THE DAIDALOS ARCHITECTURE FOR MOBILITY AND QoS

Telemaco Melia*, Marco Liebsch*, Pablo Serrano [†] , Albert Banchs [†]

[*] Network Laboratories, NEC Europe Ltd, Heidelberg, Germany
[†] Universidad Carlos III de Madrid, Leganés (Madrid), Spain

Keywords: Maximum 5 keywords placed before the abstract.

Abstract

Telecom operators and internet service providers are heading for a new shift in communications paradigms. The forthcoming convergence of cellular and wireless data networks is often manifested in an "all IP approach" where all communications are based on an end-to-end framework of IP protocol. The approach to network design becomes user and service-centered where continuous reachability of mobile users and sustained communication capabilities are default requirements for a prospective architecture.

In this paper we describe a network architecture which is able to provide seamless communication mobility, triggered either by the user or by the network, across multiple technologies. The architecture implements an approach to heterogeneity management in which supporting functions are divided between the end-devices and the network infrastructure. It includes a cross-layer design of a mobile terminal and of the supporting functions located in the network. The architecture allows for media independent handover and supports optimized and localized mobility management. The main focus of the paper is on major technical highlights of mobility as well as quality-of-service management subsystems for converged networks, and experimental results are provided to support these highlights.

1 Introduction

The 3G standard for wireless communications combines high-speed mobile access with IP-based services. Tomorrow's customers will expect the network, and in particular its technological structure, to "disappear" and be of no concern. Along these lines, previous work [2][3][4] have proposed solutions that support seamless mobility based on the Internet Protocol version 6 (IPv6). While supporting common or similar functions with parts of the 3GPP architecture, these solutions are based on IP-like technologies. In particular, while previous work has shown that the basic concepts are viable, the Daidalos project [5] has moved to a comprehensive approach to provide seamless and pervasive end-to-end services across heterogeneous technologies, offering a broad range of services accessible anytime and anywhere regardless of the wireless or wired technology.

Daidalos architecture proposes an enhanced IPv6 mobility platform that fulfils MARQS requirements. Fast intra- and inter--technology handovers are a solution to the requirement of seamlessness. While IETF protocols have been proved to work properly in a "stand-alone" manner, their applicability in integrated environments requires further enhancements. For next generation integrated systems, additional requirements are the optimization of resource usage (e.g. bandwidth in wireless environments), scalability for an increasing number of customers, and increased network flexibility. Previous studies and prototypes [4][6][7] already demonstrated the feasibility of parts of these integrated IP-based solutions. Measurements showed that non-mobility aware applications, both TCP and UDP based, can provide reliable services, with no packet loss and seamless roaming in these environments [8], [9].

In this paper we show how this architecture has improved, enhanced and optimized existing micro-mobility schemes to work integrated with the macro-mobility management scheme of the de facto standard Mobile IP version 6 [10], and in this way, we illustrate how it improved the overall system scalability. Support for both network initiated and mobile terminal initiated handovers (increasing network flexibility – different handover technologies have different handover concepts) integrated with QoS resource management is achieved. The next section presents an overview of our scalable architecture. Sections III describe how it operates during the handover decision and execution. Section IV describes complexity issues related with our implementation. The paper concludes with section V.

2 Daidalos Mobility Architecture Overview

The Daidalos mobility architecture aims to provide an efficient and scalable integration of multiple network technologies, with sustained QoS support. The (simplified) general view of the architecture is illustrated in Figure 1. As it can be observed from the figure, the architectural design follows a hierarchical structure: the network of each Mobile Operator consists of a core network (two such networks, from different operators, are represented in the figure) and a set of access networks. The access networks contain multiple Access Routers (ARs), with multiple radio Access Points (APs) each. The architecture supports multiple access technologies, including WLAN (802.11), WiMAX (802.16), TD-CDMA and DVB.

Each access network is called a region. Resources in each region are independently managed by an Access Network QoS Broker (QoSB-AN), providing a first scalability step. Resources in the core are managed by the Core Network QoS Broker (QoSB-CN), which communicates for end-to-end QoS with the QoSB-ANs of the Mobile Operator's network as well as with the QoSB-CNs of the other operator's networks.

