The costs and benefits of combining different IP mobility standards

Antonio de la Oliva[a], Ignacio Soto[b,*], Maria Calderon[a], Carlos J. Bernardos[a], M. Isabel Sanchez[a,c]

[a]*Departamento de Ingeniería Telemática, Universidad Carlos III de Madrid*
[b]*Departamento de Ingeniería de Sistemas Telemáticos, Universidad Politécnica de Madrid*
[c]*Institute IMDEA Networks*

Abstract

Triggered by the demand of ubiquitous Internet connectivity and the availability of different wireless technologies, several IP mobility support protocols have been standardized in the past. Each solution provides a specific functionality (e.g., host or network mobility) and/or requires operations of particular nodes (e.g., client or network based). The current trend is towards the co-existence of these solutions, though the impact of doing so has not been yet fully understood. This article briefly reviews key standards for providing IP mobility support, identifying scenarios where combining them is necessary. We analyze the functionality offered by each combination and its performance cost in terms of protocol overhead and handover latency, highlighting the associated benefits and costs. This analysis is complemented by an experimental evaluation that supports our findings. The conclusions of our study indicate that combining different mobility mechanisms has a non-negligible impact on both the communication overhead and handover latency. This highlights the need for developing techniques to alleviate the costs of the combination. The recently proposed Distributed Mobility Management scheme exhibits some interesting properties that may help solving some of the identified combination shortcomings, as analyzed in this paper.

Keywords: IP Mobility, Network-based Localized Mobility, Host Mobility, Network Mobility, Distributed Mobility Management

*Corresponding author. Tel: +34 915495700 ext. 3057; fax: +34 913367333

Email addresses: aoliva@it.uc3m.es (Antonio de la Oliva), isoto@dit.upm.es (Ignacio Soto), maria@it.uc3m.es (Maria Calderon), cjbc@it.uc3m.es (Carlos J. Bernardos), mariaisabel.sanchez@imdea.org (M. Isabel Sanchez)

1. Introduction

Users demand Internet access everywhere and anytime. Current smart-phones and hand-held devices are equipped with multiple technologies – e.g., 3G and IEEE 802.11 Wireless LAN (WLAN) – as a solution to provide ubiquitous Internet access. The Internet Protocol (IP) constitutes the common building block for the provision of voice and data services, independent of the access technology. Users' mobility has triggered the need for new IP mobility mechanisms that enable terminals to move and change their point of attachment without affecting to their connectivity or the applications' behavior. Driven by the requirements posed by the different scenarios where connectivity is demanded, the Internet Engineering Task Force (IETF)[1] has standardized several IP mobility solutions. Initially, global mobility protocols were designed to allow seamless roaming (keeping global reachability and session continuity) within the whole Internet, both for single mobile hosts (e.g., Mobile IPv6 [1]) and moving networks (e.g., Network Mobility Basic Support [2]). More recently there has been an increasing interest in solutions that provide mobility support within a part of the network by means of functionality residing only on the network infrastructure, what is called network-based localized mobility. Proxy Mobile IPv6 [3] enables a mobile host to roam within a local domain without changing its IP address. Another example is N-PMIPv6 [4], a proposal that enhances Proxy Mobile IPv6 to better integrate mobile networks, by fully supporting terminals to roam between fixed and mobile access routers.

All these IP mobility protocols (which are briefly explained in Section 2) are going to co-exist because each of them addresses different requirements. For example, the 3GPP[2] specification [5] deals with the integration in its mobile architecture of access networks not based on 3GPP technologies. To support the mobility of a terminal, this specification considers the possibility of choosing between two types of IP mobility protocols: network-based or with mobility functionality in the terminal. The decision is taken depending on the capabilities of the network and the terminal. Selecting the most suitable mobility protocol among those available is the goal of the defined mechanisms in current 3GPP specifications, but a next step going beyond that may be required, as these protocols (and others men-

[1] http://www.ietf.org/
[2] http://www.3gpp.org/

tioned above related with network mobility) could also be used simultaneously in certain scenarios to maximize the overall performance or provide additional functionality. So it is interesting not only to choose among mobility protocols, as in the 3GPP specification, but also to be able to decide to combine them. To enable this, we need to study the different aspects involved in the combinations of IP mobility protocols. This article studies these combinations, highlighting the functionality they provide, the scenarios in which their deployment make sense, and the support required from the network and the user terminal (Section 3). We also analytically characterize the associated costs of the combinations – in terms of protocol overhead and handover latency (Section 4). In order to validate the theoretical results, we have performed also an experimental evaluation (Section 5). From this analysis we conclude that the combination of different mobility solutions has a cost that must be considered when designing mobility solutions that combine different protocols. Finally the article outlines an approach to solve the problem of efficiently combining different mobility support mechanisms in a general way (Section 6). This approach is based on generalizing the characteristics of a particular instantiation of the Distributed Mobility Management concept, a recent trend to manage mobility that is under study by the IETF.

2. Overview of IP Mobility Protocols

2.1. Mobile IPv6 and NEMO Basic Support protocol

Mobile IPv6 (MIPv6) [1] enables global reachability and session continuity by introducing the Home Agent (HA): an entity located at the home network of the Mobile Node (MN) which anchors the permanent IP address used by the mobile node, called Home Address (HoA). The home agent (see Figure 1) is in charge of defending the permanent IP address of the mobile node when the mobile node is not at home, and redirecting received traffic to the current location of the mobile node. When away from its home network, the mobile node acquires a temporal IP address from the visited network – called Care-of Address (CoA) – and informs the home agent about its current location. An IP bi-directional tunnel between the mobile node and the home agent is then used to redirect traffic from and to the mobile node. There is also optional support to avoid this suboptimal routing and enable the mobile node to directly exchange traffic with its communication peers – called Correspondent Nodes (CNs) – without traversing the home agent. This additional support is called Route Optimization (RO), and allows the mobile node to also inform a correspondent node about its current location. MIPv6 is a

client-based solution because the mobile node is required to perform specific IP procedures to support its own mobility.

