

An Experimental Cross-Layer Approach to Improve the Vertical Handover Procedure in Heterogeneous Wireless Networks

Rosario G. Garroppo, Stefano Giordano, Stefano Lucetti, Giuseppe Risi, Luca Tavanti

Original scientific paper

Abstract – Users of next generation wireless devices will be likely to move across a heterogeneous network environment. This will give them the possibility to always exploit the best connection to the global Internet. In order to keep a seamless connection, the handover between different access technologies, also known as vertical handover, must be as smooth as possible. The current evolution of network architectures toward an all-IP core favours the use of the Mobile IPv6 protocol to handle such handovers. However, this protocol still presents several drawbacks, mainly related to the assumption of static devices and wired connections. Hence we have designed and implemented a software module that exploits information from the lower layers (e.g. physical) to extend the capabilities of Mobile IPv6 to wireless environments. We have then evaluated both the plain Mobile IPv6 and our proposed implementation over an experimental testbed. The outcome of the assessment proves the effectiveness of our solution and reveals the possibility to perform a seamless vertical handover in heterogeneous wireless networks.

Index Terms – Mobile IPv6, vertical handover, handover decision, IEEE 802.11 WLAN.

I. INTRODUCTION

The number of mobile devices embedding multiple interfaces for different wireless access technologies (e.g. cellular, Wi-Fi, Bluetooth) is growing fast. This will enable users of next generation wireless networks to enjoy ubiquitous access to a plethora of services in an *always best connected* (ABC) mode. ABC will give them the possibility to keep alive their ongoing communications while moving freely across heterogeneous networks, with seamless service continuity achieved without the need of any active switching operation.

In this context, the main obstacle is the loss of service that may occur during the handover process, i.e. when the terminal leaves the current network to join another network (or network operator). More precisely, we speak of *vertical handover*, since it takes place in an environment of overlaid networks and involves a change in the access technology. The vertical handover process is deeply different from the traditional *horizontal handover*, in which the user just moves from one base station or access point to another, both belonging to the same network and usually to the same operator (as it happens e.g. in GSM and UMTS networks). In this case, the network drives and handles most of the operations and automatically

cares for the routing of packets, user authentication, accounting, and so on. In contrast, a vertical handover requires an active participation of the mobile terminal, which must choose the network to connect to depending on the user requirements, the available resources, and other parameters. The terminal also has to inform all the devices involved in the communication of its new position, caring to provide updated routing information and other useful data. Hence, a vertical handover poses much more issues for the user to experience a seamless service. In the remainder of the paper we only address the vertical handover, which we will refer to with the simple term handover.

The occurrence of a service rupture mainly depends on the type of handover. Basically, we can distinguish a *soft* handover and a *hard* handover. In the first case, the user's terminal decides whether and when to change network. Thus it can leave the old network after having performed all the configuration and signalling to be admitted and connected to the new network. In this way it is possible to completely avoid losses. In contrast, a hard handover is often caused by the current network that suddenly becomes unavailable. The mobile device has therefore to find and join a new network in the fastest possible way, in order to reduce the period of service interruption.

To ease the handover process, different access systems such as cellular, wireless LANs, short-range devices, and even wired access (e.g. Ethernet, xDSL) should be coordinated to deal efficiently with various environments and user demands. The current evolution of wireless networks toward all-IP architectures can offer a common platform for this harmonization. Solutions will be possible which completely integrate with natively IP-based wired networks. In this context, the support for terminal mobility can be handled by the Mobile IPv6 protocol (MIPv6, see [1]), which may work with any kind of application over any kind of access network.

Yet, MIPv6 specification and realizations suffer from some shortcomings. MIPv6 has been considered mainly for the horizontal handover, and it has become a valuable solution for global mobility. However, the latency introduced in the handover process may become a significant hurdle in a situation of micro mobility, such as office or home environments, where more stringent requirements are often in place.

The paper describes how we have addressed this issue, focusing in particular on reducing the handover decision time. We have at first analysed the performance of the current implementation of MIPv6 by means of a geographical testbed.

Manuscript received November 30, 2005; revised March 6, 2006.

Rosario G. Garroppo, Stefano Giordano, Stefano Lucetti, Giuseppe Risi, and Luca Tavanti are all with the Dept. of Information Engineering, University of Pisa, Via Caruso 16, 56122 Pisa, Italy.

E-mail: {r.garroppo, s.giordano, s.lucetti, g.risi, luca.tavanti}@iet.unipi.it

We have then designed and developed a software module, called IMM (an acronym for Interface Management Module), that has been integrated in the MIPv6 stack. The goal of IMM is extending the Mobile IPv6 to better support interface selection and movement detection. The same geographical testbed has been used to evaluate the enhancements brought by our solution.

The paper is organised as follows. Section II outlines recent work in the field of vertical handover. Section III describes and characterizes the handover process, including a short description of the Mobile IPv6 protocol. The proposed solution is presented in Section IV, followed by a description of the experimental testbed in Section V. Section VI reports the outcome of our tests, and finally Section VII concludes the paper.

II. RELATED WORK

The interoperability of cellular networks and wireless LANs has recently become a hot topic for standardization bodies such as 3GPP, ETSI and IETF. The European Telecommunications Standards Institute (ETSI) first introduced a classification of the coupling between a WLAN and a cellular network [2]. When *loosely coupled*, the two networks are independent of each other and only connected through the global Internet for exchanging signalling data. User data transit over separate paths. In *tightly coupled* configurations, the WLAN is just an extension of the cellular radio access subsystem. Therefore it must implement and support most of the features of the cellular network. Both signalling and user data transit over the core of the cellular network. Each approach has its pros and cons. In the present paper we focus on the loosely coupled solution, which we believe is more general, given that in most cases WLANs and cellular networks are managed by different operators. Note that in such a scenario the procedures of authentication, authorization and accounting also play an important role. Still, from the point of view of the Mobile IPv6, these are activities of the upper layers, and therefore we did not insert them in our study.

The Third Generation Partnership Project (3GPP) is now studying how to exploit WLAN access to enhance the services offered to users of 3G networks. In a recent specification [3], 3GPP plans to reuse the functionalities of the cellular network (e.g. subscription, authentication, authorization, accounting, routing) in the WLAN. This will be achieved relying solely on the existing capabilities of the WLAN, without imposing new constraints on it.

The Session Initiation Protocol (SIP) has been proposed for the provision of seamless handover for real-time services. Both 3GPP and the Internet Engineering Task Force (IETF) have converged on it. However the two proposals do not use the same features of SIP, hence they pose compatibility problems [4]. In addition, 3GPP has not yet specified how to use SIP to provide service continuity. Other issues come from the fact that SIP works at application level and does not integrate with lower layer protocols. This may cause extra delays due to the traversal of the whole OSI protocol stack. Working at lower layers may thus result more effective.

Solutions at IP layer are mostly based on Mobile IP, either

in version 4 or 6. Clearly, Mobile IPv6 (MIPv6) looks more promising given the ongoing transition towards the IPv6 Internet. Several works have already evaluated the performance of this protocol and proposed enhancements. However, with specific regard to the issues of vertical handover, the number of contributions is still limited. Chakravorty et al. [5] carried out an experimental study on mobility between GPRS and 802.11 networks. The authors highlighted the problems raised by the disparity in the characteristics (in terms of bandwidth and round trip time) of the two networks. They also proposed and evaluated some optimization techniques (such as Fast Router Advertisements, Client-based Router Advertisement Caching, Client-Assisted Simulcast of Binding Updates) at network layer. For example, Fast Router Advertisement can minimise the execution time of the handover by increasing the frequency of router advertisements, but may also cause overhead problems in narrowband WWAN links. As the authors themselves point out, more optimizations are still possible, particularly in the direction of cross-layer solutions.