The architecture is based on widely accepted standards for mobility and QoS. Mobility is implemented by means of the MIPv6 protocol [9], with Fast Handover extensions [10], and QoS is based on the DiffServ architecture [12]. However, additional mechanisms that integrate and complement MIPv6 and DiffServ are needed in order to achieve the objective of providing QoS to mobile users while optimizing the overall performance. Such mechanisms have been designed in the architecture. Specifically, the architecture includes mechanisms for enhanced handover decision and execution [13]. Furthermore, it is also able to consider L2 QoS mechanisms of the various wireless technologies.

Handover decisions in our architecture are sustained both by measurements on signal quality as well as L3 QoS measures (such as load and resource availability). Handovers can be started either by the terminal or by the network. We refer to the former as a mobile terminal initiated handover (MIHO) and to the latter as a network initiated handover (NIHO). Handover execution is improved with functions for maintaining quality during handovers, along with tight coupling with QoS functions.

In the following we describe with further detail the enhanced functionalities of the architecture. The modules needed to instantiate the functionalities discussed above are illustrated in Figure 2, and organized according with their physical location.

- **Enhanced MIHO decisions.** Handover decisions in the case of a MIHO are enhanced with the objective of ensuring that from all possible AP candidates the “best” is chosen. The module responsible for the handover decision at the MT is the Intelligent Interface Selection (IIS). This module relies on the Mobile Terminal Controller (MTC) to obtain the information it uses to take a decision. This includes signal quality measurements, obtained from the Mobility Abstraction Layer (MAL), as well as L3 measures, such as load of the APs, retrieved from the candidate APs. The latter information is obtained from QoS Abstraction Layer (QAL) in the neighbouring ARs, and conveyed by means of the Candidate Access Router Discovery Protocol (CARD) [14] to the MT. With this information, the target AP for the handover is chosen by the IIS so that both signal strength and QoS requirements are met in the new AP, thus guaranteeing appropriate operation and service quality after the handover. MIHO execution is then triggered by the Mobile Terminal Controller (MTC).
- **NIHO functionality.** The enhanced MIHO functionality ensures that handover decisions are taken optimally according to local information, but does not guarantee that the overall distribution of resources will be optimal from an operator perspective – which is essential for a realistic network. In order to achieve this, NIHO support is required to allow the optimization of the overall capacity in the access networks of the mobile communication infrastructure by properly balancing the load among the various APs of a region. For this purpose, the Performance Management (PM) module at the QoSB-AN collects information about the load of the different APs and the radio link quality between the MT and its candidate APs, and based on this information (eventually) reorganizes the wireless connections. Information on the load of the APs is obtained by the Performance Attendant (PA) modules at the APs, from their interface with the QAL, and delivered to the PM. Signal-strength quality measurements are taken by the MMs, filtered out and aggregated by the Aggregation Module (AM), provided to the PA at the AR and from there conveyed to the PM. Based on all these data, the PM then reorganizes the

connections of all MTs for an optimized global performance. This reorganization takes into consideration QoS beyond the wireless access, by means of the interaction between the PM and the QoSB Engine at the QoSB-AN. The NIHO execution is then triggered by the communication between the QoSB and the FHO execution module at the AR, through the Advanced Router Mechanism (ARM) managing QoS inside the AR.

- **Seamless Handover execution.** In the execution of a handover involving the old AR (oAR) and the new AR (nAR), continuity of communication is required to be maintained. The architecture is based on the Fast Handover for Mobile IPv6 protocol [10]. To perform a low latency, lossless, handover this is enhanced with Duplication and Merging (D&M) functions. These functions improve performance by duplicating the packets addressed to the MT at the old AR to avoid packet loss. In order to set-up the MT's context in the nAR, the Context Transfer (CT) function is used to transfer the mobility related state (including security information). The resource reservation for the MT in the nAR is also performed before executing the handover.
- **Quality of Service.** QoS is based on the DiffServ architecture. Admission control and resource reservation are handled by the QoSBs, which act jointly to perform QoS reservations over an end-to-end path. QoS reservations at the routers are performed through the interaction between the QoSB Engine at the QoSB-AN and the ARM module at an AR, which performs the reservation via the QoS Manager (QoSM). Similarly, reservations in the wireless access part are performed through the interaction between the QoSB Engine at the QoSB-AN and the ARM module at the corresponding AR; the latter communicates with the QoSM which in its turn communicates with the QAL at the AR. QoS reservations in the wireless access are then performed by the QAL modules at the AP and MT, with which the QAL module at the AR communicates.
- **Multiple technology support.** The support of multiple technologies in the architecture is provided by means of a modular design based on the use of Abstraction Layers (AL) - the Mobility Abstraction Layer (MAL) and the Quality Abstraction Layer (QAL). These ALs interface with drivers of the different technologies and offer a unique interface to the upper layer modules of the architecture, while hiding the specifics of the underlying technologies. The QAL offers a technology independent interface for QoS functions such as set-up of a QoS connection or the measurement of the available resources in an AP. Similarly, the MAL offers a technology independent interface for mobility related functions such as the execution of a handover or measurement of signal strength received at MT.