The Network Mobility Basic Support (NEMO B.S.) protocol [2] extends MIPv6 to support the movement of a whole network (also referred to as a NEMO or mobile network), by the router of the network – called Mobile Router (MR) – taking care of the mobility management (i.e., mobility signaling and tunnel setup) of the network on behalf of the nodes of the network, called Mobile Network Nodes (MNNs). The IP addresses of these nodes belong to the Mobile Network Prefix (MNP) of the NEMO that is anchored at the home agent of the mobile router. There is no route optimization support standardized for NEMO B.S. Regarding mobility, the NEMO B.S. is a client-based solution as well, because it is also based on mobility functionality in the mobile node, a router in this case.

2.2. Proxy Mobile IPv6

Proxy Mobile IPv6 (PMIPv6) [3] is a network-based localized mobility management protocol. This means that user terminals are provided with mobility support without their involvement in the mobility management and signaling, as the required functionality is relocated from the mobile node to the network. In particular, movement detection and signaling operations are performed by a new functional entity – called Mobile Access Gateway (MAG) – which usually resides on the Access Router (AR) for the terminal (see Figure 1). In a Localized Mobility Domain (LMD), which is the area where the network provides mobility support, there are multiple MAGs. A MAG learns through standard terminal operations (such as router and neighbor discovery) or by means of link-layer support about the movement of a mobile node and coordinates routing state updates without any IP mobility support from the terminal.

Terminals are assigned a particular IP prefix in the localized domain, which remains to be the same even when they connect to a different MAG, so a terminal does not change address while visiting the domain. The IP prefixes used by the terminals are anchored at an entity called the Local Mobility Anchor (LMA), which plays the role of a local home agent, although required signaling exchanges are performed with MAGs, instead of with the user terminals. Inside the localized domain, the traffic of terminals is transferred using bi-directional tunnels between the LMA and the MAGs, so terminals can keep the same IP address without affecting the routing of their traffic. Proxy Mobile IPv6 is based on an extension of the Mobile IPv6 signaling.

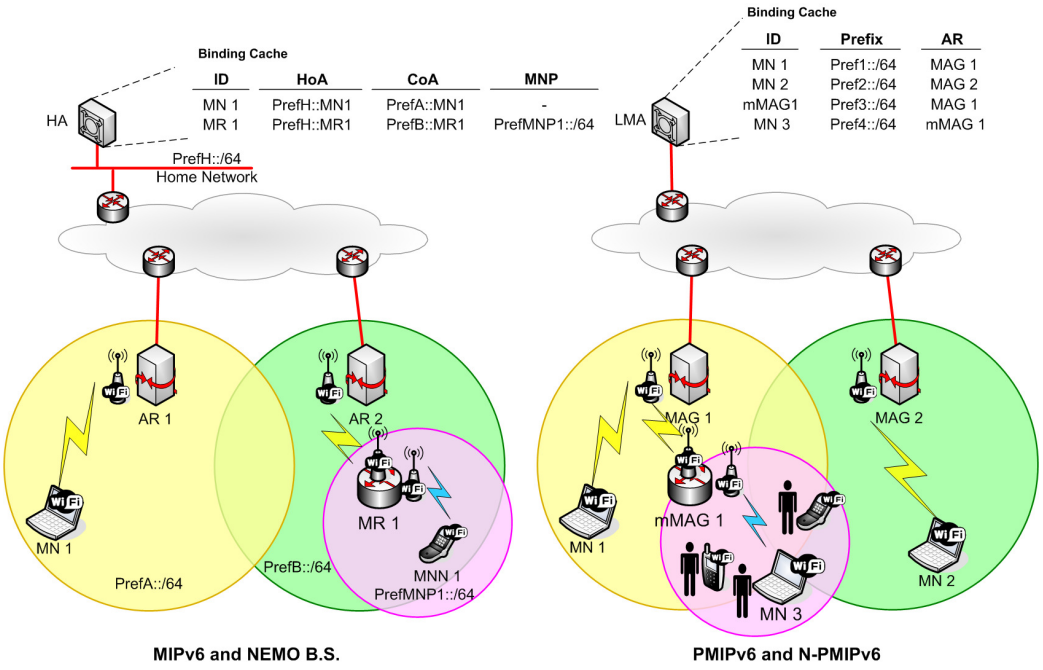

Figure 1: MIPv6, NEMO B.S., PMIPv6 and N-PMIPv6 overview

2.3. NEMO-Enabled Localized Mobility support (N-PMIPv6)

N-PMIPv6 [4] fully integrates mobile networks with Proxy Mobile IPv6. N-PMIPv6 is not a standard, but we include this protocol in our analysis because it is an interesting extension to Proxy Mobile IPv6 specially designed for communications in public transportation systems.The basic idea is to extend a localized mobility domain to include mobile networks as well, so a user terminal is not only able to roam between fixed gateways (i.e., MAGs that do not move, as in conventional PMIPv6), but also between fixed and mobile gateways (called mMAGs, which are also able to roam within the domain), without changing the IPv6 addresses they are using (see Figure 1). A moving gateway (i.e., an mMAG) behaves as a mobile node from the viewpoint of fixed gateways, since moving gateways move between different fixed gateways while keeping the same IP address. Besides, a moving gateway behaves as a regular gateway from the perspective of mobile nodes, and extends the localized domain by providing attached terminals with IPv6 prefixes of the domain, and by forwarding their packets through the localized mobility anchor (i.e., the LMA). An additional bi-directional tunnel between the moving gateway and the localized mobility anchor is used to hide the network

topology, and avoid changing the particular prefix assigned to a terminal while roaming within the same domain. The target scenarios are public transportation systems, in which fixed MAGs are deployed in stations and moving MAGs in vehicles (buses, trains, for example).