One such solution has been proposed by Ylianttila et al. [6]. The authors have defined a “transition region” as the area where the signal varies around a predefined threshold. They have then proposed two solutions that analyse how the effective data rate, the terminal speed, and the handover delay act on throughput when the mobile device enters or leaves the transition region. The analysis considers both the received signal strength (RSS) and the dwell time (the period during which the terminal uses a high data rate even though the RSS is below the threshold). An RSS-based algorithm compares the received signal strength with the threshold to decide to start the handover. The aim of this algorithm is to maximize the throughput in the transition region by choosing the link with the broadest bandwidth. Unfortunately, given the high variability of the RSS, this kind of algorithms induce the so called “ping-pong” effect. More handover processes are carried out, with a considerable increase in the latency.

III. THE VERTICAL HANDOVER PROCESS

In the present section we will give a short outline of the Mobile IPv6 protocol (the details can be found in [1]) and a characterization of the handover process. This characterization is by now rather common in literature (see e.g. [5][7]).

A. The Mobile IPv6 protocol

Mobile IPv6 has been introduced to support IP mobility of user devices, which in this context are named Mobile Nodes (MNs). When a MN moves to another network, it sends Binding Update (BU) messages to its Home Agent (HA) and Correspondent Nodes (CNs). The CNs are nodes the MN is communicating with. When they receive a BU, they should update the path over which the data packets are sent, in order to reach the MN in the new network. However, it may happen that a CN is not able to understand the BU messages (e.g. if it does not support Mobile IPv6). In such a case, data traffic is sent to the HA that redirect it to the MN. The task of the HA is in fact to trace all the MN movements in order to keep the continuity of the connections. These procedures are followed for both horizontal and vertical handovers.

There are two factors that make optimizing the handover procedure a hard task. The first is related to the implementation of Mobile IPv6, the other is intrinsic of the MIPv6 protocol.

MIPL is a free implementation of Mobile IPv6 developed at the Helsinki University of Technology [8]. It does not support any kind of dynamic decision strategy to select the interface to use for network connection. Each active interface has a statically assigned priority, which is set when the MIPL module is first loaded into the kernel. The interface with the highest priority is marked as preferred (or primary) and selected for communication. The procedure to assign priorities is pseudo-stochastic; in fact, successive tries may lead to different interface priority assignments, hence to a different primary interface. The assigned priorities depend neither on the characteristics of the interface (wired vs. wireless, nominal data rate, etc.) nor on its status (active or inactive). This rigid scheme clearly does not consent any efficient management of the interfaces.

Mobile IPv6 decides to perform an handover on the basis of the Neighbour Unreachability Detection (NUD) algorithm [11]. The MN sets a timeout of 5 seconds after receiving each unsolicited Router Advertisement (RA) on the link of the interface with the highest priority. The originator of those RAs is the MN's Default Router (DR). When the timeout expires, the MN sends a Neighbour Solicitation (NS) message to the DR. If it does not receive a Response Router Advertisement (RRA) in 200 ms, the MN assumes that the Default Router is no longer reachable and starts a survey on all its interfaces to acquire a new DR. This is performed sending Router Solicitation (RS) messages over all active interfaces and waiting for a Response RA. The interface that becomes the new primary interface is the one with the highest priority among those from where the MN has received a solicited Router Advertisement. Unfortunately, the delay introduced by this simple Layer 3 movement detection algorithm is pretty high and can easily spoil any technique for reducing the handover execution time.

B. Characterizing the vertical handover process

The vertical handover process in MIPv6 can be divided into two phases: *handover decision* and *handover execution*.

The first phase should produce a decision regarding both an "if" and a "when" the handover must be executed. The presence of overlapping networks should be exploited to always have the best connection. The user can choose his entry point in a heterogeneous network scenario with multiple points of access. However, moving to a different network should produce no noticeable service interruption. Hence, what networks are available and the quality of the links to the access points are factors that must be continuously monitored. In addition, user preferences, such as choosing the cheapest operator, or keeping the quality of the service at a desired level, must also be considered. All these data are the input of a decision algorithm that determines whether to perform an handover or not. An accurate choice of the quantity and quality of the data and the adoption of a smart algorithm are therefore the key to reducing the handover latency.

The handover execution phase encompasses all the

operations and signalling to attach the device to the new access router. It includes the registration procedures at the HA and with the CN, and the execution of the Return Routability procedure as well (registration of the Care-Of Address, CoA – see [1] for the details). It can be considered concluded when the first data packet is received through the new access router.

Both phases jointly contribute to the total handover latency. This time can alternatively be divided into three parts:

- *Detection Time (DT)*. It is the time required by the MN to become aware of its movement, or to detect physical events that may lead to a handover (e.g. entering/leaving the coverage area of an access point).

- *Configuration Time (CT)*. This measures the time the MN needs to find a new access router (by receiving Router Advertisements, RAs), to update its routing tables and to configure its interfaces with a new Care of Address (CoA). A CoA is a temporary address used to deliver the traffic to the MN registered in the new network.

- *Registration Time (RT)*. It is the time that elapses between the forwarding of a Binding Update to the HA and CN and the moment the first data packet is received from the CN over the new route. The Registration Time is typically bounded by the Round Trip Time (RTT) experienced towards the HA and CN. Depending on the link, it may be very high. The most common technique to optimize the RT is using the link with the lowest RTT (usually the WLAN link) for the registration phase. Optimizing the handover execution, however, is out of the scope of this paper.

IV. THE INTERFACE MANAGEMENT MODULE

To overcome these drawbacks, we have developed a kernel module for the dynamic management of the interfaces of the MN. This module, named Interface Management Module (IMM), works in cooperation with the Mobile IPv6 implementation, extending its functionalities to better support interface selection and movement detection. The goal is to optimize the handover decision phase by means of a cross-layer approach. As shown in Fig. 1, IMM communicates with both the Link Layer and the Mobile IPv6 stack. Information gathered from the Link Layer is used to drive the selection of the primary interface.

Similarly to MIPL, IMM also sorts the available interfaces, assigning a different priority to each of them. Modifications to Mobile IPv6 implementation are then minimal, since MIPL already performs interface selection based on priority ranking. In contrast to MIPL, however, IMM takes into account the features of each network interface and can dynamically change the priorities. This implies that information on channel quality or link state must be continuously available to assign each interface the proper priority.

The current implementation of IMM assumes the MN is in a three layer overlay network. A LAN (Ethernet) is the lowest overlay, which provides a high speed link but allows no movement. An 802.11b WLAN is the intermediate overlay. It offers a reasonable movement opportunity and a fairly good data rate. The GPRS cellular network is the highest overlay, with low data rate but geographical coverage. IMM then characterizes the WLAN interface with the signal to noise ratio (SNR) of the radio channel (collected every 100 ms), and

the Ethernet link simply by its status (either up or down). The GPRS link is assumed to be always available, so there is no need to monitor it.

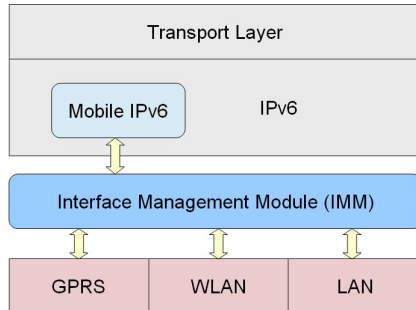


Fig. 1. Architecture of the IMM.

