last scenario is what many might consider as the real mobile WiMAX scenario, but all five scenarios are “standards-compliant.”

Although three different types of handovers are defined in [2], *Hard Handover* (HHO), *Macro Diversity Handover* (MDHO), and *Fast Base Station Switching* (FBSS), only HHO is mandatory for all mobile WiMAX equipment. This type of handover is often referred to as a *break-before-make handover*: first, the MS disconnects from the serving BS and then connects to the target BS. Because of the short disconnection period, packets may be lost; HHO is less sophisticated than either MDHO or FBSS and may be inappropriate for some applications. The MS must also register with the target BS and reauthenticate with the network, typically meaning further delays before actual data exchange can (re)start. If multiple handover types are supported and enabled, the BS decides which type should take precedence over the other. MDHO and FBSS are enabled or disabled during the registration of the MS with the BS.

Figure 5 illustrates the five stages of a successful HHO in mobile WiMAX. The first stage is to select the target BS cell based on information about the network topology surrounding the serving BS through periodically broadcasted neighbor advertisements. The advertisements include the same information on the serving BS neighbors that the *Downlink Channel Descriptor* (DCD) and *Uplink Channel Descriptor* (UCD) messages of the neighboring BSs would include. For example, a neighbor advertisement message includes channel information of the neighboring BSs so that the MS can synchronize with them and perform scanning operations to evaluate their suitability as potential targets for a HO.

Figure 5: The Five Phases of a Successful HHO



The second phase is to make the actual decision to initiate the handover procedure, when a certain network (say, congestion in the serving cell requires load balancing) or channel condition threshold (for example, low received *Signal-to-Interference + Noise Ratio* [SINR] in the current cell) is crossed. The actual decision to start the message exchange for the MS to migrate from the radio interface of the serving BS to the radio interface of another BS can be made by the MS, BS, or the network. In the third phase, the MS synchronizes with the downlink transmission of the target BS and obtains the transmission parameters for the downlink and the uplink. The time consumed to perform the synchronization procedure depends on the amount of information the MS received about the target BS in the neighbor advertisement messages prior to the handover. The average synchronization latency without previously acquired information about the target BS ranges from two to three frame cycles, or approximately 4 to 40 ms depending on the OFDMA frame duration used in the system. The more extensive the channel parameter list received in the neighbor advertisement messages prior to the handover, the shorter the time to achieve the synchronization.

After synchronizing, the MS and the target BS initiate the ranging procedure. During this fourth step in HHO, MS and BS exchange the required information so that the MS can reenter the network. The target BS can request information about the MS from the (previously) serving BS and other network entities. Again, the more information made available to the target BS, the shorter the time to reenter the network, because the target BS may skip some steps from the network (re)entry procedure described earlier. In short, sharing context information before the actual handover optimizes the handover procedure and decreases its latency. In the last step of a HHO, the MS context at the serving BS is terminated and resources are released.