3. Combining IP Mobility Protocols

Each of the mobility support protocols described in the previous section are designed to be used independently. However there are circumstances in which two or more of them can be combined. In most cases the combination is the result of individual actions of the different actors –users, operators– involved in the scenario, with each of them deploying a solution to fulfill their own requirements. For example a client-based solution can be set up by a user requiring global mobility, but then the user's MN could visit a network where the operator has deployed a network-based solution to provide mobility support to its visiting nodes. On the other hand, the combination can also be planned to get together different functionalities, for example network mobility and host mobility. The basic combinations do not require modifying the individual protocols. Although they are used together, they are not aware of each other and they do not have explicit mechanisms to cooperate, so there is no increased complexity because there is no new functionality implemented in the involved nodes. We next describe and analyze different combinations of IP mobility protocols, explaining the motivation for each combination, the functionality resulting from that combination, and the additional complexity, if any, that each of the particular combinations brings.

3.1. MIPv6+PMIPv6

A mobile node uses MIPv6 to obtain global mobility support (i.e., it can roam to any visited network while keeping global reachability and session continuity). On the other hand, an operator deploys PMIPv6 to offer local mobility support enabling local roaming (i.e., within the domain) without requiring any support from the user terminals. In this scenario, a MIPv6 node may visit the PMIPv6 domain.

The operation of MIPv6 in the mobile node when visiting a PMIPv6 access network is the same as when visiting any other foreign network: initially, after attaching to the domain, the mobile node gets an IP address (i.e., to be used as its care-of address), and registers it in its global mobility agent (i.e., the home agent), to bind this temporal address to its permanent address (i.e., home address). Since the IP address used in the PMIPv6 domain remains the same while roaming within

this domain, movements are transparent to the mobility management software in the user terminal (i.e., MIPv6). Furthermore, the terminal can also move to an access network outside the domain while keeping ongoing sessions. This is done by the terminal getting another temporal address (i.e., a CoA) from the new access network and using MIPv6 to keep its global mobility agent (i.e., the home agent) updated with its new location.

In this combination of IP mobility protocols, no explicit cooperation among the involved protocols or extra complexity is required, so each of the mobility protocols functions as usual, not even being aware of their simultaneous operation.

3.2. NEMO B.S.+PMIPv6 (+MIPv6)

A mobile router uses NEMO B.S. to obtain global mobility support for itself and the network behind it. A node inside the mobile network can be a normal IP node without mobility support, if it is not going to move away from the mobile network. It can also be a node with MIPv6 to have global mobility support by itself, i.e., to be able to leave the mobile network and roam to other networks. In addition, an operator deploys PMIPv6 to offer local mobility support enabling local roaming (i.e., within the domain) without requiring any support from visiting nodes (hosts or routers).

A particularly relevant example of this scenario is the provision of Internet connectivity in public transportation systems (e.g., buses) where users benefit from seamless access using mobility unaware devices, while the network mobility support (i.e., the mobile router) takes care of managing the mobility on behalf of the terminals. Some of the access networks the mobile network may visit could also provide PMIPv6 support. In this situation, where NEMO B.S. and PMIPv6 protocols are combined, when the mobile router enters the localized domain, it gets a temporal address (to be used as its care-of address) from the domain and registers this address in its global mobility agent (i.e., the home agent), binding the mobile network prefixes managed by the mobile router (that is, the IPv6 prefixes used inside the mobile network) to its current location (i.e., its care-of address). Since this new acquired IPv6 address is provided by the PMIPv6 domain, it does not change while the mobile network roams within the localized domain, and therefore its movements are transparent to the mobility software in the mobile router (i.e., NEMO B.S.). Moreover, the mobile network is able to roam not only within the localized domain but also outside the domain, thanks to the NEMO B.S. operation that provides global mobility support. A user terminal in the mobile network (i.e., attached to the mobile router) will not be able to leave

the mobile network without breaking its ongoing sessions unless this terminal has MIPv6 support.

As in the previous case, in this combination of IP mobility protocols, no explicit cooperation among the involved protocols or extra complexity is required, so each of the mobility protocols functions as usual, not even being aware of their simultaneous operation.

3.3. MIPv6+N-PMIPv6

This combination is very similar to the first one (MIPv6+PMIPv6). A mobile node uses MIPv6 to obtain global mobility support (i.e., it can roam to any visited network while keeping global reachability and session continuity). In addition an operator deploys N-PMIPv6 to offer local mobility support enabling local roaming (i.e., within the domain) without requiring any support from user terminals. With N-PMIPv6 this local mobility domain is composed of fix and moving access gateways. In this scenario, a MIPv6 terminal may visit the N-PMIPv6 domain.

In this scenario a user terminal can both move within a localized domain – without changing its IP address (which is used as care-of address) – and can also leave the domain without breaking any ongoing communications, by acquiring a new temporal address (i.e., a care-of address) from the new access network and using MIPv6 to register this temporal address in its global mobility agent. The difference with the first combination is that here the localized domain integrates both fixed gateways (MAGs) and moving gateways (mMAGs), so that a user terminal is able to roam between fixed and mobile access infrastructure within the domain without involving/requiring any IP mobility support in the terminal (thanks to the use of N-PMIPv6 protocol). Whenever the terminal changes its location within the domain, the new access gateway (i.e., fixed or mobile) will update the terminal's location in the mobility anchor (i.e., LMA).

An example of this scenario could also be a public transportation system, where mobility unaware devices would not only get Internet access while moving (e.g., in a bus or train) or while waiting at the station platforms, but also while roaming between fixed and mobile access infrastructure (e.g., getting on or getting off a bus). Additionally, the use of MIPv6 would also enable a mobile node to roam outside the localized domain, for example, when leaving the public transportation environment.

As in the previous case, in this combination of IP mobility protocols, no explicit cooperation among the involved protocols or extra complexity is required, so each of the mobility protocols functions as usual, not even being aware of their simultaneous operation.