In the following sections, the above functions and their interactions are described in detail. First, we describe the operations related to the decision of performing a handover for the mobile initiated and network initiated cases (MIHO and NIHO). Then, we address the process of handover execution, which is almost identical for both cases.

3 Handover Decision and Execution Functions

In existing IP based architectures, handovers are typically initiated by terminals upon detecting that the quality of the signal received from the AP degrades below a certain threshold. However, there are situations in which it is desirable that handovers are triggered by the network, such as e.g. when the load is not optimally distributed among various APs. The proposed architecture thus supports both MIHO and NIHO, and the decision processes consider these two handover types.

Mobile Initiated Handover Decision. The MIHO operation is depicted in Figure 3. The main decision to be taken in this case is the choice of the new AP to handover. In the preparatory phase of this process, the MT discovers the available candidate APs. This is done by means of the CARD protocol [14]. With this protocol, neighbouring ARs exchange QoS information about attached APs and provide it to the MT (message 1 in Figure 3). This information is sent from candidate ARs (including the future nAR) to the current AR (oAR), and from there it is conveyed to the MT (messages 2 and 3, respectively). Once the MT has obtained information about available candidate APs and their QoS, it proceeds to measure their signal strength. The decision of which APs are to be measured is made by the MTC based on the CARD information (message 4) and is provided to the drivers of the respective technologies via the MAL (messages 5 and 6). The measured signal qualities are reported back to the MTC (messages 7 and 8) and this information, together with the QoS related information obtained previously, is then provided to the IIS (message 9). Based on these data, the IIS decides which is the most appropriate AP for handover. Based on this information, the IIS chooses the most appropriate candidate Ap and informs the MTC (message 10). At this point the handover execution process starts (see Mobile Initiated Handover Execution below).