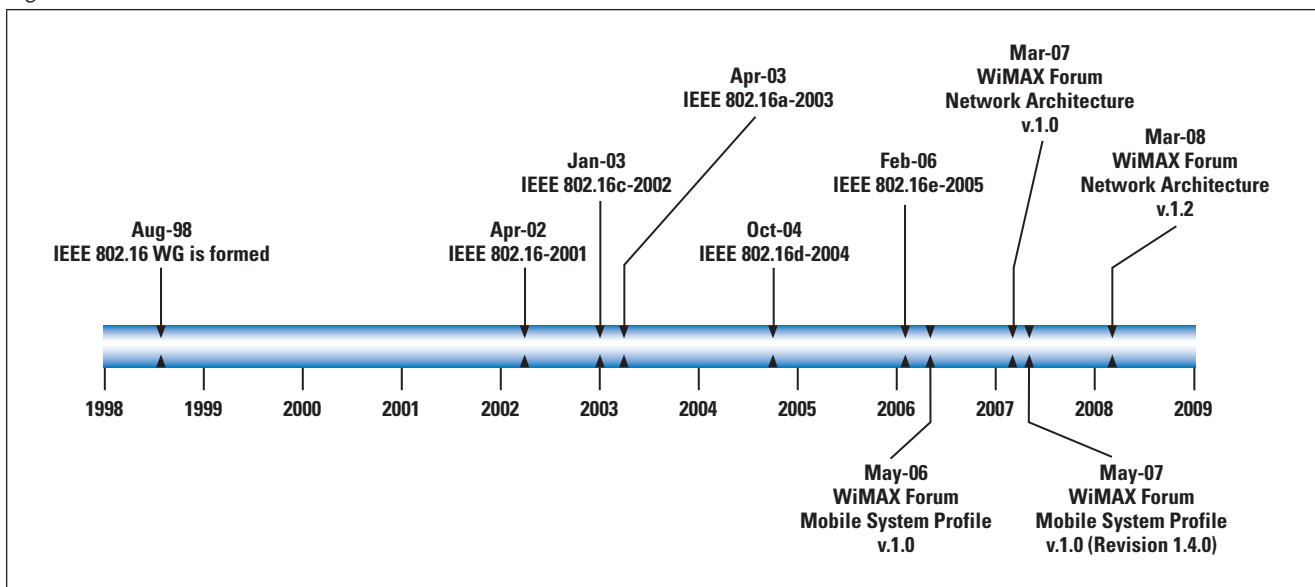
If MDHO and FBSS are supported, the following stages, in addition to those already described in the HHO procedure, must be performed: (a) decision to enable MDHO or FBSS, (b) diversity set update, and (c) anchor BS selection. In macrodiversity communications the MS maintains a connection to one or more serving BSs simultaneously, enabling soft or make-before-break handovers. In [2], the transition of the MS from the air interface of one or more serving BSs to the air interface of one or more target BSs is referred to as a MDHO. The MS and the BS both maintain a list called the *diversity set*, which includes all serving BSs involved in the MDHO communication. The MS maintains both uplink and downlink unicast connections to all the BSs in the diversity set, and one of the serving BSs is defined as the anchor BS. Note that all BSs involved in the diversity set use the same set of CIDs for the connections established between the MS and the serving BSs.

In FBSS, the MS transmits to and receives data from a single serving BS during any frame period. The BS, to which the MS has the connection to at any given frame, is called the *anchor BS*. The MS maintains a diversity set, which includes all active BSs in its range, and can change its anchor BS on a frame-by-frame basis, based on certain criteria. The transition from the serving anchor BS to the target anchor BS in FBSS is done without invocation of the normal handover procedure, and only the anchor BS update procedure is needed. After all, the MS has collected all required information about all BSs during the diversity set update ranging procedures.

Mobile WiMAX vs. HSPA

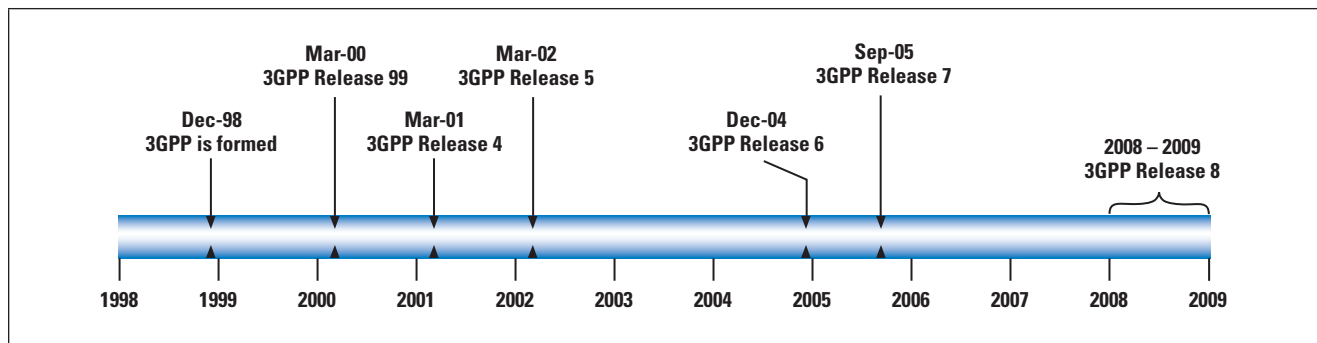
Mobile WiMAX and *High-Speed Packet Access* (HSPA) are expected to be the two major contestants in the rapidly growing wireless broadband market. The two, however, come from different origins. Figure 6 summarizes the evolution toward mobile WiMAX. It all started with the establishment in August 1998 of the IEEE 802.16 working group, which published its first standard (IEEE 802.16-2001) in April 2002. This first version defines a single carrier system operating in the 10- to 66-GHz frequency band and only under *line-of-sight* (LOS) conditions. The IEEE 802.16c-2002 amendment detailed system profiles for the original standard based on the 10- to 66-GHz frequency band. IEEE 802.16a-2003 introduced support for 2- to 11-GHz frequencies and *non-line of sight* (NLOS) operation, and adopted the use of OFDM and OFDMA. IEEE 802.16d-2004^[1] consolidated all these previous versions and amendments in a single document, and further enhanced the system. Fixed WiMAX is based on IEEE 802.16d-2004, [3], and [4]. Mobile WiMAX is based on the IEEE 802.16e-2005 amendment^[2], which introduced mobility support, as well on [3] and [4].