3.4. NEMO B.S.+N-PMIPv6 (+MIPv6)

In this combination, as in the previous one, an operator deploys N-PMIPv6 to offer local mobility support enabling local roaming (i.e., within the domain) without requiring any support from the user terminals. But, in addition, the operator also deploys NEMO B.S. mobile router capabilities in the moving gateways, which enable the corresponding mobile networks to be able to move outside the localized domain while keeping ongoing sessions. This could be a common configuration if the mobile network needs to move out of a domain (e.g., a bus leaves the N-PMIPv6 localized domain deployed in a city and connects to another network operator). In this combination, thanks to the use of the N-PMIPv6 protocol, the localized domain integrates both fixed gateways (MAGs) and moving gateways (mMAGs), so that a user terminal is able to roam between fixed and mobile access infrastructure within the domain without changing its IP address. The terminal can also be connected to a mMAG that moves outside the localized domain and, thanks to the use of NEMO B.S. functionality, this movement will be transparent to terminals in the mobile network, i.e., they will not need to change their IP addresses. The terminal can also use MIPv6 to obtain global mobility, i.e., to be able to roam outside the access infrastructure provided by the operator through N-PMIPv6 and the mobile networks created by using NEMO B.S.

The most efficient way of deploying this scenario is by co-locating the global mobility agent of the mobile router functionality (i.e., the home agent) and the local mobility anchor of the moving gateway (i.e., the LMA) in the same node, so they share the range of addresses to be used (i.e., the mobile network prefixes of the NEMO are part of the IPv6 address space of the localized domain and, therefore, they are topologically anchored at the LMA). With this configuration, the localized domain becomes also the home network of the global mobility support (i.e., home domain). Therefore, when the mobile network is at the home domain, packets addressed to a user terminal attached to this mobile network are forwarded as in the N-PMIPv6 simple case (i.e., through the localized mobility anchor – LMA). This means that when the mobile network is away from its home domain, a bi-directional tunnel is created between the mobile router – after obtaining a new care of address from the visited network – and its home agent, used to forward all the traffic from or to terminals connected to the mobile network. Note that in case the mobile network moves out of its home domain, the mobile router (also with moving gateway functionality) cannot act anymore as a moving gateway (either because the visited domain is not an N-PMIPv6 localized domain or because the moving gateway lacks the appropriate security associations with the localized mobility anchor of the visited domain). When the mobile network

is not at its home domain, a user terminal moving away from the mobile network would need to change its IP address, thus breaking ongoing sessions unless the mobile node has its own MIPv6 support.

In this combination a node in the network has to combine LMA (N-PMIPv6) and HA (NEMO B.S.) functionality. Additionally, the moving access gateways have to combine mMAG (N-PMIPv6) and MR (NEMO B.S.) functionality. [6] documents the issues that might arise from the interactions between PMIPv6 and MIPv6 when the LMA and the HA are co-located, being some of their recommendations applicable to the NEMO B.S. + N-PMIPv6 combination addressed in this section. The implementations of N-PMIPv6 and NEMO B.S. can work independently (actually, [6] recommends to avoid the LMA and HA entities sharing their binding cache). Nevertheless we have to guarantee the compatibility of the addressing assigned by both protocols to the same nodes. This can be done by using static pre-assignments of IP prefixes to be used by each mMAG/MR. In the mMAG/MR node, mobile router functionality can be triggered for example by changes in the used IP prefix in the outgoing interface, and moving MAG functionality can be triggered by detecting the advertisement of an IP prefix in the outgoing interface that belongs to the mMAG's home network. Other means of triggering the protocols are also possible, such as using hints from the authentication mechanism in the access network. When the mMAG/MR enters a visited network away from its home domain, it has to register the IP prefixes used inside the mobile network in the HA. When the mMAG/MR enters the home domain it has to register itself in the LMA and also it has to register the identities of the nodes attached to the mMAG. In the LMA/HA, each implementation processes its own signaling and behaves accordingly, without affecting the other one. We could make some optimizations by enabling cooperation between both protocols. For example, in the LMA/HA both implementations could share a database with information about prefixes and the identities of nodes using them. The database would be updated dynamically by both implementations. Therefore, for making this combination work, we need to combine the implementation of different protocols in the same nodes (LMA/HA and mMAG/MR) and the corresponding configuration. This means some added complexity in both the LMA/HA and the mMAG/MR. But the added complexity is not much compared with the independent implementation of the mMAG and the MR functionalities, and the operator gains the ability to offer transparent connectivity service to nodes roaming in its domain or connected to its mobile networks even when they move to other domains, and that without depending on functionality or configuration in the mobile nodes themselves. The possible use of MIPv6 in a mobile node to achieve global