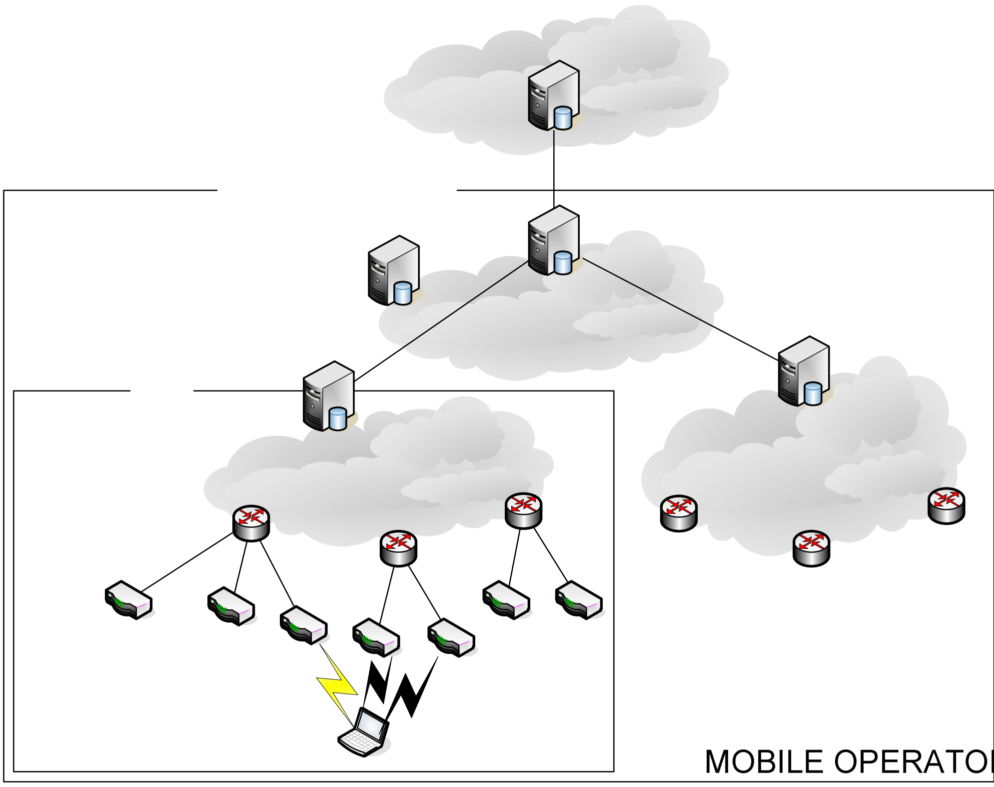

Figure 1Hierarchical structure of Mobility and QoS

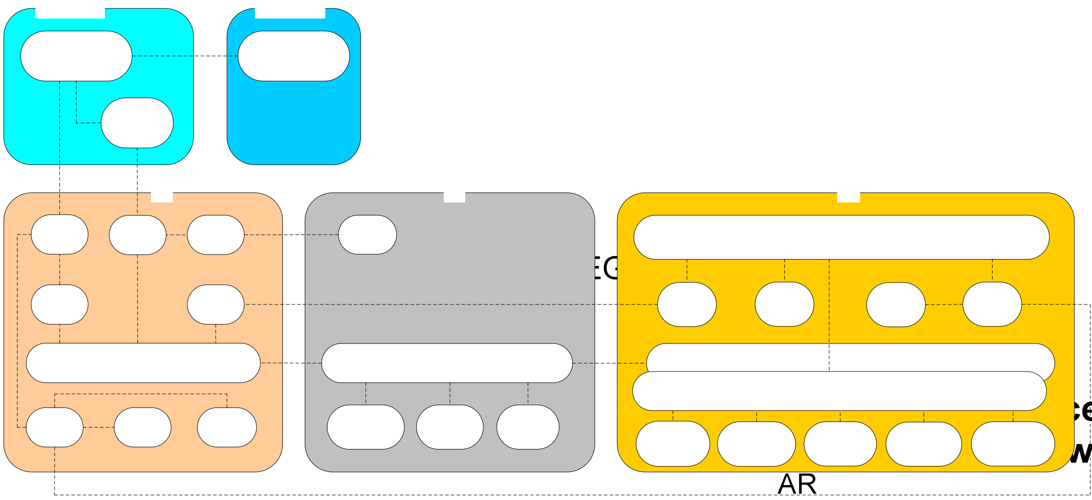

Figure 2 Architecture Modules

Network Initiated Handover Decision. The NIHO operation is depicted in Figure 4. NIHO may be triggered either by the degradation of the quality of the signal received from the MT by the AP or by the QAL upon detecting that the load of an AP exceeds a predefined threshold. NIHO decisions are taken with the goal of optimizing global performance in a region controlled by a QoSB-AN. The NIHO process is described in the following. Signal strength measurements are taken by MM modules at the APs and from there they are transferred to the AM modules at the ARs (message 1, Figure 4), which aggregate all the information received and provide it to the PA (message 2). By measuring the strength of the signal received from a given MT at all the APs of a region, it is possible to estimate all the candidate APs that provide good signal quality to this MT.

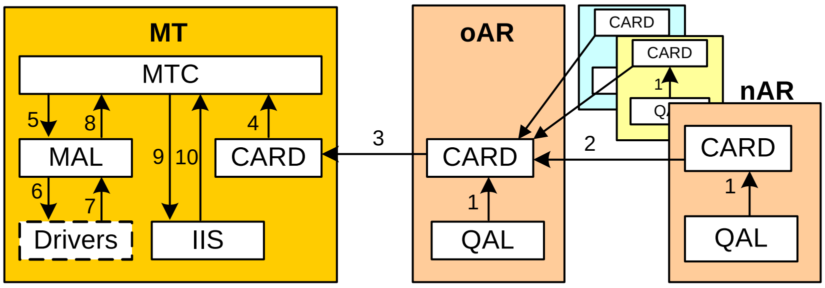

Figure 3 MIHO decision operation

In addition to signal strength data, the PA also collects QoS-related information from the QAL modules (message 3). The data of all the PAs of a region is sent to their corresponding PM module, which is located at the QoSB-AN controlling the region (message 4). With all the above information, the PM is aware (through the QoS related data) of the load of the various APs of the region, and is also aware (from the measurements taken) of the possible candidate APs that each MT may be handed over to while preserving a good signal quality. Based on these data, the PM can then decide the AP to which each MT should be attached to such that 1) load is optimally distributed among all the APs of a region, and 2) the signal strength of all connections is good. These decisions are checked against the QoSB engine in order to make sure that end-to-end QoS requirements are kept for the connections (message 5). At this point the handover execution process starts (see Network Initiated Handover Decision below).

