This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: https://doi.org/10.1007/s11235-009-9244-z
Postprint of: Mrugalski T., Wozniak J., Analysis of IPv6 handovers in IEEE 802.16 environment, TELECOMMUNICATION SYSTEMS, Vol. 45, Iss. 2 (2010), pp. 191–204

Analysis of IPv6 Handovers in IEEE 802.16 Environment

Tomasz Mrugalski, tomasz.mrugalski@gmail.com

Jozef Wozniak, jowoz@eti.pg.gda.pl

Abstract The second generation of WiMAX solutions, based on IEEE 802.16-2005 standard, offers limited mobility support. Unfortunately, after quickly changing the point of attachment on the WiMAX data link layer (DLL), very slow and inefficient IPv6 reconfiguration takes place. Delays introduced by automatic configuration (DHCPv6 and IPv6 protocols) and Mobile IPv6 can easily diminish or even render useless all benefits gained using the efficient handover performed on DLL. As handover is a crucial process in mobile cellular environments, reasons behind delays introduced by IPv6 layer mechanisms have to be analyzed and appropriate countermeasures applied.

In order to analyse influence of different factors on the handover delay a simulation environment modelling the full handover procedure in a WiMAX environment has been developed. It allows simulation and analysis of various mobility related issues, offering support for multiple base stations with groups of subscribers, both fixed and mobile, with various mobility models. Also support for tight integration with higher layers (IPv6, DHCPv6, and Mobile IPv6) is fully implemented. All stages of full IPv6 handover in IEEE 802.16 environment, focusing on major reasons of reconfiguration delays are described.

The paper presents components, functional requirements and architecture of the simulation environment, together with example simulation results. The obtained results clearly show that most significant delays are caused by the IPv6 layer. The areas of improvement in several autoconfiguration mechanisms are identified. Proposals include novel use of DHCPv6 relays for remote configuration, solving DAD delays, limiting Binding Update procedure in Mobile IPv6, and configuring routing through DHCPv6 communication.

A universal metric for assessing impact of every stage on handover efficiency is also defined. Several proposed improvements to the IPv6 handover process are evaluated. Discussion regarding possible generalization of best improvement proposals and remarks on further research areas conclude this paper.

Keywords mobility, DHCPv6, autoconfiguration, handover, IEEE 802.16, IPv6, WiMAX, simulation environment, analysis

Gdansk University of Technology, 11/12 Narutowicza Str
80-952 Gdansk, Poland

I. Mobility – WiMAX and IPv6 perspective

The amount of digital information created, stored, retrieved and transmitted is increasing rapidly. At the same time, portable and different handheld devices are becoming smaller and more powerful. As with all electronic equipment, also mobile devices are affected by Moore's law, which states that the computing power of devices doubles every 18 months. With wireless technologies reaching their maturity, more users are expected to use mobile devices. As a direct result of both trends, users demand transmission and reception of digital data, and the popularity of various mobile oriented multimedia applications, like video on demand or VoIP is growing. At the same time, due to miniaturization and advancements in wireless electronics, new services enabling users to perform mobile computing are gaining significant advantage. From the network point of view, two requirements – delivery of large amounts of data and mobility support – are very hard to meet at the same time. That is because changing a point of attachment to the network by a mobile station is usually complicated.

In order to solve the aforementioned problem a new broadband technology, namely IEEE 802.16, has been developed. The IEEE 802.16 standard, also known under its commercial name WiMAX[1], defines mechanisms which allow Subscriber Stations (SSs) to communicate with Base Stations (BSs). Thanks to the use of advanced radio access technologies and smart bandwidth management, significant improvements have been made in transmission ranges (up to 40-50km) as well as in available throughput (up to 70Mbps). Lack of mobility support in the initial IEEE 802.16-2004 specification was solved in early 2006, when in the paper [2] was published. Numerous mechanisms supporting subscriber mobility were introduced, like Neighbor Advertisement, Scanning and Handover.

After performing data link handover and network reentry in a new location, an IPv6 node, working on top of the WiMAX SS stack, is required to make handover in IPv6 and higher layers. After handover (also after power-up or power conservation wakeup) every IPv6 node is required to confirm its old or

[1] WiMAX logo can be used by vendors, whose equipment pass specific conformance and interoperability tests.

obtain a new address and configuration parameters. That can be arranged using a stateless [5] or stateful [6] automatic configuration (often referred to as autoconfiguration) procedure. Since stateless mode does not provide means of configuring any parameters beside those regarding address and routing, it is generally agreed that stateful configuration should be used in any bigger network. See reference [13] for a detailed discussion regarding this topic. After the full configuration is completed, a given node informs its corresponding nodes[2] (CN) and the home agent (HA) about its new location, according to [7]. This procedure concludes the handover and the node becomes fully operational in its new location, regains its full communication capability, and can continue its communication activities.

Unfortunately, some IPv6 protocols, namely DAD and DHCPv6 were not optimized taking into account mobility support and fast handover purposes. Therefore delays introduced by each of those protocols impacts handover delays significantly. To analyze mobility induced delays in IPv6 over WiMAX, a new simulation environment has been developed. It provides support for multiple base stations with multiple subscribers, both fixed and mobile. Also support for close integration with higher layers (IPv6, DHCPv6, and Mobile IPv6) as well as several mobility models is implemented. Section IV provides additional insights into this simulation environment.

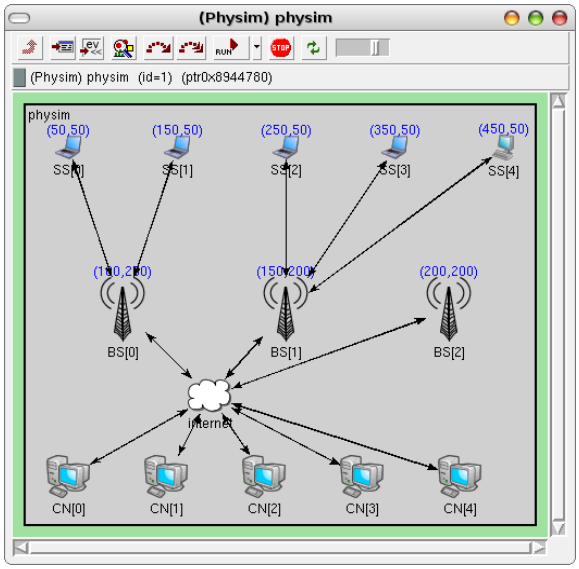

Figure 1: Example of a WiMAX network

In the following section (Section 2), an overview of mobile WiMAX is presented. Mobility delays and lack of communication capability is defined in Section 3. In Section 4 the developed environment with its purpose, functional requirements and architecture is presented. Current state of implementation, testing, and future improvement areas and example simulation results are thoroughly discussed in Section 5.