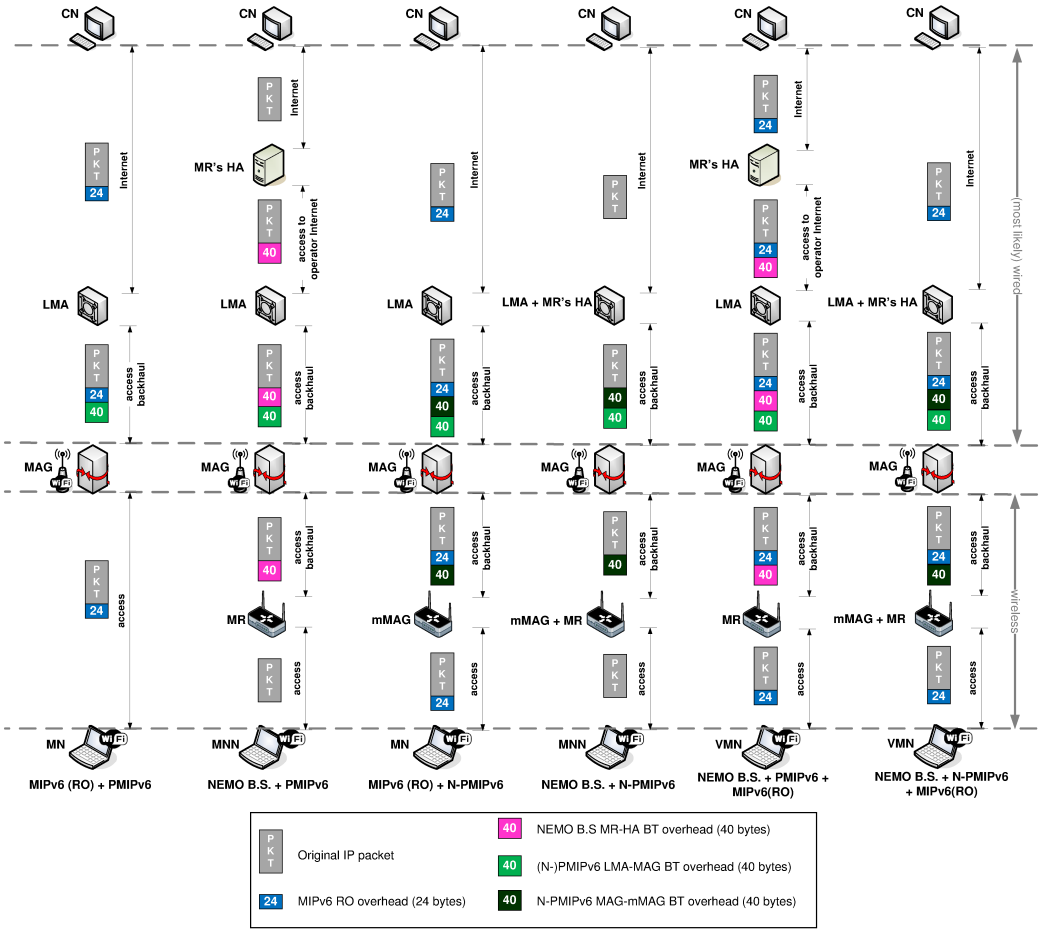

Figure 2: Overhead of the different combinations

mobility by itself is independent of the N-PMIPv6+MIPv6 solution, both work unaware of each other, so there is not added complexity in this case.

4. Performance Analysis

We next present the results of an analytic performance evaluation of combining different IP mobility solutions, in terms of protocol overhead and handover latency.

4.1. Overhead

In this section we analyze the overhead – in terms of control headers (tunnels) – of each combination. In this analysis we focus not only on the amount of added packet overhead, but also on which segments of the network suffer from this extra overhead, as the impact is more important if the additional control information appears on segments of the network where the transmission media is wireless. In order to simplify this exercise, we limit the analysis of the MIPv6 overhead to the case where Route Optimization (RO) is enabled. The difference between using RO mode or bi-directional tunneling (BT) mode is basically the following: in BT mode, the overhead is higher (40 bytes, instead of 24), but it only appears between the terminal and its home agent, while in route optimized mode the overhead is present along the complete path (i.e., between the terminal and the correspondent node).

4.1.1. MIPv6 + PMIPv6

In this case, 24 bytes of additional overhead are added in the whole path between the mobile node and the correspondent node, due to the use of Mobile IPv6 (in RO mode), plus an IPv6 tunnel (40 bytes) between the localized mobility anchor and the gateway (MAG) where the mobile node is attached to (due to the use of Proxy Mobile IPv6). It is important to note that, out of the overall overhead, only the 24 bytes added by Mobile IPv6 are present in the wireless access.

4.1.2. NEMO B.S. + PMIPv6 (+ MIPv6)

Two different tunnels are involved to enable the communications of the mobile network: one between the mobile router and its home agent (due to the use of NEMO B.S.), and another between the localized mobility anchor and the gateway serving the mobile router (due to the use of Proxy Mobile IPv6). Thus, there are up to 80 additional bytes of overhead in some wired segments of the path (when both tunnels are present), and up to 40 bytes in the wireless access (due to the use of NEMO B.S.), though not in the last wireless hop between the user terminal (i.e., the MNN) and its access router (i.e., the MR). Note that this last wireless hop is where the effect on battery consumption and bandwidth waste is likely to be more significant. A third overhead component (24 bytes in route optimized mode) is required if the terminal attached to the mobile network is itself a MIPv6 mobile node outside its home network.

4.1.3. MIPv6 + N-PMIPv6

In this case, where a mobile node is attached to a moving gateway, three overhead components are required: one (24 bytes) between the mobile and the correspondent node (due to the use of Mobile IPv6 in route optimized mode), an IPv6 tunnel (40 bytes) between the localized mobility anchor and the fixed gateway where the moving gateway is attached to, and a second tunnel between the localized mobility anchor and the moving gateway.

4.1.4. NEMO B.S. + N-PMIPv6 (+ MIPv6)

When a mobile network is attached to a moving MAG (which is at its home N-PMIPv6 domain), and assuming a deployment scenario in which the localized mobility anchor and the home agent of the mobile router are co-located (see Section 2.2), two IPv6 tunnels are required: one between the localized mobility anchor and the fixed gateway serving the moving gateway, and a second between the localized mobility anchor and the moving gateway. If the user terminal that is getting access through the mobile network is a mobile node running Mobile IPv6 (which is outside its home network), an additional overhead component (24 bytes) is required (due to the use of Mobile IPv6 in route optimized mode).

Figure 2 shows the overhead of all the analyzed combinations over the different network segments. Depending on the combination under consideration, we can have up to three extra headers in the wired access network backhaul, up to two in the wireless access backhaul (i.e., between the fixed access network and the moving MAG/mobile router), and up to one in the last wireless hop to the terminal.