Figure 6: The Road Toward Mobile WiMAX



HSPA is a set of technological enhancements to the already widely deployed *Wideband Code Division Multiple Access* (WCDMA) cellular networks defined by the *Third Generation Partnership Project* (3GPP). Figure 7 illustrates the WCDMA specification evolution. The origins of HSPA can be traced in the foundation of 3GPP in December 1998. The original aim of 3GPP was to develop a third-generation WCDMA system, and in the process, HSPA was introduced. In March 2000, Release 99, the original standard specifying the WCDMA system, was published. A year later, the first enhancements were published in Release 4, which introduced, among others, an IP-based core network. Release 5 introduced *High-Speed Downlink Packet Access* (HSDPA) and defined the *3GPP IP Multimedia Subsystem* (IMS). *High-Speed Uplink Packet Access* (HSUPA) and some further improvements to HSDPA were defined in Release 6 (December 2004). Release 7 further enhanced QoS support and defined mechanisms to decrease network latency. Release 8 is expected to be published in 2008, and it will include specifications for the next step, called 3GPP *Long-Term Evolution* (LTE). LTE is meant to deliver maximum cell throughputs an order of magnitude larger than HSPA.

Figure 7: The Evolution of the 3GPP WCDMA Standard



Mobile WiMAX evolved out of a broadband wireless LAN/MAN technology, and vendors currently report that it can deliver maximum cell capacities of 46 and 7 Mbps in downlink and uplink transmissions, respectively. However, mobility management is a later addition and, according to Maravedis, by September 2007 only 12 percent of all deployed *Customer Premises Equipment* (CPE) was IEEE 802.16e-2005-compliant^[16]. On the other hand, HSPA is based on a solid foundation of mobility management techniques with wide deployment in cellular networks around the globe, but can currently deliver maximum cell throughputs of only 14.4 and 5.8 Mbps in downlink and uplink transmissions, respectively.

Either commercial or trial networks of both technologies have already been implemented all over the world. However, according to the *Global Mobile Suppliers Association* (GSA), HSPA networks have yet to be deployed in China and India, both of which are large and rapidly growing market areas for wireless communications. According to Maravedis, both India and China have at least WiMAX trial deployments in place.

As mentioned already, the vast majority of current WiMAX deployments do not support mobility. Up to now, fixed WiMAX has been used mainly for last-mile broadband connectivity for sparsely populated rural areas. The largest commercial IEEE 802.16e-2005-compliant system is currently the *Wireless Broadband (WiBro)*^[17] network in South Korea, which supports simple mobility up to 60 km/h. Even though WiMAX and WiBro are both based on the same standards, WiBro was developed by the South Korean telecommunications industry before the WiMAX Forum adopted mobility support for its system profiles. WiMAX and WiBro are often cited as separate technologies, even though cooperation is in place in order to assure interoperability between the two.

Summary

In this article we presented an overview of mobile WiMAX, a much-heralded technology for next-generation mobile broadband networks; mobile WiMAX is an intricate system. We introduced WiMAX and the role of the WiMAX Forum, and summarized the important points of the WiMAX network reference model and the PHY and MAC layers. We addressed mobility support, but not the security aspects. Finally, we briefly compared WiMAX with HSPA, presenting their respective evolutions and illustrating their worldwide deployments. We hope that this article will serve as a valuable primer, and we highly recommend that those interested in the mobile WiMAX technology check the bibliography.