The paper describes several selected stages of full IPv6 handover in IEEE 802.16 networks, focusing on areas causing essential reconfiguration delays. In order to properly evaluate their influence on the overall handover performance, a metric for assessing the impact of every stage is defined. Several novel improvements to the IPv6 handover process are also proposed and evaluated. Simulation results and conclusions are presented in subsequent sections. A discussion of possible generalization of improvement proposals and further research areas concludes this paper.

II. MOBILE WIMAX AND IPV6

WiMAX is the commercial name of network solutions based on the IEEE 802.16 standard ([1]). Its specification is constantly assessed, tested and improved. Considered as a wireless replacement for DSL lines, WiMAX lives up to the expectations. Offering a range of up to 50km, with throughput up to 70Mbps and good handling of NLOS (non line of sight) scenarios, it seems to be the perfect network solution for suburban and rural areas. One major flaw that was quickly identified is that the original WiMAX specification did not support mobility; hence the most significant improvement made to the specification is support for mobile stations. An example WiMAX network is presented in Fig. 1. Since fixed, IEEE 802.16 based, network solutions started to appear in 2004, intensive work has been undertaken to provide mobility support. As a result in late 2005, a mobility-supporting specification was released [2]. Currently IEEE 802.16 based network solutions are rapidly gaining acceptance, both in academic and telecommunication sectors. Most of the already deployed solutions are fixed, but road maps of major telecom corporations indicate that mobile versions will be commercially available in a very near future [12], [11]. Large-scale deployment of mobile WiMAX solutions is expected to occur within 2 years[3]. In a similar time range, a new version of the currently omnipresent IPv4 protocol, designated next generation IP (or IPv6) will also gain acceptance. Although defined in 1996 ([3]), its rate of adoption has been somewhat slower than initially anticipated. However, there are strong indicators suggesting that massive migration to dual-stack (i.e. supporting both IPv4 and IPv6) or IPv6-only will occur within 2 years. The most important driving force behind this is the United States' Department of Defense (DOD). According to a DOD

[2] CN is a peer node with which a mobile node is communicating; see [5].

[3] As of November 2008, consumer networks have been deployed in Amsterdam and Boston.

Memorandum dated June 9th, 2003 [10], DOD should switch its internal IT structures and all their contractors should also provide all services using IPv6 by the year 2008. Another – and perhaps more fundamental – cause of accelerating migration is uneven distribution of IPv4 addresses. Because the Internet began in the USA, most of the IPv4 user space is allocated for the USA and, to some lesser extent, Europe. From the pool distribution point, the poorest region of the world is Asia. Rapidly developing large countries like India or China require vast amount of address space, a desire that cannot be satisfied with depleting IPv4 pool. The rates of IPv4 allocations are similarly scarce in Africa, but due to a minimal demand, that is considered a minor issue. Adoption of IPv6 technology in several Asian countries is among the highest in the world. According to some sources, exhaustion of the unallocated IPv4 address pool will happen in March 2010 ([16]). It appears reasonable to assume that a significant part of all mobile WiMAX stations will be dual-stack or even IPv6 only. Therefore authors chose IEEE 802.16-2005 as a PHY/MAC layer and IPv6 as a network layer. The move to IPv6 has an interesting aspect, in that IPv6 provides extensive automatic configuration mechanisms, including stateless (i.e. router advertisements [5]) and stateful (dynamic host configuration protocol - DHCPv6, [6]) automatic configuration, Duplicate Address Detection, [5], and Mobile IPv6 [7].

III. Problem statement

Different network layers will produce dramatically different delays during handover. The network and MAC layers in WiMAX have been developed with mobility support and fast processing in mind. Therefore, introduced delays are considered small (reaching a few hundred milliseconds, usually below 100ms). Unfortunately, IPv6 protocol was not designed in this manner. Several steps introduce delays that are very large from the mobility point of view (one second or more). For example, the DHCPv6 server discovery phase takes exactly one second as clients are required to wait for possible responses from other servers, even when one or more servers have already responded (according to DHCPv6 specification).

The handover procedure in the PHY and MAC layers alone is quite complicated. It consists of the following steps:

- **Neighbor Discovery (WiMAX)** – The base station periodically transmits Neighbor Advertisement messages containing information related to other base stations nearby.
- **Scanning (WiMAX)** – A subscriber station, after obtaining information about potential handover targets, performs scanning. This is a temporary detachment from the current base station. During the scanning phase, the subscriber tries to adjust its radio to receive information from other base stations and assess its signal strength and quality. After scanning is complete, the subscriber gains knowledge about the target base station. Using various metrics (signal strength, signal to noise ratio etc.), the subscriber sorts the base station list.
- **Handover (WiMAX)** – A subscriber sends a list of desired target base stations to its serving base station. The serving base station can modify this list and send it back. The actual detachment is signaled by an HO-IND message transmission performed by the subscriber. After this transmission, all connections to the serving base station are removed and the subscriber loses its communication capability.
- **Network Re-entry (WiMAX)** – After adjusting the radio to the target base station's frequency and modulation, subscriber initiates network re-entry. Depending on network configuration and management, this can be a highly optimized re-entry involving the exchange of just four messages. When the target base station has no a priori information about this particular subscriber, full network entry must be performed.

In the optimistic case, when intradomain handover takes place (i.e. handover between two base stations governed by the same operator), the network operator has information about the current subscriber location. It is possible to adjust routing strategies, so the subscriber will be able to send and receive IP datagrams without changing its IP address. This is only an option and even during handover between the same operator's base stations it is not always reasonable (e.g. due to large number of subscribers and thus complicated routing tables may degrade routing efficiency and manageability, therefore operators may want to limit excessive routing modifications.)