In the design of IMM, we have chosen to favour the fastest links as long as they are available. This implies that the lowest overlay layers should always be preferred. Thus, if all the interfaces are active, the Ethernet link has the highest priority, the 802.11 link the intermediate, and the GPRS link the lowest. The priority assignment algorithm, which drives the handover decision in dynamic conditions, is illustrated in Fig. 2. To determine the priority for the WLAN interface, we used a hysteresis mechanism. The current SNR is compared with two thresholds: SNR_{HIGH} and SNR_{LOW} , with $SNR_{HIGH} > SNR_{LOW}$. If the current SNR becomes greater than SNR_{HIGH} , the link is assumed to be good and its priority is set higher than GPRS but lower than Ethernet. If SNR moves below SNR_{LOW} , the WLAN link is reckoned to be poor and the interface priority becomes the lowest among the active interfaces. The choice of employing two thresholds is due to the need of reducing the “ping-pong” effect. The priority level assigned to the GPRS interface is kept constant, as the link is always available. As for the Ethernet link, if it is up, it gets the highest priority. If the link is broken, the interface is removed from the list of active interfaces. Whenever the link is restored, the interface is put back in the active interface list and is configured with the highest priority.

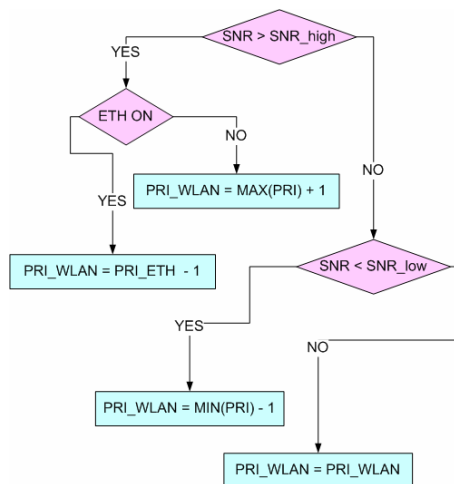


Fig. 2. The algorithm of priority updating.

When the state of any interface changes, the priorities are

updated. This may cause a change in the primary interface, hence an handover may be necessary. The handover process is started as soon as the MN, after updating the priorities, receives a Router Advertisement on an interface whose priority is higher than the currently preferred interface. With regard to the default MIPv6 procedure, employing the IMM no longer requires waiting for the timeout expiration of unsolicited RAs, and there is no need to transmit NS and RRA messages. Hence we expect that the proposed solution will significantly reduce the handover latency.

V. THE EXPERIMENTAL TESTBED

In our experimental testbed we have included a wired LAN, a wireless 802.11 network and wide area connectivity (WWAN). The testbed is depicted in Fig. 3. 6net is a native IPv6-based network built within a European project to test a variety of new IPv6 services and applications [9].

WWAN connectivity is provided by the GPRS (General Packet Radio Service) network operated by TIM, one of the four Italian cellular operators. Starting from the terminal, the GPRS network infrastructure includes a Base Transceiver Station (BTS), a Base Station Controller (BSC), the Service GPRS Support Node (SGSN), and eventually the Gateway GPRS Support Node (GGSN). The GGSN is responsible for acting as a gateway between the GPRS packet network and the public Internet. All these elements are property of the cellular operator and we had no control on them. Instead, the rest of the testbed has been deployed in our laboratory. It includes a number of desktop PCs, one IEEE 802.11 Access Point (AP) and a laptop acting as the MN. Three different IPv6 subnets have been configured. Subnets A and B are used to complete the overlay network, whereas Subnet C hosts the CN. Subnet A is the Home Network. Each PC is running a Linux GNU/Debian distribution; the MN, CN and HA run the MIPLv1.1 implementation of Mobile IPv6 [8].

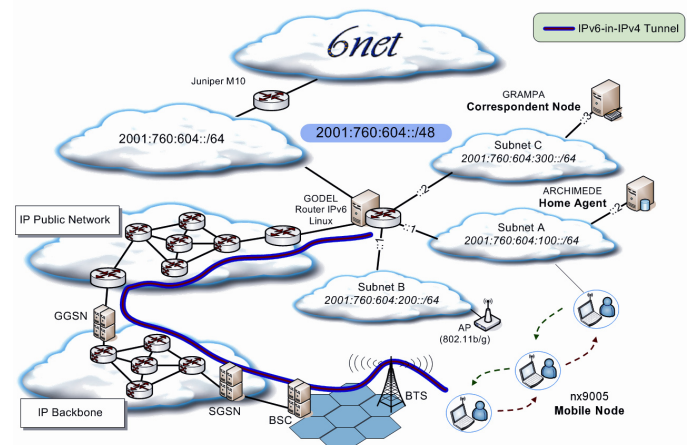


Fig. 3. The experimental testbed.

The PCs involved in the experiment are:

- *nx9005*. This is the Mobile Node. It is equipped with an Ethernet interface connected to Subnet A, a PCMCIA 802.11b interface to connect to the AP belonging to Subnet B, and a PCMCIA Globetrotter GPRS network card to connect to the GPRS network.

– *Granpa*. This is the CN. It represents a generic host of the Internet, and runs a number of services, among which an FTP server and the MGEN traffic generator.

– *Archimede*. It acts as the Home Agent for the MN. It is attached to Subnet A.

– *Godel*. It acts as an IPv6 router, and does not need any Mobile IPv6 knowledge.

Since the GPRS network currently supports only IPv4, we set up a static tunnel of IPv6 in IPv4 between our router (*Godel*) and the MN. All IPv6 traffic coming from the MN through the GPRS interface is encapsulated into IPv4 packets. Once they reach *Godel*, they are decapsulated. The same happens in the opposite direction for all the traffic addressed to the GPRS interface of the MN.

A. The Measurement Tool

The MAGNET [10] toolkit has been chosen to gather a detailed description of the active TCP sessions, to collect the statistics of the Mobile IPv6 protocol and to identify the messages exchanged during the three phases of the handover procedure. This measurement tool offers a framework to monitor any arbitrary kernel event of the Linux operating system and is available as a patch for the 2.4.19 kernel. An extension to MAGNET [11] has been necessary in order to support the 2.4.22 kernel, which is required by the MIPLv1.1 implementation. Furthermore, the TCP tracking engine has required a further extension in order to export all the TCP related variables and statistics, such as the Congestion Window Size, the Retransmission Time Out, the Smoothed Round Trip Time, as well as all the events related to the Mobile IPv6 protocol.

VI. EXPERIMENTAL ANALYSIS

To evaluate the improvements of the proposed IMM over the plain Mobile IPv6, we have performed a number of tests using both UDP and TCP traffic. Within the reference scenario of Fig. 3, the MN moves across the overlay network triggering the handovers. At first, the MN is connected to its Home Network (Subnet A) through the Ethernet interface. At the same time it is also under the coverage area of both the WLAN and the GPRS networks. Next, it gets disconnected from Subnet A. This forces a first handover towards the WLAN access network (Subnet B). Lastly, the MN is moved away from the AP, thus leading to a second handover towards the GPRS network. The MN is then moved along the inverse path, to perform two more handovers, the first from the GPRS network to the WLAN, and the second from the WLAN to Ethernet, when it is finally reconnected to its Home Network.