Bibliography

- [1] IEEE 802.16 Working Group, “IEEE Standard for Local and Metropolitan Area Networks, Part 16: Air Interface for Fixed Broadband Wireless Access Systems,” IEEE Standard 802.16-2004, October 2004.
- [2] IEEE 802.16 Working Group, “IEEE Standard for Local and Metropolitan Area Networks, Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems, Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands,” IEEE Standard 802.16e-2005, February 2006.
- [3] WiMAX Forum Network Working Group, “WiMAX Forum Network Architecture—Stage 2: Architecture Tenets, Reference Model and Reference Points—Release 1, Version 1.2,” WiMAX Forum, January 2008.
- [4] WiMAX Forum Network Working Group, “WiMAX Forum Network Architecture—Stage 3: Detailed Protocols and Procedures—Release 1, Version 1.2,” WiMAX Forum, January 2008.

- [5] K. Pentikousis, “Wireless Data Networks,” *The Internet Protocol Journal*, Volume 8, No. 1, March 2005, pp. 6–14.
- [6] IEEE 802.16 Working Group, “IEEE Standard for Local and Metropolitan Area Networks. Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems. Amendment 3: Management Plane Procedures and Services,” IEEE Standard 802.16g-2007, December 2007.
- [7] C. Perkins (Ed.), “IP Mobility Support for IPv4,” RFC 3344, August 2002.
- [8] D. Johnson, C. Perkins, and J. Arkko, “Mobility Support in IPv6,” RFC 3775, June 2004.
- [9] K. Leung, G. Domemety, P. Yegani, and K. Chowdhury, “WiMAX Forum/3GPP2 Proxy Mobile IPv4,” Internet-Draft, Work in Progress.
- [10] R. Droms, “Dynamic Host Configuration Protocol,” RFC 2131, March 1997.
- [11] R. Droms (Ed.), J. Bound, B. Volz, T. Lemon, C. Perkins, and M. Carney “Dynamic Host Configuration Protocol for IPv6 (DHCPv6),” RFC 3315, July 2003.
- [12] S. Thomson and T. Narten, “IPv6 Stateless Address Autoconfiguration,” RFC 2462, December 1998.
- [13] K. Sollins, “The TFTP Protocol (Revision 2),” RFC 1350, July 1992.
- [14] J. Case, M. Fedor, M. Schoffstall, and J. Davin “A Simple Network Management Protocol (SNMP),” RFC 1157, May 1990.
- [15] J. Postel and K. Harrenstien, “Time Protocol,” RFC 868, May 1983.
- [16] K. Pentikousis, J. Pinola, E. Piri, F. Fitzek, T. Nissilä, and I. Harjula, “Empirical Evaluation of VoIP Aggregation over a Fixed WiMAX Testbed,” Proceedings of The 4th International Conference on Testbeds and Research Infrastructures for the Development of Networks & Communities (TRIDENTCOM), 18–20 March, 2008, Innsbruck, Austria.
- [17] Telecommunications Technology Association, “Specifications for 2.3GHz Band Portable Internet (WiBro™) Service,” TTA Standard TTAS.KO-06.0082/R1, December 2005.

JARNO PINOLA received his M.Sc. from the University of Oulu, Oulu, Finland, in Spring 2008. During his studies, he specialized in telecommunication systems and wrote his Master's Thesis on mobility management issues in wireless broadband systems. Currently he is working as a Research Scientist at VTT Technical Research Centre of Finland in Oulu, Finland. He can be contacted via e-mail at: [**jarno.pinola@vtt.fi**](mailto:jarno.pinola@vtt.fi)

KOSTAS PENTIKOUSIS studied computer science at Aristotle University of Thessaloniki, Thessaloniki, Greece (B.Sc. 1996), and Stony Brook University, Stony Brook, New York, USA (M.Sc. 2000, Ph.D. 2004). He is a tenured Senior Research Scientist at VTT Technical Research Centre of Finland, in Oulu, Finland. He has published internationally in several areas, including mobile computing (mobility triggers, multiaccess, media-independent handovers, and energy consumption); transport protocols; applications; network traffic measurements and analysis; and simulation and modeling. Visit [**http://ipv6.willab.fi/kostas**](http://ipv6.willab.fi/kostas) for more information and contact details.