Interdomain handover, which is more difficult than intradomain handover, must be analyzed. When a subscriber completes network reentry, the higher layer (i.e. IPv6), must be reconfigured. According to IPv6 standards ([3], [4], [5]) the following steps are necessary:

- **Stateless autoconfiguration (IPv6)** – The station must wait for a Router Advertisement (RA), a message announced periodically by routers. It is possible to request such a message by sending a SOLICIT message. RA will contain information about locally available prefixes and further autoconfiguration instructions. It also allows subscriber routing to be configured properly.

- **Stateful configuration (DHCPv6)** – Only some basic parameters can be configured using stateless configuration, so stateful configuration is required to obtain such parameters as IPv6 addresses, DNS configuration, SIP domains and server. Stateful configuration is performed according to the DHCP for IPv6 protocol (often abbreviated as DHCPv6).
- **Location update (Mobile IPv6)** – After mobile station receives a new IPv6 address and configuration parameters, it must inform its home agent and corresponding nodes of its new point of attachment, and thus of its new address. After this step is complete, mobile station is finally able to resume communication.

It is essential to realize that not all handover steps are causing handover delays. The goal of this work is to identify reasons of major delays during full handover in an IPv6 capable mobile WiMAX stations**.** To assess the impact of every part of the handover procedure on the communication capability, an advanced simulation environment is required.

A full scale handover process described above is long and complicated. During certain steps, like scanning or IPv6 autoconfiguration, the SS is unable to maintain communication. Since real time data transfer, e.g. videoconferences or voice connections, constitute the major types of applications foreseen for WiMAX, interruptions in data transfer are highly problematic, as they lead to deterioration of service quality. To some lesser extent this disadvantage also affects video on demand streaming. Packet drops and delays in that case can easily be counteracted by data buffering in the SS. However, interactive multimedia scenarios (e.g. using VoIP) do not allow extensive data buffering as adding an extra buffering delay becomes a source of service quality degradation.

An approach proposed in this paper consists of two phases. During the first phase, actions or procedures undertaken by network elements are assessed. Procedures that meet the following criteria are good candidates for optimizations:

- **Blocking property** – during such an action communication with CNs is not possible;
- **Action necessity** – means that this action cannot be omitted as it is required by the handover process;
- **Action duration** – a considered action introduces a significant delay i.e., time intervals between communication opportunities.

To conveniently assess and compare radically different handover phases (actions), we propose a metric called Handover Delay. It is expressed in milliseconds and specifies how long an IPv6 node does not have full communication capability due to an analyzed method. X is the metric value, while *HD(.)* stands for its symbolic designation.

$$X = HD(\text{action})\ [\text{ms}] \qquad (1)$$

In general, methods with lower HD are considered "better", as they introduce shorter delays. If a method (action) allows an IPv6 node to communicate immediately, with no handover delay, its HD value is equal to 0ms, so it does not hinder communication in any way and thus requires neither optimization nor improvements.

The second phase of the proposed approach concerns improvements for actions – methods with highest HD metrics. This metric is also used to assess benefits from the proposed improvements. Full handover procedure may also be evaluated in a similar manner. As this paper describes on-going research, improvements are proposed only for selected SS and BS actions.

IV. HANDOVER – THE IEEE 802.16 PERSPECTIVE

During normal operation, the SS is associated with one base station - the Serving Base Station (SBS). Because of degraded signal quality, unsatisfactory quality of service or network operator enforcement, the SS may intend to migrate to another base station - the Target Base Station (TBS). The IEEE 802.16-2005 standard provides various mechanisms that make such a handover possible.

A. Neighbor Advertisements and scanning

The Serving Base Station possesses information regarding neighboring base stations. The list of available target base stations is periodically announced to the associated SSs. That list does not provide a complete set of target base stations. SS and TBS might be on the opposite sides of the area covered by an SBS, so communication between them might be impossible. Also one operator will advertise only its own base stations, although there might be other nearby BSs operated by another service provider. As during neighbor advertisements SS maintains full communication capability, HD metric value of this mechanism equals 0. Therefore this method does not require optimization.

To verify the current status of all BSs within its range, the SS performs scanning. This procedure includes temporary disassociation with SBS and an attempt to receive transmissions from other BSs. This feature allows not associated SSs to receive transmissions and perform signal strength measurement and quality assessment.

The SS cannot maintain communication with its CNs or its BS during the scanning procedure. Duration of the scanning period is variable as a single scanning period length, number of iterations, as well as interval between consecutive scanning periods are requested by the SS and provided by the

BS. Therefore it can adjust scanning procedure parameters to currently supported services. HD metric of this method is difficult to be determined precisely as in theory scanning periods can be arbitrarily long. In practice scanning periods will rarely exceed 20ms. If longer scanning is required, it can be split into several shorter periods. Those facts indicate that scanning does have significant impact on handover delays.

B. Handover

After scanning, the SS has reliable information about possible handover targets. It evaluates all candidates and chooses one i.e., the TBS, as its new network attachment point. The SS announces its decision to its SBS by requesting handover and after receiving confirmation is ready to detach. Optionally, BS can notify the Target Base Station via backbone network, so necessary preparation can be arranged in advance. The exact moment of detachment is chosen by the SS. SS maintains full communication capability until the HO–IND (Handover Indication) message is sent. therefore this step does not require any optimization.

C. Network reentry

After leaving its SBS, the SS adjusts its radio to receive transmissions from the TBS and performs network reentry at the new location. Normal network entry procedure consists of several steps. Both (Anonymous and handover) ranging procedures are intended for radio tuning, and connection identifiers update. *Negotiate Basic Capabilities* is performed to find the common denominator between BS and SS capabilities, like ARQ support and the list of supported modulations. Optional cryptographic protection is achieved via *Privacy Key Management*. *Registration* performed afterwards specifies additional information about IP protocol version used, vendor specific information etc. The final step is to create connections by using *Dynamic Service Flow* procedure. After successful registration the subscriber is logged into the network but is able to communicate only with its base station for control purposes. To send and receive data traffic, service flows must be created. Since service flows are unidirectional, at least one uplink and one downlink flow must be created.