The modified version of MAGNET has been used to obtain all TCP parameters and to collect the exact timestamp of each UDP packet along with its sequence number. It has also been used to track all MIPv6 messages related to the handover procedure. These have been reported in the figures as vertical lines. For the sake of clarity, however, we have drawn only the most meaningful, omitting for example the messages of the Return Routability and Care-of Address registration procedures (Home Test Init, Care-of Test Init, Home Test, Care-of Test, etc.).

The handovers are considered completed when the MN

sends the Binding Update (BU) message to the CN (in same cases when the data transfer is resumed). This triggers the execution of the Route Optimization procedure and the subsequent routing of the packets along the new route.

Note that in most cases the duration of the experiments based on the MIPL and IMM implementations are different (since MIPL often takes longer) and the points of occurrence of the handovers are not the same. Hence the scales on the horizontal axis of many of the following figures are different too.

A. TCP traffic

The TCP session consisted in downloading a 50 MB data file from the Correspondent Node towards the MN. Statistics on the TCP connection have been measured both at the sender side (the CN) and at the receiver side (the MN). Each handover is detailed and commented in the following subsections.

A.1. Ethernet to WLAN Handover

The first handover involves a passage from the wired connection to the wireless LAN. This is obtained by simply unplugging the cable from the mobile device. Thus we can measure how promptly the mobile procedures react to a sudden change in the network state.

When the plain MIPL is in place (see Fig. 4), we can see how the default timeout of the NUD procedure causes a long service interruption. The last RA is received shortly after 1 second from the beginning of the test, and the device is disconnected from the Ethernet at 1.6 s. The Router Solicitation is sent out at 6.3 s, and the data transfer is resumed only at 8.1 s, after the registration of the mobile device with the CN. We can then compute the total handover duration time, that is 6.5 seconds.

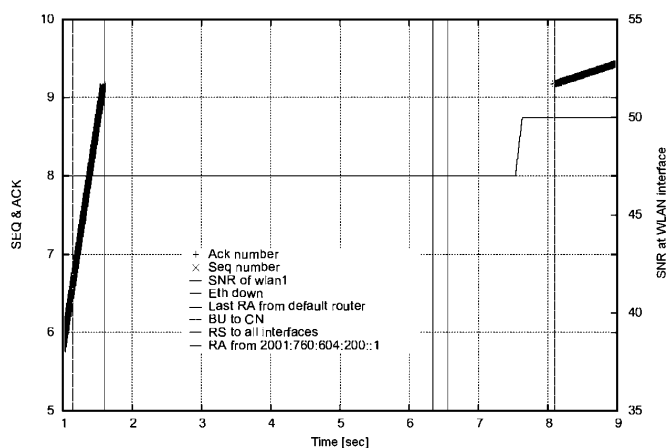


Fig. 4. Relevant TCP traffic data and SNR at the MN for the Ethernet to WLAN handover with MIPL.

The improvement brought by IMM is highlighted in Fig. 5 (note that the scale on the abscissa is different from Fig. 4). The device is notified of the lost availability of the wired connection (at 1.95 s). It immediately reacts by sending an RS over all interfaces. At 2.5 s it receives a RA from the WLAN interface so it can start the configuration and registration

phases, that finally lead to resuming the data transfer at 5.2 s. The overall length of the handover is reduced to 3.3 seconds, most of which employed by the registration procedure.

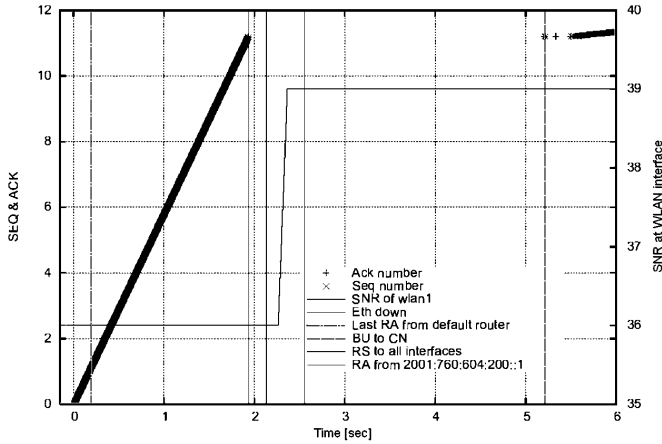


Fig. 5. Relevant TCP traffic data and SNR at the MN for the Ethernet to WLAN handover with IMM.

A.2. WLAN to GPRS Handover

As in the previous case, when employing MIPL, the WLAN interface is used until the timeout on the receipt of a Router Advertisement transmitted by the Default Router in Subnet B expires. Analysing Fig. 6, we can see that the SNR assumes increasingly lower values, thus leading to many packet losses and retransmissions. This has an impact on the CN as well. Fig. 7 reports traffic data registered at the CN. In particular, some points in the interval 43-62 s, that refer to transmitted TCP segments, have no corresponding ACK. Those data packets, in fact, had never been received by the MN (see Fig. 6, where there are no data points between 43 and 62 seconds). This leads to the expiration and subsequent increase of the retransmission timeout (RTO), as can be seen in Fig. 8, that plots the values of all relevant TCP parameters at the CN (note that SRTT and RTO are plotted on a logarithmic scale, while the grid refers to the linear scale of CWND and Ssthresh).

The handover process, however, is not started until 49.5 s, when a first Router Solicitation is cast by the MN on every interface. Unfortunately an RA is received over the WLAN, which is then kept as the primary interface even if no data can flow through it due to the very low SNR.

This behaviour finds an explanation in the IEEE 802.11b standard [14]. It should be noted that RAs have a multicast destination address. According to the standard, the AP maps this packet to a broadcast MAC frame and decreases the transmission bit rate (to 2 Mbps) for backward compatibility with non-b/g cards. This allows the RAs to be received with higher probability than data packets, which are unicast and therefore transmitted by the AP at the highest rate and with a less robust modulation.

The expiration of the second RA timeout is thus necessary to restart the procedure and bring to a complete handover. The Binding Update is thus sent only at 55.5 s, and the file transfer is resumed through the GPRS networks at 62 s. Hence we can assume the handover decision time, i.e. the time to understand that a handover is needed, to go from the last received packet

(at 43 s) to 49.5 seconds, and the handover execution phase to last 12.5 seconds. Summarizing, the whole handover process using the plain MIPv6 protocol takes about 19 seconds. During this period the file transfer is suspended, five retransmissions are necessary to deliver a packet to the MN, and TCP needs roughly 3 more seconds to reach a regime over the GPRS network.

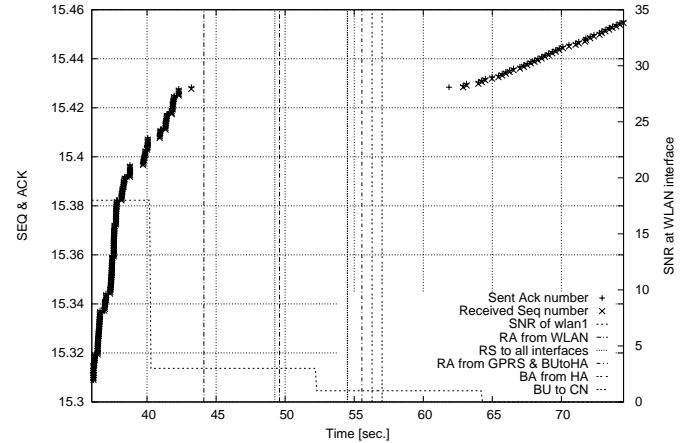


Fig. 6. Relevant TCP traffic data and SNR at the MN for the WLAN to GPRS handover with MIPL.