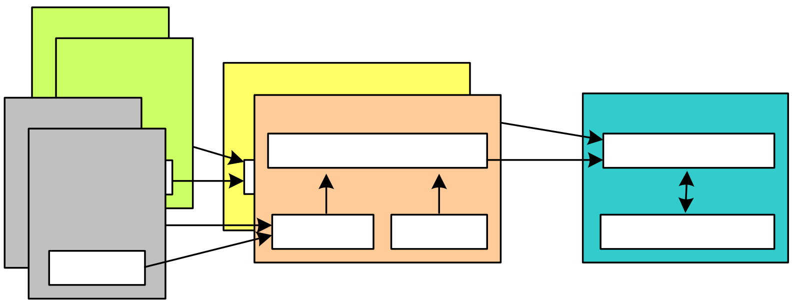

Figure 4 NIHO decision operation

Mobile Initiated Handover Execution. The Mobile Initiated Handover execution is performed as follows. In a MIHO, the MTC, upon receiving information from the IIS to trigger the handover procedure, activates the FHO module in the MT (message 1) to initiate a handover. It then sends a RouterSolicitationProxy message (message 2) to its current AR with the chosen candidate ARs. The AR, through the ARM module, forwards the request for approval to the QoSB (messages 3a and 3b). Upon receiving this request, the QoSB verifies the availability of the required end-to-end QoS, informs the nAR (message 4) of the QoS requirements (messages 4a), and sends back to the oAR its handover decision (message 5a and 5b). After the oAR processes message 5, it triggers the CT, instructs the D&M to start duplication, and informs the MT that it can now move to the network provided in the message ProxyRouterAdvertisement (messages 6a, 6b, 6c). As soon as the MT receives a ProxyRouterAdvertisement, it starts the merging process (messages 7a, 7b, 7c and 7d). Then, the MT sends a FastBindingUpdate (FBU) message (message 8a) to the oAR, which informs the QoSB of the MT's decision (message 8b and 8c). This is followed by disconnection from the current link and attachment to the new one (message 9a and 9b). Upon connection to the nAR interface, the MT sends a FastNeighbourAdvertisement (FNA) message (message 10). The nAR then informs the QoSB that the MT is attached on the new link (message 11), and therefore this indication is then forwarded to the oAR (message 12) in order to delete reservations (messages 12a), and stop D&M (messages 12b, 12c and 12d). After this process, the oAR informs the QoSB (message 13) that the reservation release actions have been successfully performed.

AP

AP

AR

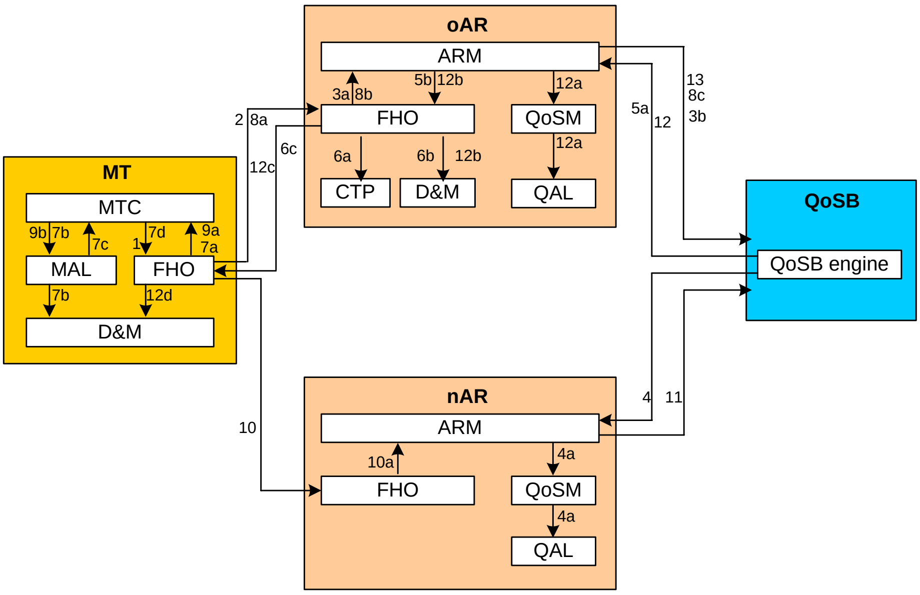

Figure 5 FHO Operation

Network Initiated Handover Execution. When the network initiates a NIHO, the procedure is similar to the one described above. However, now the MT does not send any RouterSolicitationProxy (message 2) and the ProxyRouterAdvertisement (message 6c) contains a flag indicating that this is a NIHO and information of the network to where the MT has to move. The remaining process is basically the same.