In order to gain understanding on the effect of the mobility overhead with user traffic, we have taken data from a real access network deployment offering Internet access during a conference (ACM CoNEXT 2008 [7]). The average packet size for UDP or TCP traffic is 710 bytes (including all headers). For the case of three additional overhead components (104 bytes), the extra headers account for a waste of 14.6% of the bandwidth. If two extra headers are required, the waste is between 9% and 11.26% (for the cases of 64 and 80 bytes of overhead, respectively). Finally, if only one overhead component is needed, the bandwidth waste is between 3.38% and 5.6% (for the cases of 24 and 40 bytes). Nevertheless, note that these figures just represent a mixture of user traffic – composed mostly of HTTP data – in a conference. In mobile scenarios the overhead penalty will tend to be worse, for example with the expected increase of VoIP traffic. It is worth highlighting that the extra headers – in addition to the bandwidth waste – also involve the extra energy consumption required to transmit them, which is signifi-

cant in wireless environments. Moreover, it is commonly argued that the problem is not so important in the access network backhaul, because it is usually wired and bandwidth is not severely limited, but wireless multi-hop access networks are becoming increasingly popular, which weakens this reasoning.

4.2. Handover latency of IP mobility protocols

The handover delay of mobility protocols can be expressed as a combination of independent factors such as the layer-2 handover time, the movement detection time, the IP configuration time, and the specific mobility signaling delay. We briefly describe first each of these independent factors.

1. Layer-2 handover time (T_{L2}^{ho}). It is defined as the time required by layer-2 technology to perform a handover (i.e., disconnecting from its current point of attachment and connecting to a new one). In the case of an IEEE 802.11-based layer-2 technology, this time usually involves the channel scanning for candidate Access Points (APs), plus the time required for re-association. As presented in [8], this delay can be modeled with a Beta probability distribution function. In [9] it is shown how its mean value can be reduced up to 50 ms by appropriately selecting the number of channels being scanned. For the handover delay analysis we conduct later, we take this last number as the mean value of the Beta distribution, while the standard deviation of the resulting distribution is of 3.66 ms.

2. Movement detection time (T_{MD}). This delay corresponds to the time required by the terminal to detect that it has moved to a different layer-3 point of attachment. This detection can be performed by functionality at the IP layer or assisted by layer-2 mechanisms. If we focus on IPv6 mechanisms, movement detection can be done in different ways. The most simple (and the most widely supported) consists in using Routing Advertisement (RA) messages. An access router periodically multicasts unsolicited RA messages. Typically, the time interval between these advertisements follows a uniform distribution: whenever a router advertisement is sent from an access router, a timer is set to a uniformly distributed random value [10] between the configured *MinRtrAdvInterval* (R_m) and *MaxRtrAdvInterval* (R_M). Although IPv6-based movement detection mechanisms are well known and supported, new optimized mechanisms to detect the connection of a terminal to a new point of attachment have been and are still

being developed. That is the case of a mechanism known as link layer triggers, of which IEEE 802.21 Media Independent Event Service is a good example. This technology enables lower layers to notify the occurrence of a certain event (such as attachment or disconnection) to higher layers, e.g., the mobility management protocol. For the purposes of this work we consider the use of layer-2 assisted movement detection since it is the mechanism introducing the smallest delay, almost negligible.

3. IP configuration time (T_{IP}). This is the time required by the IP stack to configure a new IP address and update the forwarding table. This time depends on the hardware and operating system since this operation is generally performed by the kernel. It should be noted that this delay is not always present in a handover event, as the network-based mobility protocols ensure that the IP address, as well as the default router, of the moving terminal remain the same while roaming within the domain. On the other hand, if client-based mobility is used, the MN needs to configure a new IP address if the former one is no longer valid, and signal it to the anchoring point, e.g., in MIPv6 the MN needs to configure a new care-of address and send a Binding Update message to its HA. In the rest of the article we assume that the IP configuration time is very short (negligible), since the operations of configuring an address and updating the routing in the IP stack should not require a long time in modern computers or hand-helds. It is also worth noticing that in this work we assume the use of IPv6 stateless auto-configuration mechanisms (SLAAC). The use of a different IP address configuration mechanism (e.g., DHCP) is likely to incur in higher delays.

4. Signaling delay ($T_{BU/BA}$ or $T_{PBU/PBA}$). This is the time required to update the local/global mobility agent (i.e., LMA/HA) with the new location. It highly depends on the distance between the entities participating in the mobility management: the terminal/gateway/mobile router on the one side and the localized mobility anchor/home agent on the other side. In order to model this behavior we use measurements taken from the PingER (Ping end-to-end reporting) project[3]. We take the average value of the reported values for 3 types of scenarios characterized by the distance between the communication peers. In particular, we take a "local" delay characterized

[3] http://www-iepm.slac.stanford.edu/pinger/

by an average value of 5.37 ms, a "regional" delay of 18.32 ms and a "continental" delay of 138.79 ms. To characterize the delay in an Internet path, these values are used as the mean of a Weibull distribution with variance provided by Hurst parameters of 0.8, 0.65 and 0.5 for local, regional and continental delays respectively [11] [12].

After explaining the different independent factors that are common to the handover time for all mobility solutions, we focus in the following on understanding the handover delay for the different IP mobility management protocols (Mobile IPv6, NEMO B.S., Proxy Mobile IPv6 and N-PMIPv6).

1. *MIPv6/NEMO B.S.* The delay incurred by a Mobile IPv6 terminal performing a handover (in bi-directional tunnel mode) or by a mobile router, can be expressed as:

$$T(MIPv6\ BT/NEMO) = T_{L2}^{ho} + T_{MD} + T_{IP} + RTT(MN/MR, HA), \quad (1)$$

   where RTT(MN/MR,HA) represents the round trip time between the mobile node (or router), and the corresponding home agent.
   In case route optimization mode is used in Mobile IPv6, we assume the mobile node is performing optimistic registration [13], which means that the route optimization related signaling is performed in parallel with the registration signaling with the home agent. For this case, an additional $RTT(CN, HA)$ component should be added to the delay shown in Eq. (1).