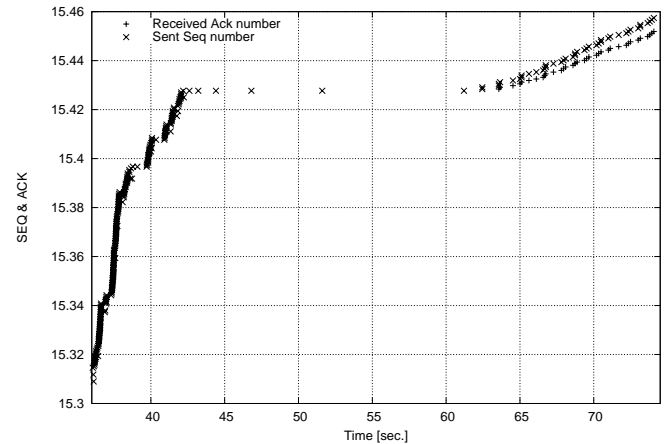


Fig. 7. Relevant TCP traffic data at the CN for the WLAN to GPRS Handover with MIPL.

Fig. 9, Fig. 10 and Fig. 11 refer to the handover performed using the proposed IMM. The first noticeable thing is that the data transfer is interrupted for a very short time. The WLAN interface is used while the measured SNR is higher than the SNR_{LOW} threshold. As soon as this condition becomes false (at 36.6 s), IMM marks the GPRS interface as preferred. So, when the MN receives an RA message through the GPRS network, it decides to start an handover (at 38.8 s), which is completed in less than 2 seconds.

Note that when the MN switches to the GPRS network, the CN does not receive immediately the ACK packets (refer to Fig. 10, at 38.8 s). Hence, when the RTO expires, the CN retransmits the first unacknowledged packet and the Contention Window parameter (CWND) is set to 1 (Fig. 11, time 39.1 s). When the CN receives an ACK by means of the new connection, the CWND is increased and the file transfer is recovered.

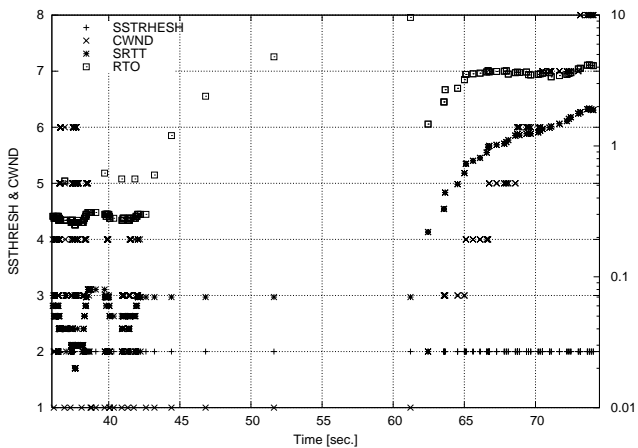


Fig. 8. TCP parameters at the CN for the WLAN to GPRS Handover with MIPL.

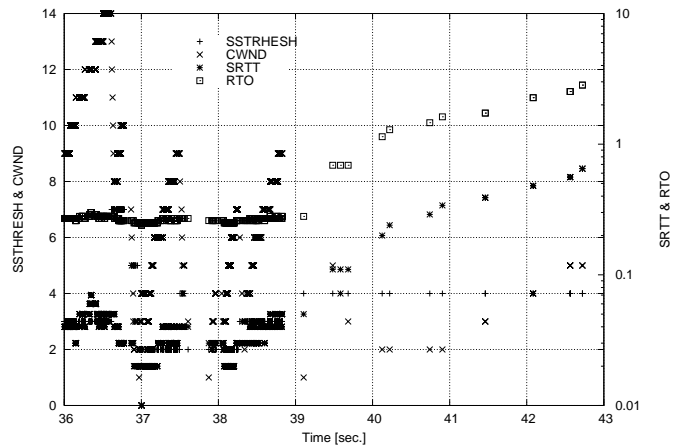


Fig. 11. TCP parameters at the CN for the WLAN to GPRS handover with IMM.

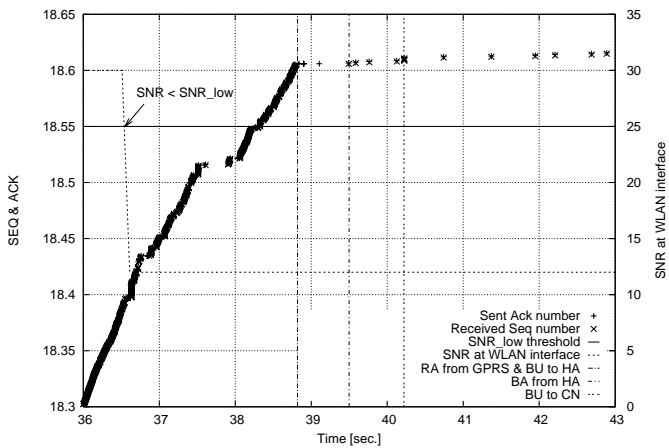


Fig. 9. Relevant TCP traffic data and SNR at the MN for the WLAN to GPRS handover with IMM.

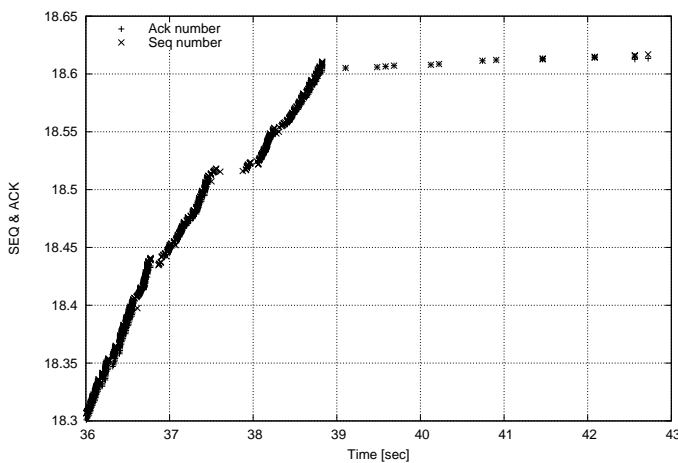


Fig. 10. Relevant TCP traffic data at the CN for the WLAN to GPRS handover with IMM.

A.3. GPRS to WLAN Handover

The behaviour of the standard MIPL implementation is reported in Fig. 12. The MN performs the handover as soon as it receives a Router Advertisement over the WLAN interface. However, the link quality is not yet good enough to sustain the communication. Hence, after a few sent TCP ACKs, the data transfer is suspended. It takes about 10 seconds to reach a zone where the SNR is sufficient to correctly receive at least some packets. The overall handover time can therefore be around 18 seconds. Note that in this case, assuming that casting a BU to the CN concludes the handover is not correct, as no data is moved after that event. Rather a service interruption occurs, and therefore we must consider the interruption time as the handover duration.

When using IMM (see Fig. 13), the handover can only be triggered after the SNR at the MN has become higher than the SNR_{HIGH} threshold (for the moment, IMM just marks the WLAN as preferred). The link quality towards the 802.11 AP is then good enough to guarantee the delivery of the TCP packets. So, when a RA is received on the WLAN interface, the configuration and registration procedures are started. They are then completed in a very short time thanks to the good quality and speed of the new interface. Note that during the handover packets are received on both wireless interfaces. This realises a soft handover which allows to keep the service interruption at a minimum. From the user perspective no interruption is even experienced.