Fortunately the IEEE 802.16 standard provides numerous optimizations in this procedure, designed with mobility in mind. Therefore, the full procedure as described above is usually performed only once, during the first network entry. Further reentries are less time consuming. Exact duration of reentry process is considered one of the most important parameters of a Base Station and is highly dependent on the amount of information about a given SS that this BS has before the actual reentry takes place. Conservative estimation shows that HD metric value can, in some cases, reach even several hundreds of milliseconds. However, there are also solutions with HD less than 100ms. Further optimizations require detailed analysis of the medium access mechanisms and are outside the scope of this paper.

V. DELAYS IN THE IPV6 LAYER

When the SS's data link layer changes the point of attachment, the upper layers must also be reconfigured. After moving to its new location, the IPv6 node must obtain a new address and configure routing parameters. Stateless or stateful autoconfiguration is used to obtain those new parameters. In IP protocol, the address serves two goals: equipment identity (nodes are identified by their addresses) and user location (an address determines the node's location). Therefore a mobile node is required to inform its CNs about a new location. As important parts of this reconfiguration process were not designed with mobility in mind, they introduce significant delays.

A. Router Discovery

Routing in IPv6 networks is configured using stateless autoconfiguration mechanism called Router Advertisements [5]. Router periodically announces advertisements describing what prefixes are available directly on the link and what routes are reachable via router. Those advertisements might also be requested by nodes, that are not willing to wait for the next unsolicited announcement. Although this procedure can be executed quickly, its main disadvantage is that no other configuration parameters, except routing, might be obtained. This implies the lack of basic parameters like DNS server addresses, VoIP configuration and other. The necessity to transmit extra messages and wait for responses introduces an additional delay every time the subscriber moves to a new location. Its magnitude depends on how fast the base station is able to grant the requested bandwidth and how fast the router is able to send responses. HD metric is estimated to be within a 40ms range.

B. DHCPv6

To set the remaining configuration parameters, stateful configuration is used. It can be achieved by using Dynamic Host Configuration Protocol for IPv6 [6]. Although more complex than stateless autoconfiguration, it offers far broader functionality: it conveys numerous additional configuration options, maintains control over address assignment, provides authentication, etc.

Initial stateful automatic configuration is divided into two phases. The first one is server discovery, during which a client sends a message informing, how many and what configuration parameters it is interested in. As DHCPv6 protocol offers server redundancy, all available servers respond with advertisements containing their proposed addresses and configuration parameters. As specified in [6], a

client waits one second to allow all servers to generate and send answers. The second phase is the actual configuration. The node chooses one of the available servers and requests configuration, which is granted by the server. Clearly, HD metric of the basic DHCP configuration is over 1000ms, so this is a major area for possible improvement.

C. Duplicate Address Detection

After new IPv6 address is obtained, according to [5] node is required to verify if this new address is not used. This procedure is mandatory for all IPv6 nodes starting to use new address, regardless of whether it was obtained through stateful or stateless autoconfiguration. When a handover (considered a network interface reinitialization) is completed, a node must initiate a DAD procedure, which introduces another 1000ms delay. Again, HD metric is over 1000ms.

D. Location update in Mobile IPv6

After successfully obtained and verified addresses, a node must inform its home agent and CNs about its new location. This procedure is also known as *Binding Update*. Although this procedure consists of an exchange of only two messages, they are not exchanged locally. Assuming that a message processing time is very small (thus can be neglected), this procedure takes a full round trip time from the current mobile node location to its CN. This might be the longest step in the whole handover process, described in the last two sections.

VI. Areas of improvement

As discussed in the previous sections, there are several mechanisms that introduce significant delays to the handover procedure. To mitigate, or in some cases, even eliminate such delays, a number of new improvements are proposed.

E. Remote DHCPv6

During a normal handover procedure, the data link layer (i.e. IEEE 802.16) initiates and performs the handover procedure. After it is completed, the network layer (i.e. IPv6) handover is performed. Doing it sequentially causes the delays introduced by each layer to add up, resulting in a large overall delay. To avoid this, data gathered by IEEE 802.16 may be used to exercise some preparatory steps before actual switching takes place. Although mainly dealing with horizontal handover (between two locations with the same network access type), the 802.21 Media Independent Handover framework proposed in [19] may be used for L2 handover notifications. A subscriber knows its target location before actual handover occurs. This prior knowledge may be exploited to initialize a connection with a DHCPv6 server, located within the destination network. As all BSs are connected to a common network (e.g., Internet or an ISP's network), it is possible to perform a connection between base stations using a backbone network.

To initiate and maintain such communication, existing DHCPv6 relays may be used, albeit in a modified form. In a classical configuration, relays work as intermediaries between clients and servers. From the client's perspective, direct communication with a server or via relays is indistinguishable. Relays act as representatives of the server. From the server's perspective, a client is connected to the remote link. By modifying relays' behaviour, it is possible to use them to forward data from a client to the server and vice versa. In this proposed scenario, a client is aware of the relays. It sends messages to relays and expects them to be forwarded to the specific remote server. Thus relays act as representatives of clients. From the server's perspective, a client is connected directly to the local link. To achieve such operation, clients, relays and servers must support this new mode. As this modification causes DHCP configuration to be performed while still maintaining full connectivity, it effectively cancels any negative impact on the handover delays. Thus HD metric value of this new improved DHCP configuration is zero.

F. DAD elimination

Every IPv6 node after receiving a new address must perform Duplicate Address Detection. It is intended to detect possible duplicates i.e., other nodes that use this recently acquired address. According to [5], a node is supposed to wait 1000ms for an unlikely response. Also, as this is the first IPv6 message to be sent from an interface after reinitialization, the node should also delay the transmission by a random delay between 0 and 1000ms.

In a real-life environment, address duplicates are extremely rare and usually a sign of a severe network misconfiguration or a malicious attack. In the latter case to spoof duplicate address, an attacker must have already penetrated the network. After a link has been compromised, the attacker can spoof numerous error conditions and DAD will not prevent him from doing so.

Therefore one proposal to limit the delays is to omit the DAD procedure completely. Such a radical proposal lowers HD metric of the DAD phase to zero.

G. Server side DAD

In some cases skipping DAD procedure completely may not be the best course of action. To expedite automatic configuration process, a server may maintain a small pool of IPv6 addresses that are checked against duplication. After startup, the server selects several addresses and performs DAD for each address added to this pool. After successful validation, an address is ready to be leased by clients. It is essential for the server to passively

participate in the DAD procedure for those addresses, i.e. to answer for possible DAD messages sent by other nodes trying to use such address. Before the server sends information about particular lease to the client, the assigned address should be removed from the server's interface (i.e. the server should stop responding to DAD messages sent to this specific address).