2. *PMIPv6*. If Proxy Mobile IPv6 is used to manage the mobility of the user terminal, the handover delay can be expressed as:

$$T(PMIPv6) = T_{L2}^{ho} + T_{MD} + RTT(MAG, LMA) \quad (2)$$

   In this case, in addition to the time required for the layer-2 handover, movement detection and authentication, we also add the signaling delay between the gateway and the localized mobility anchor. In case the MN is entering the localized domain for the first time, an additional T_{IP} component should be added to the delay shown in Eq. (2).

Table 1: Delay of the combination of different mobility solutions

Entity Moving		MIPv6+ PMIPv6	NEMO B.S.+ PMIPv6	MIPv6+ N-PMIPv6	NEMO B.S.+ N-PMIPv6	NEMO B.S.+ PMIPv6+MIPv6	NEMO B.S.+ N-PMIPv6+MIPv6
MN	within LMD	T(PMIPv6)	N/A	T(PMIPv6)	T(PMIPv6)	*MR* $\rightarrow$ *MAG*: T(MIPv6+PMIPv6) *MAG* $\rightarrow$ *MR*: T(MIPv6)	T(PMIPv6)
	to LMD	T(MIPv6+PMIPv6)	N/A	T(MIPv6+PMIPv6)	N/A	T(MIPv6)	T(MIPv6+PMIPv6)
MR	within LMD	N/A	T(PMIPv6)	T(PMIPv6)	T(PMIPv6)	T(PMIPv6)	T(PMIPv6)
mMAG	to LMD	N/A	T(NEMO+PMIPv6)	NA	T(PMIPv6)	T(NEMO+PMIPv6)	T(PMIPv6)

3. *N-PMIPv6*. The difference in terms of handover delay between the case of using N-PMIPv6 and the regular Proxy Mobile IPv6 case is just the additional hop between the moving MAG and the fixed MAG that has to be traversed. We consider this difference negligible in the next calculations.

4.3. Handover Latency for the Combinations of IP Mobility Protocols

We analyze next the resulting handover delay for the different combinations of mobility protocols when either the mobile node or the mobile router/moving gateway moves. This is done for the cases when the handover is performed within the Local Mobility Domain (LMD), or entering the localized domain. In all the following situations we assume that the visited domain supports Proxy Mobile IPv6 or its extension N-PMIPv6.

Table 1 summarizes the resulting handover delays for the different mobility combinations.

4.3.1. MIPv6+PMIPv6

This combination corresponds to a user terminal with Mobile IPv6 support that may hand off to a localized domain. In this case there are two possible mobility scenarios: *i)* The terminal moves within a localized domain, hence the mobility of the terminal is handled within the domain using Proxy Mobile IPv6 (hence the handover delay is equal to the PMIPv6 case), or *ii)* the terminal enters a localized domain, then the terminal handles itself using Mobile IPv6 – the global mobility, performing a handover to the localized domain. In this case the use of both mobility solutions is done in a sequential way: the terminal first configures an address that belongs to the prefix obtained through the operation of Proxy Mobile IPv6, which is then used as its care-of address by the Mobile IPv6 protocol running on the terminal. The handover delay of this approach, when bidirectional tunnel mode is used, can be expressed as:

$$
\begin{aligned}
T(MIPv6+PMIPv6) &= T_{L2}^{ho} + T_{MD} + T_{IP} + RTT(MAG, LMA) \\
&\quad + RTT(MN, HA) \\
&= T_{L2}^{ho} + T_{MD} + T_{IP} + RTT(MAG, LMA) \\
&\quad + RTT(MN, LMA) + RTT(LMA, HA) \\
&\simeq T_{L2}^{ho} + 2 * RTT(MAG, LMA) \\
&\quad + RTT(LMA, HA).
\end{aligned}
\tag{3}
$$

We separate the RTT between the MN and the HA in two parts: the RTT between the MN and the LMA, and the RTT between the LMA and the HA. This is because when the MN is in a PMIPv6 domain, its traffic, including the MIPv6 signaling, has to go through the LMA. Separating the RTT in two parts allows us to consider the influence of the distance between the LMA and the HA. Additionally, the RTT between the MN and the LMA is equal to the RTT between the MAG and the LMA plus the delay in the hop MAG-MN that we consider negligible.

In the case of optimistic route optimization mode, an additional $RTT(CN, HA)$ component should be added to the delay shown in Eq. (3).

4.3.2. NEMO B.S.+PMIPv6

In this combination, instead of a user terminal running Mobile IPv6, the entity moving is a mobile router running the NEMO B.S. protocol. As in the previous section, two situations may arise: *i)* the mobile router moves within a localized mobility domain, hence the mobility of the router within the domain is handled using Proxy Mobile IPv6, or *ii)* the mobile router enters a localized domain, so the router handles the macro-mobility (using the NEMO B.S. protocol). This case is similar to the MIPv6+PMIPv6 without route optimization, being Eq. (3) also applicable in this scenario (considering a mobile router instead of a mobile node).

$$
\begin{aligned}
T(NEMO+PMIPv6) &= T_{L2}^{ho} + T_{MD} + T_{IP} + RTT(MAG, LMA) \\
&\quad + RTT(MR, LMA) + RTT(LMA, HA) \\
&\simeq T_{L2}^{ho} + 2 * RTT(MAG, LMA) \\
&\quad + RTT(LMA, HA).
\end{aligned}
\tag{4}
$$

4.3.3. MIPv6+N-PMIPv6

This scenario considers a user terminal with Mobile IPv6 support that may move to a localized mobility domain where gateways are able to move (mMAGs), hence there are two mobile entities: the mobile node and the moving gateway. In case the terminal attaches to a moving gateway, as shown in the delay explanation for the N-PMIPv6 solution (see Section 4.2), the difference in delay between the Proxy Mobile IPv6 case and N-PMIPv6 case is the one due to an additional hop in the local network. Taking this fact into account, the combined solution of MIPv6+N-PMIPv6 presents a slightly increase in the delay: $RTT(mMAG, MAG)$ – which we consider negligible – compared to MIPv6 + PMIPv6.