A.4. WLAN to Ethernet Handover

In this last case, the handover performs well with and without the IMM module. In fact, as soon as the MN gets connected to the cabled LAN, the receipt of a Router Advertisement on this link triggers the handover procedure, which concludes very quickly and causes no service interruption or slowing down. For the sake of brevity, we do not include any figure for this test, given the strong similarity of the results for the two approaches.

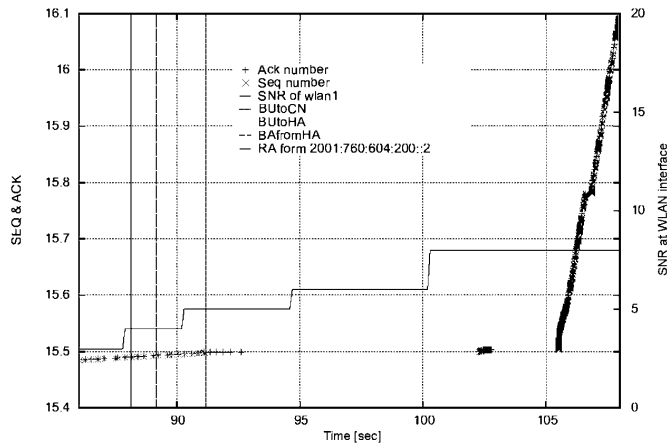


Fig. 12. Relevant TCP traffic data and SNR at the MN for the GPRS to WLAN handover with MIPL.

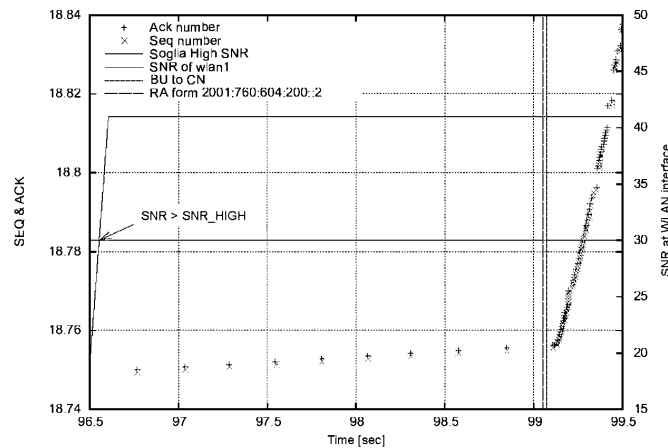


Fig. 13. Relevant TCP traffic data and SNR at the MN for the GPRS to WLAN handover with IMM.

B. UDP traffic

The impact of IMM on UDP traffic has been analysed using the same experimental testbed shown in Fig. 3. As in the TCP test, the Mobile Node moves in the overlay networks triggering four vertical handovers. An UDP session has been set up with packets generated at the Correspondent Node (using the MGEN [13] traffic generator) and directed to the MN. To avoid losses in the GPRS network, we have limited the rate of the traffic to a value lower than the GPRS available throughput. In the environmental conditions of the experiments, we managed to get a sustained data rate of about 30 kbps through the GPRS network. MGEN has consequently been configured to generate 30 packets per second, with a payload size of 64 bytes. Adding the IPv6 and UDP headers, this results in an offered traffic data rate of about 27 kbps.

The patterns of the received UDP packets at the MN along with the cumulative number of lost packets are plotted in Fig. 14 and Fig. 15, using respectively the MIPL implementation and the IMM. It is immediately evident how MIPL causes more packet losses and performs longer handovers than IMM.

The general behaviour of the two implementations during the four handovers was very similar to what has already been described in the previous Section. What makes this series of tests different, is the distinct nature and response of the two

transport protocols. Therefore, in this section we focus only on the most evident differences with TCP.

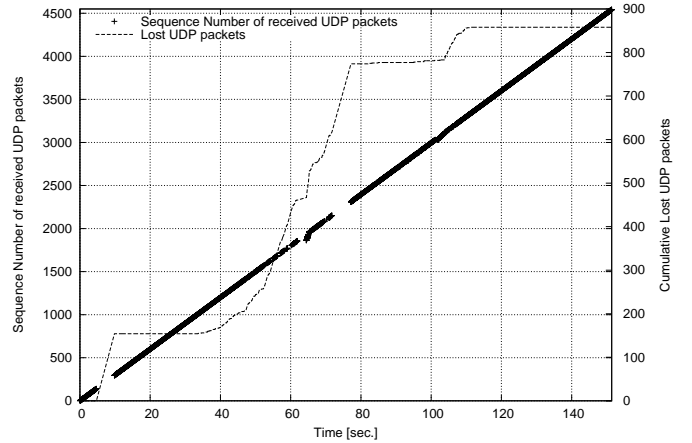


Fig. 14. Received and lost UDP packets at the MN with MIPL.

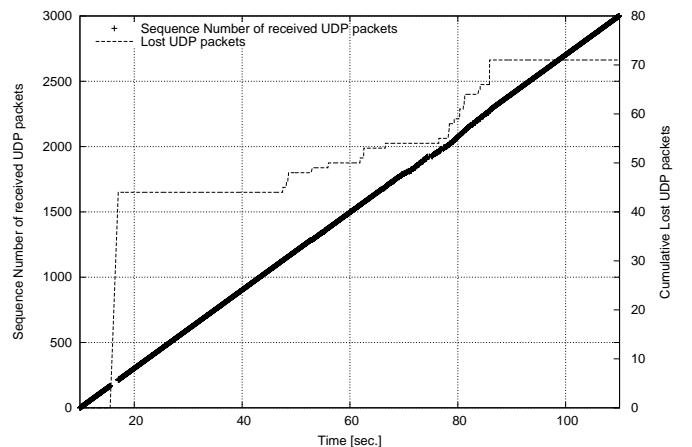


Fig. 15. Received and lost UDP packets at the MN with IMM.

B.1. Ethernet to WLAN Handover

Fig. 16 and Fig. 17 are the close-ups of Fig. 14 and Fig. 15 in the proximity of the first handover, from the Ethernet to the 802.11 network.

As for the TCP test, the NUD procedure prevents the MN to promptly react to the loss of connectivity created to the disconnection of the cable at 4.7 s. The first reaction occurs at around 8.6 s, when a Router Solicitation is cast over all the interfaces as a consequence of the expired timeout. The BU completes the handover with a total delay of about 5.2 seconds. During this period, about 160 UDP packets are lost.

When the IMM is in use (see Fig. 17), the loss of the Ethernet connection (at 15.6 s) is notified to the Mobile IPv6 engine, causing the Default Router to be set as unreachable. The MN quickly sends an RS message and after a few hundreds of milliseconds it receives an RA on the WLAN interface. The Care-of Address registration procedure is then started and at 17 s the MN is able to receive packets from the CN. The total handover latency is reduced to about 1.5 seconds, leading to a loss of only 44 packets.