4 Implementation Experiences

In order to validate the proposed architecture in terms of protocol design and basic functionality we implemented the proposed architecture and evaluated its performance in the testbed illustrated in Figure 6. Our implementation was done on Linux 2.6.8.1 with MIPL basic Mobile IPv6 support. The description of the testbed and results hereafter are based on the Wireless LAN technology, although implementation is also available for TD-CDMA with some hardware limitations.

In order to provide the mobility features required for MIHO and NIHO, we introduced modifications to the behaviour of the WLAN driver we use in MTs and in the APs. A normal WLAN driver decides about handover and executes it only based on signal strength. The handover decision and execution can be done at firmware or driver level because all the information required is available in both places. In our architecture, the handover decision is influenced by other factors and therefore the decision cannot be taken at the firmware/driver level. Besides, in our architecture the decision must be separated from the execution. In MIHO, after the decision, some preparation must be done before the execution can start; in NIHO the decision is taken in the network and the MT just has to execute the handover. For this reason, the first modification we implemented was to disable automatic handovers and use a function to force handover execution when it is required. For MIHO we also implemented a function for scanning a selected subset of channels. This allows performing signal strength measurements at the required times and limited to the channels provided by CARD, in contrast to the default automatic scanning whose execution cannot be controlled and where scanning covers all channels, a slower process that results in larger latencies. Therefore, our implementation can benefit from the information obtained via CARD to decrease layer 2 handover latency.

In NIHOs, APs are required to measure the signal strengths of the MTs connected (using different channels) to other APs. In order to make these measurements, we installed a second WLAN card at the APs whose function was to scan all channels periodically and perform passive measurements on the signal strength detected from the MTs. QoS functions, required both for MIHOs and NIHOs, were developed based on the algorithm of [15] for admission control in WLAN.

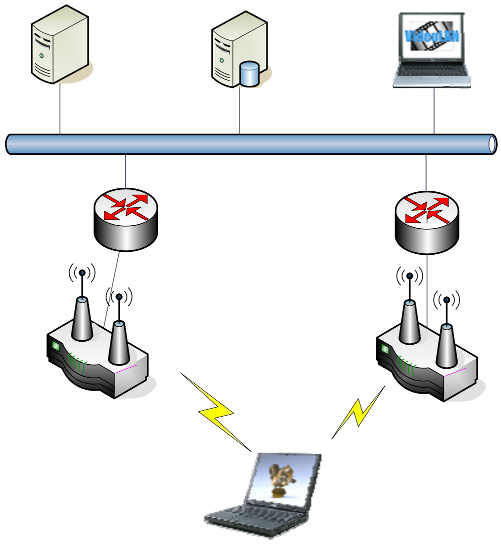

Figure 6 Testbed

With the above testbed, we performed experiments for MIHO and NIHO and obtained some preliminary results. By physically moving the MT from AP1 towards AP2, MIHOs were triggered. We measured the time that the MT takes to select the new AP and perform the signaling that precedes handover. Measured times were of about 1 second. Although strongly dependent on hardware and technology characteristics, and with the limitations inherent to a testbed, we can nevertheless argue that these times are low enough for realistic scenarios (e.g. with overlapping coverage areas of 30 m, speeds of above 100 Km/h are allowed). Handover times were also measured. Results, of about 50 ms, were also satisfactory according to the measures of [6].
NIHOs were forced by issuing a new QoS request such that one AP (AP1) became heavily loaded and one of the MTs was moved to AP2 in order to unload AP1. We measured the time elapsed between the QoS request and the beginning of the handover execution. This is the extra time that NIHO needs to collect the measurements information. The times we measured were below 2 seconds; note, however, that in a running commercial system measurements will probably be regularly collected and already present on the QoSB, and thus these times can be substantially reduced. We also measured the handover execution time. The results obtained were similar to MIHO case, as both cases involve almost the same functions, the main difference being that NIHO handovers have the advantage of not needing the scanning in the search for handover candidate APs.