To make sure that the client supports this enhancement, it should send special suboption to the server in the IA_NA option. It will inform the server that this particular client supports "server side DAD" and it is possible to grant pre-validated address.

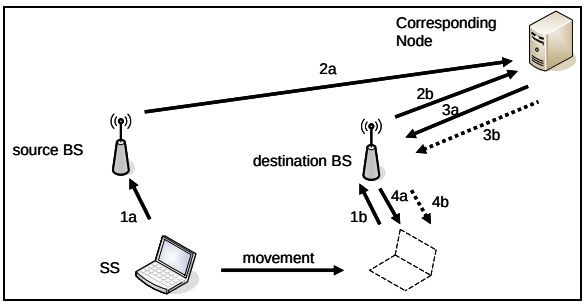

Fig. 2: Remote Location Update

H. Routing configuration

During normal configuration, IPv6 nodes are expected to wait for (or force the router to transmit) router advertisements to configure routing. From the mobile node perspective, that requirement is unfortunate as it introduces additional delays. To solve this problem, authors propose a new method for routing configuration. As clients have to exercise message exchange with DHCPv6 server, it may also provide extra information that allows clients to configure routing. To convey this information, a new DHCPv6 option – address parameters – has been proposed and support for this option has been implemented. See [8] for details.

I. Remote Location Update

Usage of the above described Remote DHCPv6 configuration provides access to a whole new class of possible solutions. One of them is a proposed improvement in Location Update, a crucial procedure in Mobile IPv6 protocol [7]. After configuration of a new address the mobile node sends necessary information to all CNs and the home agent. Corresponding nodes and the home agent update their location tables and send confirmation. This procedure often causes significant delays, as it requires quite a long round trip time to complete. Depending on the location of notified nodes, that may be several seconds.

Since a mobile node knows its new address before the actual handover takes place, it may send notification to its home agent and CNs before commencing actual handover. If calculated properly, a node will complete reconfiguration at destination location exactly between sending notification and receiving update, thus reducing HD metric to half the previous value. Unfortunately, it is not possible to reliably measure the round trip time as it fluctuates frequently. Also, since handover has been initiated, it is fair to assume that transmission conditions for one of the BSs are poor and it is likely that the transmitted message will be lost. To avoid a situation, where a mobile node moves to a new location and waits for a message from the corresponding node that may never arrive (due to the lost notification sent by the node, just before handover), it may retransmit update after completing movement at the destination location. The only drawback is a possible duplicate update, which can be safely ignored. This scenario is presented in Fig. 2.

VII. OPTIMIZATION SCENARIOS

In the paper 10 different scenarios were investigated. Each scenario, except the last, contains all optimizations introduced in previous ones.

1. **No optimization** – All possible optimizations (provided by standards as well as proposed by the authors) are disabled. During each handover, a mobile station must perform full network entry, recreate all service flows, obtain and verify its IPv6 address. There are no optimizations in the IPv6 layer. This scenario is considered the worst possible case.
2. **WiMAX optimization** – IEEE 802.16 provides an extensive set of possible optimizations. This includes context sharing between base stations, so that target base stations know in advance about incoming subscriber. If such a priori knowledge is available, a significant number of steps may be omitted, such as basic capability negotiation (SBC-REQ/SBC-RSP), registration (REG-REQ/REG-RSP) and key exchange (multiple PKM-REQ/PKM-RSP) and service flow creation (DSA-REQ/DSA-RSP/DSA-ACK). This results in significant reduction of the time required to perform 802.16 handover (i.e. PHY and MAC switch).
3. **DHCPv6: Skip initial delay** – DHCPv6 spec states that initial transmission "must be delayed by a random amount of time between 0 and 1 second". This feature is intended to prevent congestion following a power outage. Unfortunately, it introduces unacceptable delays if stateful (DHCPv6) autoconfiguration is required in mobile devices. Therefore the random delay is removed in this scenario.
4. **DHCPv6: Preference 255** – As DHCPv6 offers redundancy, it is possible for more than one server to exist on the same link. Therefore DHCPv6 offers a mechanism to discover all

servers. After transmitting a SOLICIT message, the server replies with an ADVERTISE message. This discovery phase takes exactly one second because clients are required to wait for possible responses from other servers, even when one or more servers have already responded. To omit this waiting phase, a server may send responses with the preference option set to maximum (255) value, causing the client to abort its discovery phase.

5. **DHCPv6: Rapid-commit** – After discovering a DHCPv6 server (two messages), the client requests an address (two messages), so in total four messages are exchanged. It is possible to shorten this to only two messages, with an actual assigned address may be sent in the reply to the initial SOLICIT message. This fast approach is called rapid-commit.
6. **IPv6: Skip DAD** – After obtaining an IPv6 address, according to, the node is required to perform Duplicate Address Detection – a mechanism intended to detect cases when assigned address is already used by other node. Again, this feature was not designed with mobility in mind, so it takes one second to wait for possible responses from other nodes using the same address.
7. **Server side DAD** – Duplicate Address Detection procedure is performed on the server side, as explained in section VI.C.
8. **Remote DHCPv6** – Instead of waiting for L2 handover to complete, remote DHCPv6 procedure is started before actual handover takes place. See section VI.A for details.
9. **Use address parameters in DHCPv6** – Some route information can be delivered together with addresses and other options via DHCPv6. See section VI.D.
10. **Use Remote Location Update** – Once client knows its next address while still in the old location, it can trigger Location Update procedure before actual handover takes place, as discussed in section VI.E

VIII. SIMULATION ENVIRONMENT

In order to reliably measure and analyze delays introduced by handover procedures and their impact on a lack of communication capabilities, an advanced simulation environment is necessary.

Currently there are no applicable solutions available, so the authors chose to implement a new one. To encourage open discussion and contributions to this, the authors have released this simulation environment, called Numbat, as an open source. For source code, see [9].