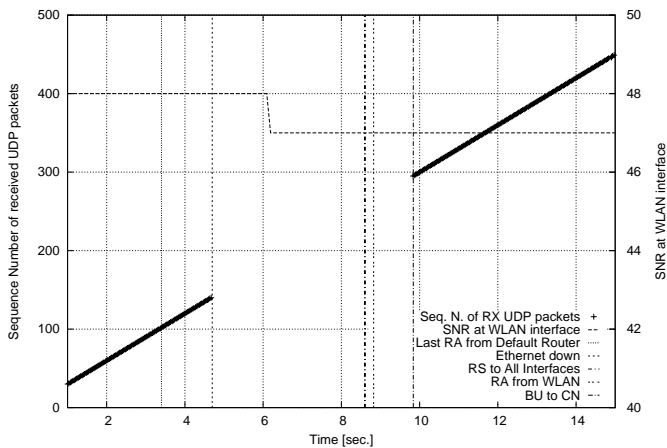


Fig. 16. Received UDP packets and SNR at the MN for the LAN to WLAN handover with MIPL.

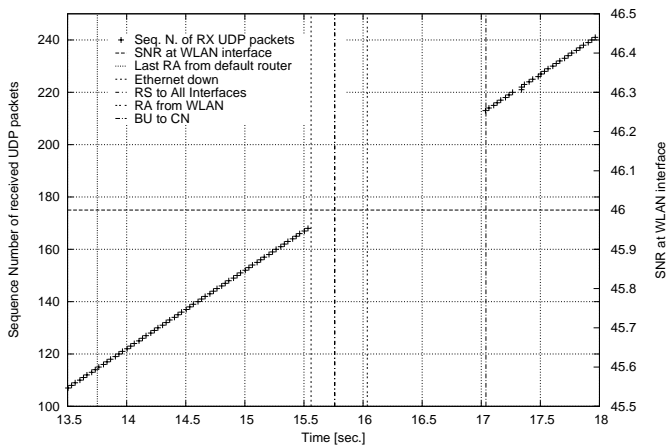


Fig. 17. Received UDP packets and SNR at the MN for the LAN to WLAN handover with IMM.

B.2. WLAN to GPRS Handover

After the completion of the first handover, the MN starts moving away from the WLAN coverage area. Fig. 18 shows the performance of the MIPL for this kind of handover. Note that the SNR falls to very low values. UDP packets sent by the CN start being corrupted and get lost. As no RAs are received, the NUD procedure forces the MN to send RS messages. Unluckily (as for the TCP-based experiment), the solicited Router Advertisements are received from both the WLAN and the GPRS interfaces. Consequently the MN remains in Subnet B (i.e. on WLAN). Only when channel conditions further degrades (leading to an almost complete loss of packets, after 70 s), does the handover take place. The overall handover duration is then more than 25 seconds, causing about 500 packets to be lost.

On the other hand, the adoption of the IMM can even allow an almost seamless handover, as shown in Fig. 19. As soon as the SNR at the MN falls below the SNR_{LOW} threshold (at 49.2 s), the priority of the WLAN interface is lowered under that of the GPRS interface. Consequently, when the first unsolicited RA is received from the GPRS network, the handover execution is triggered. The handover latency is reduced to the duration of the Care-of Address registration,

and only 6 packets are lost. The short interruption of packet reception around 53 s is not related to the handover, but is due to a temporarily low quality of the GPRS link.

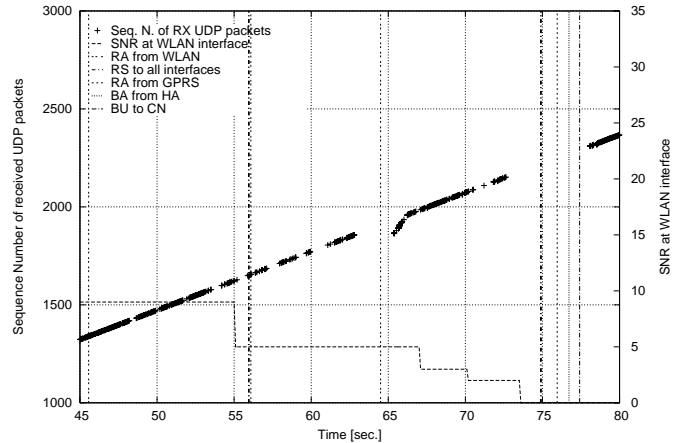


Fig. 18. Received UDP packets and SNR at the MN for the WLAN to GPRS handover with MIPL.

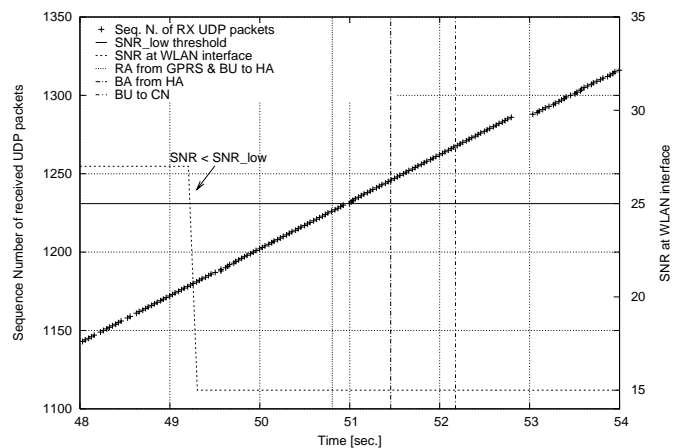


Fig. 19. Received UDP packets and SNR at the MN for the WLAN to GPRS handover with IMM.

B.3. GPRS to WLAN Handover

The behaviour of the standard MIPL implementation is reported in Fig. 20. The MN performs the handover as soon as it receives a Router Advertisement over the WLAN. However, the link quality is not yet good enough to sustain the communication. Hence several (about 80) packets are lost. In the specific try, it takes about 4 seconds to reach a zone where the SNR is sufficient to correctly receive all packets. It is noteworthy that in another measurement run (not reported), the SNR after the handover was so bad that the MN performed another handover towards the GPRS. The overall handover time reached over 20 seconds.

When using the IMM (see Fig. 21), the handover is triggered after the SNR at the MN has become higher than the SNR_{HIGH} threshold. The link quality towards the 802.11 AP is then good enough to guarantee the delivery of the streamed UDP packets. Since packets are received on both wireless interfaces, a soft handover is performed and packet losses are minimal. Only 5 packets are lost, including those that reach

the MN out of order, because of the different delay of the GPRS access network.

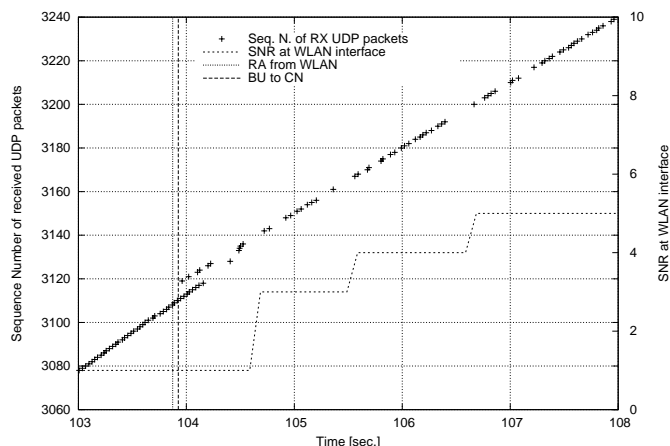


Fig. 20. Received UDP packets and SNR at the MN for the GPRS to WLAN handover with MIPL.