5 Conclusions

The architecture presented, IP-based, integrates multiple technologies in a seamless environment, very flexible in terms of the handover possibilities, MIHO and NIHO, intra- and inter-technology, with integrated QoS support. The overall design integrates and enhances some work-in-progress and trends inside IETF, 3GPP and IEEE. The architecture is inherent hierarchical, both at a mobility and QoS-support levels, making it highly scalable. On the mobility part, the architecture decouples the notion of domain (multiple domains may interoperate), the notion of global mobility (supported by Mobile IP), and the notion of local mobility (supported by micro-mobility protocols). On the QoS part, the architecture assumes a differentiated services core with fine granularity control in the borders; this fine control allows for optimized performance, both from the view of the network and from the mobile terminal.
Linking these aspects, the proposed changes to FHO protocols are essential, supporting both MIHO and NIHO. The usage of monitoring functions, and their integration on the mobility process, is also an added advantage of our work. The changes developed are conceptually simple, and can be deployed with low-cost equipment, given the scalable approach described. Our architecture allows for seamless handover across technologies, with good performance levels. The integration of any Ethernet-influenced technology is straightforward, and work is

proceeding on the integration of WiMax. In fact, the architecture can even interoperate with non-IP based architectures, using strategies similar to the 3G GGSN developments, and with broadcast technologies, such as DVB, using UDLR-based strategies. Furthermore, although not completely described here, the architecture has added advantages, such as integrated security and AAA functionalities [5].

Acknowledgement

We presented work has been realized in the framework of the collaborative project Daidalos.

References

[1] "Voice call continuity between Circuit Switched (CS) and IP Multimedia Subsystem (IMS) Study", 3GPP specification: 23.806.

[2] H.-Y. Lach, et al, "Network Mobility in Beyond-3G Systems", IEEE Communications Magazine, Vol. 41, Issue 7, July 2003.

[3] Joachim Hillebrand et al, "Quality-of-Service Signaling for Next-Generation IP-Based Mobile Networks", IEEE Communications Magazine , June 2004, pp 72-79.

[4] J. Jähnert, et al, "The 'pure-IP' Moby Dick 4G architecture", Elsevier, Computer Communications, Vol. 28, Issue 9, 2nd June 2005, ISSN –140-3664.

[5] R. L. Aguiar et al, "Designing Networks for the Delivery of Advanced Flexible Personal Services: the Daidalos approach" Proc. IST Mobile & Wireless Telecommunications Summit, Lyon 2004.

[6] C. J. Bernardos, et al, "Experimental evaluation of a handover optimization solution for multimedia applications in a mobile IPv6 network", European Transactions on Telecommunications, Vol. 16, Issue 4, Jul./Aug. 2005.

[7] Victor Marques, et al, Evaluation of a Mobile IPv6-based Architecture Supporting User Mobility, Quality of Service and AAAC in Heterogeneous Networks, IEEE Journal on Selected Areas in Communications, Special Issue "Wireless Overlay Networks Based on Mobile IPv6", Vol. 23, no. 11, Nov 2005.

[8] Xavier Perez-Costa, et al, "A performance comparison of Mobile IPv6, Hierarchical Mobile IPv6, fast handovers for Mobile IPv6 and their combination", ACM SIGMOBILE Mobile Computing and Communications Review, Volume 7 , Issue 4 , October 2003.

[9] R. Hsieh et al "Performance analysis on Hierarchical Mobile IPv6 with Fast-handoff over End-toEnd TCP", GLOBECOM, Taipei, Taiwan 2002.

[10] D. Johnson et al, RFC 3775, "Mobility Support in IPv6", June 2004.

[11] R. Koodli (ed.), RFC4068,"Fast Handovers for Mobile IPv6", July 2005.

[12] S. Blake et al, RFC 2475"An Architecture for Differentiated Services", December 1998.

[13] T. Melia et al., "Network-controlled fast handovers", Internet draft (work in progress).

[14] M. Liebsch et al, RFC4066 "Candidate Access Router Discovery (CARD), July 2005.

[15] A. Banchs et al., "End-to-end Delay Analysis and Admission Control in 802.11 DCF", to appear in Elsevier, Computer Communications.