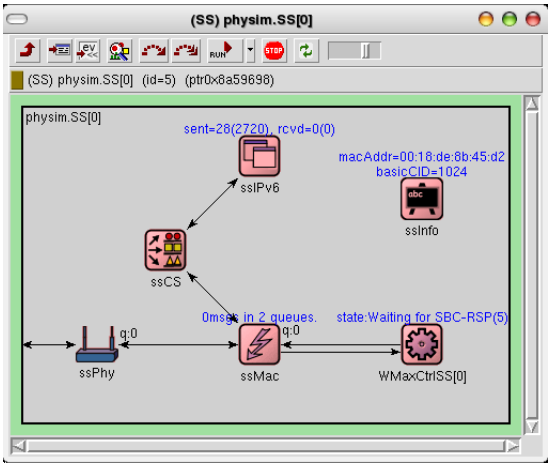

Figure 3: Subscriber station model

Numbat is a simulation environment designed and implemented in order to fulfill specific goals and requirements. Most important principles are:

- **Coverage** – Provide environment for simulation of the IPv6 capable 802.16 stations. That includes fixed subscriber stations, mobile subscriber stations and base stations.
- **Handover Oriented** – Simulate essential 802.16 mechanisms (e.g. network entry and handover procedure), but omit unnecessary mechanisms, not important from the mobility point of view (e.g. dummy implementation of the cryptographic protection, simple radio channel implementation)
- **IPv6 stack implementation** – Stateless autoconfiguration, Router Advertisements, DHCPv6 and Mobile IPv6. No full implementation is necessary, only aspects related to or affecting mobility.
- **Modular approach** – The Numbat environment is and will be under constant development. Instead of simulating one complex system, it has been split into numerous small modules, each interacting with its neighbors or parent only. To simulate complex entities easily, several simple modules may be grouped into larger complex modules.
- **Composite approach** – Although authors' interests focus on mobility and IPv6, other users may find different aspects more interesting. It must be very easy to modify or extend some parts of the environment without comprehending every detail of the whole simulation.
- **Flexibility** – Definition of the simulation parameters must be flexible and easy to modify. Therefore most of the parameters must be easily configurable.

- **Visual and text versions** – During development, debugging and presentations it is more convenient to use a graphical interface that allows visual inspection of all network elements. When simulation parameters are prepared, however, complicated simulations can take a long time to produce results. During such runs, visualization only slows down the simulation process so a command line interface is necessary. As an added benefit, a CLI allows remote execution on powerful servers.
- **Parallel approach** – Since Gdansk University of Technology provides multiprocessor clusters with up to 256 processors, it seems reasonable to take advantage of parallel processing.

After thorough evaluation of several environments, Omnet++ was chosen as a simulation engine. Compared to other (NS-2, ANVL) possible choices, Omnet++ has clear and well defined architecture, is fast (all modules are coded in C++), provides command-line and graphical interfaces, is modular, is well documented and is free for non-commercial use. Unfortunately, Omnet++ does not provide IPv6 or WiMAX simulation modules, so they had to be implemented.[4]

A. Subscriber and Base Station Architecture

IEEE 802.16 is an asymmetric protocol: a station acts differently depending on if whether it is a subscriber or a base station. However, there are some functional similarities. For example both stations have schedulers that coordinate transmission and reception of data messages. Therefore both station models have been split into similar modules.

A complicated network environment usually deals with a large number of data packets and a much smaller amount of control packets. For better performance data packets are often processed using a ”fast path” that is referred to as the ”data plane”. Control messages usually require additional processing, so their transmission and reception takes longer. This ”slow path” message processing is called ”control plane”. In the Numbat design, control plane and data plane have been split and implemented separately. The following modules have been implemented:

- **IPv6 module** – Sends and receives IPv6 messages. This is a composite module that represent full IPv6 stack. Consists of several submodules: IPv6Gen (an IPv6 traffic generator and analyzer), DHCPv6Cli (a DHCPv6 client), DHCPv6Srv (a DHCPv6 server), RaSrv (Router Advertisement server/router), RaCli (Router Advertisement client), MobIPv6Mn (Mobile IPv6 mobile node) and MobIPv6Ha (Mobile IPv6 Home Agent).
- **WMaxCS** – Convergence Sublayer. Classifies received data to corresponding connections and dispatches it to destination modules.
- **WMaxCtrl** – Represents the control plane, i.e. logic of the base or subscriber stations. All decisions are made here. It is an instance of the Finite State Machine.
- **WMaxMAC** – Represents the main part of the data plane, i.e. mainly scheduler (transmission) and dispatcher (reception).
- **WMaxPHY** – Simulator of the 802.16 PHY layer. Since PHY operation is outside the scope of discussed topics, this implementation is trivial.
- **WMaxRadio** – Present in the base station only. It is a simple simulator of the radio channel. It supports two transmission types: broadcast (one sender transmitting to many receivers, i.e. downlink: one base station to multiple subscribers) and unicast (one sender transmitting to one receiver, i.e. uplink: one subscriber transmitting to one base station).

Since base stations act as relays and do not generate any data traffic on their own, additional subscribers have to be added to simulate ”background” traffic. A fixed subscriber can be configured very easily to act as a traffic generator/analyzer.

The Omnet++ environment includes FSM implementation, but the provided interface does not offer the required flexibility. The most serious flaw in the Omnet’s FSM is that staying in the same state is not supported (state exit and entry must be performed). Therefore a new state machine framework has been developed.

B. Mobility model

There are several possible approaches to model mobility in Numbat:

- **Location based** – A mobile subscriber can change its physical location and periodically perform scanning. When it detects that there is a better base station than the one currently associated, it initiates handover. This model is more realistic, but it requires planning of base station locations and defining subscriber station path. Both problems are not trivial and can blur simulation results easily.
- **Time based** – It is possible to define that, regardless of its location, subscriber initiates a handover after a certain amount of time. This model is a simplification, but it is very useful for

[4]There is a separate project based on Omnet++ environment called INET, which provides IPv6 implementation. However, its complicated nature and fragmentation made it useless for the authors’ purposes.

scenarios focused on the handover procedure itself.

The example results presented in this paper are based on the latter approach. Handover is executed exactly three seconds after previous handover of the 802.16 layers has been completed. Delay values related to DHCPv6 protocol simulation are real values, measured on real hardware. The precise logging mode, which uses microseconds instead of the usual y:m:d h:m:s format has been implemented in Dibbler – a DHCPv6 implementation. During simulation data packets were sent with random sizes between 64 and 128 bytes. Packet sizes are not significant as they are used mainly as indicators if SS is able to communicate or not. Each scenario was run 10 times and the results were averaged.