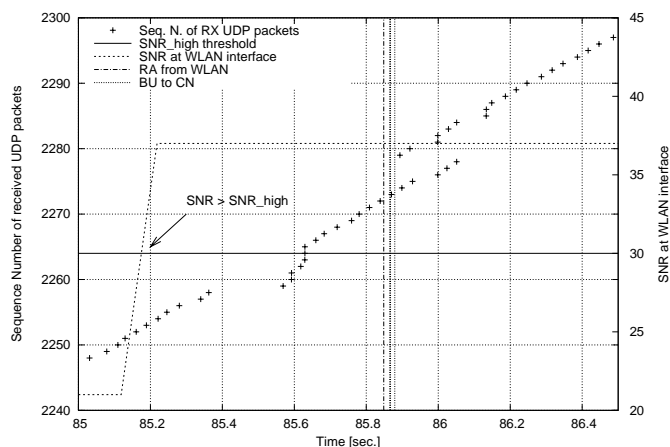


Fig. 21. Received UDP packets and SNR at the MN for the GPRS to WLAN handover with IMM.

Also note that the MN sends the Binding Update message to the CN twice, since it receives a packet from the CN through the GPRS network after having sent the first BU. This is due to relevant delay introduced by the GPRS network. Consequently there are some UDP packets “on-air” that reach the MN after the handover has successfully concluded.

B.4. WLAN to Ethernet Handover

As for the TCP test, the handover performs well with both implementations. Hence there is no need to report any figure.

VII. CONCLUSIONS

We have set up an experimental testbed to analyse the performance of the Mobile IPv6 protocol in heterogeneous wireless networks. In particular, we have focused on the issues raised by the vertical handover. From the analysis, we have found that the main cause of the excessive latency is the handover decision. In the standard MIPL implementation, the simple Layer 3 Neighbour Unreachability Detection (NUD) algorithm is the entity that triggers the execution of the

handover. The NUD procedure, however, has been designed for static nodes, and does not allow a prompt reaction to sudden changes in the state of the MN interfaces.

Therefore, we have developed a cross layer mechanism, the Interface Management Module (IMM), which takes advantage of the information on the state of all the available network connections. The proposed module has been evaluated in a geographical test-bed and has achieved a noteworthy reduction of the handover latency. This subsequently leads to decreasing packet losses in presence of UDP traffic and to a quicker recovery of TCP sessions. Additionally, with regard to the WLAN-to-cellular handover, inserting an hysteresis on the levels of SNR that trigger the handover has allowed to avoid the “ping-pong” effect that afflicts the standard MIPv6 handover procedure.

We can therefore conclude that the use of dynamic, state-dependent interface selection algorithms can dramatically enhance the performance of the MIPv6 vertical handover. By realizing a software module that implements such algorithms (the IMM), we have proved that this is a feasible idea, that supports both TCP and UDP traffic. Moreover, if sufficient intelligence is deployed in the IMM, phenomena like the “ping-pong” effect can be avoided. Also, the proposed solution offers enough flexibility to account for more information than the simple interface state. For example user preferences (cost, QoS requirements, etc) can be easily included in the software.

REFERENCES

- [1] D. Johnson, C. Perkins, J. Arkko, “Mobility support in IPv6”, IETF RFC 3775, June 2004.
- [2] ETSI TR 101 957, “Requirements and Architectures for Interworking between HIPERLAN/3 and 3rd Generation Cellular Systems”, August 2001.
- [3] 3GPP, “Group Services and System Aspects; 3GPP Systems to Wireless Local Area Network (WLAN) Interworking; System Description (Release 6)”, TS 23.234, v.6.2.0, Sept. 2004.
- [4] F.G. Marquez, M.G. Rodriguez, T.R. Valladares, T. de Miguel, L.A. Galindo, “Interworking of IP multimedia core networks between 3GPP and WLAN”, *IEEE Wireless Communications*, Vol. 12, Issue 3, June 2005.
- [5] R. Chakravorty, P. Vidales, K. Subramanian, I. Pratt, J. Crowcroft, “Performance Issues with Vertical Handovers: Experiences from GPRS Cellular and WLAN hot-spots Integration”, *Proceedings of IEEE PerCom*, March 2004.
- [6] M. Ylianttila, J. Mäkelä, K. Pahlavan, “Analysis of handoffs in a location-aware vertical multi-access network”, *Computer Networks*, Vol. 47, Issue 2, Feb. 2005.
- [7] M. Bernaschi, F. Cacace, G. Iannello, S. Za, A. Pescape, “Seamless internetworking of WLANs and cellular networks: architecture and performance issues in a Mobile IPv6 scenario”, *IEEE Wireless Communications*, Vol. 12, Issue 3, June 2005.
- [8] MIPL Mobile IPv6 for Linux, available at: <http://www.mobile-ipv6.org>.
- [9] 6net – Large Scale International IPv6 Pilot Network, available at: <http://www.6net.org>.
- [10] Monitoring Apparatus for General kerNel Event Tracing (MAGNET). Available at: <http://public.lanl.gov/radiant/research/measurement/magnet.html>.
- [11] MAGNeT+, available at: <http://netgroup-serv.iet.unipi.it/magnet>.
- [12] T. Narten, E. Nordmark, W. Simpson, “Neighbor Discovery for IP Version 6 (IPv6)”, IETF 2461, December 1998.

- [13] The Multi-Generator (MGEN), available at: <http://pf.itd.nrl.navy.mil/mgen/mgen.html>.
- [14] IEEE 802.11b-1999, "Wireless LAN MAC and PHY specifications: Higher speed Physical Layer (PHY) extension in the 2.4 GHz band".



Rosario G. Garroppo received the Laurea (M.Sc.) degree cum laude in Telecommunication Engineering in 1995 and the Ph.D. degree in Information Engineering in 1999 from the University of Pisa. Since December 2001, he is Assistant Professor at the University of Pisa. His research interests are in traffic modeling, QoS guarantee in mobile networks, performance evaluation of telecommunication networks by means of simulation tools, traffic control in 802.11, and wireless sensors networks.



Stefano Giordano received the Laurea (M.Sc.) degree cum laude in Electronics Engineering in 1990 and the Ph.D. degree in Information Engineering in 1994, both from the University of Pisa. Since 2001 he is Associate Professor at the Department of Information Engineering of the University of Pisa (Telecommunication Networks Group) where he gives lectures in "Telecommunication Networks" and "Teletraffic Engineering". His research and professional interest is the area of broadband multimedia communications.



Stefano Lucetti received his M.Sc. (Laurea) degree cum laude in Telecommunication Engineering in 1999 and his Ph.D. in Information Engineering in 2003, both from the University of Pisa. He is now a postdoctoral fellow with the Telecommunication Networks Group of the University of Pisa. He has worked on teletraffic engineering and efficient radio resource management in cellular networks. Currently, his research interests are mainly focused on mobility support in IPv6 networks and QoS issues in 802.11 wireless networks.



Giuseppe Risi received the M.Sc. degree in Telecommunication Engineering from the University of Pisa in 2001. He is now a research assistant with the Telecommunication Networks Group of the University of Pisa. His research interest is in the areas of mobility support in IPv6 networks and Voice over IP.



Luca Tavanti received the M.Sc. degree in Telecommunication Engineering from the University of Pisa in 2000. He has then spent six months at the BT Labs in Ipswich (UK), and two years with Simtel spa (Florence, Italy), working on monitoring of cellular networks. Since January 2004 he is a Ph.D. student at the Dept. of Information Engineering of the University of Pisa. He is actively working in the area of wireless networks, mainly focusing on QoS provision and traffic control in WLANs, WMANs and ad-hoc networks.