IX. VALIDATION

There are several ways to assess effectiveness of proposed handover modifications. The first one is to construct a theoretical model that will describe analysed scenarios. However, it is extremely difficult to develop an analytical model reliably describing such a complex environment as an IEEE 802.16 network with multiple entities. Even worse, the IPv6 layer provides an additional level of complexity. Therefore this approach appears to be hardly feasible. The second way to assess the usefulness of modifications is to develop a simulation environment that will emulate all affected processes. Although complicated, this task is feasible. In general, simulation results can be accepted to prove usefulness of new solutions. Therefore this approach has been selected as a primary validation method. To further reinforce our claims, another verification method has also been used., namely some parts of the proposed improvements were included in a real DHCPv6 implementation.

Ten different scenarios (described in Section VII) were prepared. First five scenarios contained optimizations provided by existing standards, while scenarios six to ten provided proposed improvements.

The Numbat simulation environment provides several mobility models. As this research is focused on the handover process itself, rather than decision if and when initiate handover, handover performed after specified timeout was selected. Numerous parameters were observed, like the number of packets transmitted and received by a mobile SS and its corresponding node, average packet delay, number of packets dropped by a subscriber (due to lack of communication capability), received bits per second by SS, handover preparation time, 802.16 network reentry time, DHCPv6 configuration time, IPv6 reconfigure time and the duration of communication capability periods.

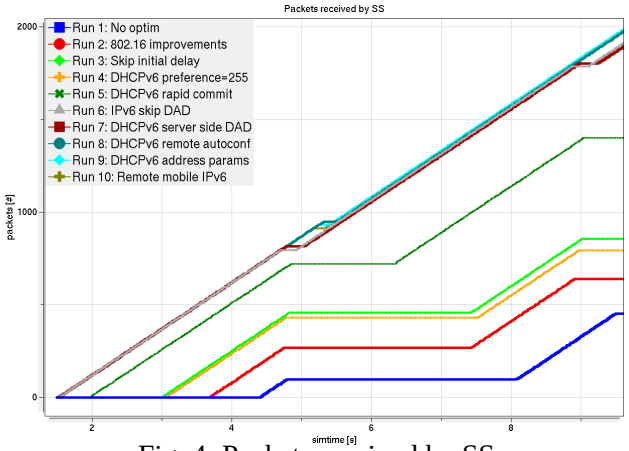

Fig. 4: Packets received by SS

It is clear that with enlarging the set of improvements, the number of received IPv6 packets increases. That may be considered as a first indication that the proposed mechanisms behave as expected. When handover delay is limited, the network is able to deliver more packets in the same time. Handover occurrences can be clearly identified as horizontal lines in the diagram. This relation, for all analyzed modifications, has been presented in Fig. 4.

It is also worth noting that the number of bytes, received per second by the SS, has significantly increased. Intervals of transmission inactivity on Run 1 are clearly visible. For the mobility model used, the last scenario not only allows transmitting more data in a continuous manner, but also makes it possible to complete a greater number of handovers. (In our mobility model, a next handover is performed 4 seconds after the previous L2 handover was completed.)

In every case data transmission begins after initial network entry. Differences start to appear during the first handover. In the worst case (Scenario 1), full handover procedure takes almost as long as handover intervals, so time for real data exchange is minimal. Skipping initial DHCPv6 delay considerably decreases handover time vs. transmission time ratio (Scenario 3). It appears that the greatest advantage is to further optimize DHCPv6 exchange by using rapid-commit option or maximum (255) server preference option value. Another significant improvement is gained by using 802.16 handover optimizations. Most of them assume that target base station has *a priori* knowledge about the incoming subscriber. To gain such knowledge, some off-the-network communication framework between base stations is required. It appears that in real life-solutions, operators will deploy such operation, administration and maintenance (OAM) entities that also serve several other purposes, such as accounting and network management. Deploying such entities,

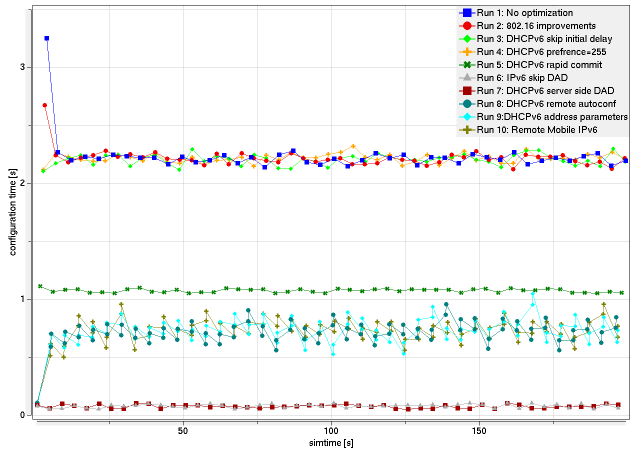

Fig. 5: DHCPv6 configuration time.

called ASN Gateways, seems inevitable for networks more complicated than a single base station. With such information about an expected subscriber station, a base station is able to skip most otherwise required steps. This in turn shortens the reentry procedure, which results in a greater number of received packets

The last analyzed optimization is related to the Duplicate Address Detection mechanism in IPv6. Since there are 2^{128} addresses available in the IPv6 address space, it seems highly unlikely that duplicates will ever occur, except during malicious attacks. A prevention mechanism, designed by IPv6 spec authors, allows for protection against such cases. Unfortunately it introduces one-second delay between address assignment and its actual exploitation.

Comparing all scenarios together it is clearly visible that the most useful optimizations are possible in the IPv6 related areas (DHCPv6 and DAD). However, to obtain the best results it is strongly recommended to combine DHCPv6, IPv6 and 802.16 optimizations.

As an intermediate value, DHCPv6 configuration time (i.e. time required to complete DHCPv6 message exchange) was also measured (see Fig. 5). In general, with more improvements, this time decreases. There is an exception, though. All scenarios that use remote autoconfiguration take longer to complete. That is understandable, as messages have to be exchanged between BSs. It is also crucial to understand that during remote autoconfiguration, a subscriber station maintains full communication capability, thus HD metric is zero.

X. CONCLUSIONS AND FUTURE RESEARCH

The conducted research focused on WiMAX IEEE 802.16-2005 and DHCPv6 implementations. All major elements of the WiMAX stack were modeled and tested, and were proved to operate properly.

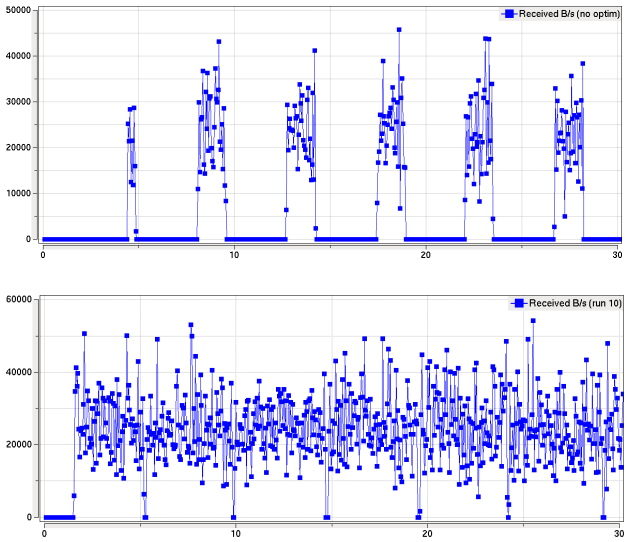

Fig. 6: Received bytes per second by mobile node. a) before optimization b) with all optimizations and proposed enhancements enabled

The current IEEE 802.16-2005 standard, covering two lower layers in the ISO/OSI protocol stack, offers quite good mobility support and there are no significant areas of necessary improvements. However, from the mobility perspective, IPv6 standard does not support real-time mobility. IPv6 protocols (like DAD or DHCPv6) introduce significant, and often unnecessary, delays. To overcome at least some of these delays, several new methods were proposed. All proposals were validated and verified, using the Numbat simulation environment. Results obtained with the use of Numbat strongly suggest that the largest delays are caused by the DHCPv6 protocol and Duplicate Address Detection.

Combining optimizations on several layers, both offered by already existing standard as well as proposed by authors, gives essential improvements in data transmission efficiency in a mobile environment. Graphs obtained for 2 mobile stations are presented in Fig. 6. One can see for the most advanced optimization scenario (all improvements enabled; see Fig. 6b) transmission breaks are almost entirely eliminated.

It is worth noting that the proposed HD metric allows estimating essential delays introduced by different actions during handover. The next step to ultimately confirm usefulness of the discussed proposals is to implement them in a real DHCPv6 environment and to validate them in a controlled network. Some features, e.g. routing configuration via DHCPv6, was implemented already as part of the Dibbler[5] project [6] – an actual implementation of the DHCPv6 protocol, developed at Gdansk University of Technology since 2003. Since it is

[5]Available for Windows and Linux, accepted in 5 linux distributions, confirmed use in 29 countries.

widely used software, users' feedback will eventually confirm the usefulness of proposed features in a real-life environment.

As proposed enhancements appear to be reasonable solutions, work is in progress to specify an RFC draft and to submit it to IETF as an independent proposal.

ACKNOWLEDGMENT

This work was supported in part by the Polish National Center for Research and Development under the PBZ grant MNiSW – 02/II/2007. The authors also would like to thank Andrzej Bojarski and Maciej Jureko for their contribution to the Numbat development.

REFERENCES

[1] IEEE working group, "IEEE 802.16-2004: IEEE Standard for Local and Metropolitan Area Networks – Part 16: Air Interface for Fixed Broadband Wireless Access Systems", IEEE, Oct.2004

[2] IEEE working group, "IEEE 802.16-2004: IEEE Standard for Local and Metropolitan Area Networks – Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems", IEEE, Dec.2005

[3] S. Deering, R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC2460, IEEE, Dec.1998

[4] T. Narten, E. Nordmark, W. Simpson "Neighbor Discovery for IP Version 6 (IPv6)", RFC2461, IETF, Dec.1998

[5] S. Thomson, and T. Narten "IPv6 Stateless Address Autoconfiguration", RFC2462, IETF, Dec.1998

[6] R. Droms, Ed. "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC3315, IETF, Jul.2003

[7] D. Johnson, C. Perkins, J. Arkko "Mobility support in IPv6", RFC3775, IETF, Jun.2004

[8] T. Mrugalski, "Dibber – a portable DHCPv6", project homepage, http://klub.com.pl/dhcpv6/, Nov.2008

[9] T. Mrugalski, "Numbat – mobile IPv6 in WiMAX environment", project homepage, http://klub.com.pl/projects/numbat/, Nov.2008

[10] M. Kraus, "DOD: Transition to IPv6", http://www.usipv6.com/2003arlington/presents/ Marilyn Kraus.pdf, Dec.2003

[11] WiMAX Forum, "Mobile WiMAX – Part II: A Comparative Analysis", http://www.intel.com/netcomms/technologies/ wimax/mobile wimax p2.pdf, Apr.2006

[12] K. Tae Lee, "Create the future with Mobile WiMAX", Communications Magazine, May 2007

[13] J. Wozniak, K. Nowicki, T. Mrugalski, "Mobile users issues, in micro and macro scale, in IP networks", SIS2004, Lodz, Sep.2004

[14] A. Conceiao, J. Li, D. Florencio F. Kon, "Is IEEE 802.11 Ready for VoIP?", Microsoft Research, IEEE International Workshop on Multimedia Signal Processing, 2006

[15] P. Matusz, P.Machan, J.Wozniak, "Analysis of profitability of intersystem handovers between IEEE 802.11b and UMTS", LCN'03, IEEE International Conference, Oct.2003

[16]Wikipedia "IPv4 address exhaustion, http://en.wikipedia.org/wiki/IPv4 address exhaustion, Aug.2007

[17]J. Wozniak, T. Mrugalski "Numbat – extensible simulation environment for mobile, IPv6 capable IEEE 802.16 stations", ATNAC'2007, Christchurch, NZ, December 2007

[18]J. Wozniak, T. Mrugalski "How to Improve the Efficiency of IPv6 Handovers in IEEE 802.16 Networks", ATNAC'2008, Adelaide, Dec.2008

[19] IEEE 802.21 working group, http://www.ieee802.org/21/