If the entity moving is a moving gateway, it can move either within the localized mobility domain or from outside to the domain.

N-PMIPv6 protocol allows the moving gateway to roam within the domain (using the Proxy Mobile IPv6 protocol). Hence, in the case the moving gateway roams within the domain, the handover delay is equal to the one obtained in Proxy Mobile IPv6. The case where the moving gateway moves to a different localized mobility domain is not considered here, since there is no mobile router functionality in the moving gateway and, therefore, mobility cannot be granted.

4.3.4. NEMO B.S.+N-PMIPv6

This combination considers user terminals without Mobile IPv6 support and moving MAGs that also incorporate NEMO B.S. functionality, so they can hand off outside the localized mobility domain. As in the previous section, in this scenario two entities are able to move, the user terminal and the moving gateway. In case the terminal moves, it can attach to another access router belonging to the same domain and the mobility is handled by the network-based signaling (N-PMIPv6).

The case where the terminal hands off from outside to the domain is not applicable since the terminal does not have mobility support in this case and, hence, it cannot hand off from outside the localized mobility domain. If the entity moving is the moving gateway, it can move *i)* within the domain, using the N-PMIPv6 protocol that allows the moving gateway to roam within the domain, or *ii)* to the domain from the outside (in this case the moving gateway is outside the domain and it performs a handover to it). As the moving gateway is part of the domain, the prefix used by its mobile router functionality belongs to the localized domain. Therefore it only requires performing a Proxy Mobile IPv6 registration in order to hand off to the domain.

4.3.5. NEMO B.S.+PMIPv6+MIPv6

This scenario encompasses a user terminal supporting Mobile IPv6, and a localized mobility domain where there is a mobile router attached. The terminal can move within the domain or can move to it from the outside. In case the terminal is attached to the domain, it can be anchored to a gateway or to the mobile router. For the scenario where the terminal is attached to the mobile router and hands off to a gateway, the terminal uses both Mobile IPv6 and Proxy Mobile IPv6 to reconnect to the domain, i.e., it is equivalent to the MIPv6+PMIPv6 scenario. On the other hand, if the terminal is attached to a gateway and hands off to the mobile router, it must use Mobile IPv6 to regain connectivity, since the address provided by the mobile router does not belong to the domain, but to the home network of the router.

The case where the terminal is moving to the domain from the outside is equivalent to the MIPv6+PMIPv6 case. In this mobility combination scenario the mobile router can also move, being this case equivalent to the NEMO B.S.+PMIPv6 scenario.

4.3.6. NEMO B.S.+N-PMIPv6+MIPv6

This scenario considers a terminal with Mobile IPv6 support and a localized mobility domain where there are moving gateways (mMAGs) with mobile router (NEMO B.S.) functionality. This scenario can be analyzed as a combination of the previous scenarios. The case where the terminal moves to the domain, is equivalent to the combination MIPv6+N-PMIPv6, while the case where the mobile router moves is equivalent to the NEMO B.S.+N-PMIPv6 combination.

4.4. Delay performance analysis

Figure 3 allows us to gain some insight about the impact of the combination of different mobility solutions on the delay experienced during handovers. It represents, for different mobility solutions and different topologies, the percentage of handovers with a delay below a particular threshold, namely 150 ms. This value is a reasonable disruption time for most applications, assuming the use of some buffering function to minimize packet loss during the interruption caused by the handover in the communication [9]. The first two bars on the left of the figure are the percentage of handovers whose delay is below 150 ms for NEMO B.S. or MIPv6 in bi-directional tunneling mode (BT), depending on the distance between the Home Agent and the Mobile Node/Router. Using global mobility the MN/MR can move everywhere. Additionally, the visiting network may provide local mobility support. This allows better efficiency in handovers inside the local

domain (see the bars referring to PMIPv6 in Figure 3) but at the cost of a longer handover delay to move into the local domain, as shown in the bars referring to the MIPv6 BT/NEMO B.S.+PMIPv6 combination. This cost increases with the distance between the Mobile Anchor Gateway and the Localized Mobility Anchor. For example, a MIPv6 terminal that is not using Route Optimization and it is at regional distance from its HA has a probability close to 95% of having a delay below 150 ms when executing a handover to an access network without local mobility support; but when the access network has local mobility support the probability is reduced to the 55-90% range depending on the distance between the MAG and the LMA in the local domain and the distance between the LMA and the HA. It is also worth noticing that the effect of the delay MAG-LMA is greater than the effect of the delay LMA-HA. This is because the path MAG-LMA is traversed both for the PMIPv6 and for the MIPv6 signaling while the path LMA-HA is only traversed by the MIPv6 signaling – see Eq. (3). Figure 3 also shows the performance of the handovers of MIPv6 terminals with Route Optimization, and of the handovers of MIPv6 terminals with Route Optimization moving to a PMIPv6 domain.

To fully understand the importance of these results we need to consider two aspects. First, when evaluating handover performance we need to focus on the longest handover that the terminal can suffer in any type of handover, because that will determine the performance and the needed mechanisms (e.g., buffering) to avoid interruptions in the terminals communications. The second aspect is the frequency of the handovers: if some type of handover was very unusual, we could accept having a worse performance in that type of handover. The frequency of the type of handovers depends on the scenario, but we argue that the trend in mobile communications networks is towards very dynamic mobility scenarios in which nodes will change of access network very frequently according to the access network availability and the terminal requirements, so probably every type of handover will become usual.

5. Experimental analysis

We next complement the findings of our analytic study, by performing an experimental evaluation using Linux-based implementations of IP mobility protocols and conducting several experiments under different scenarios.

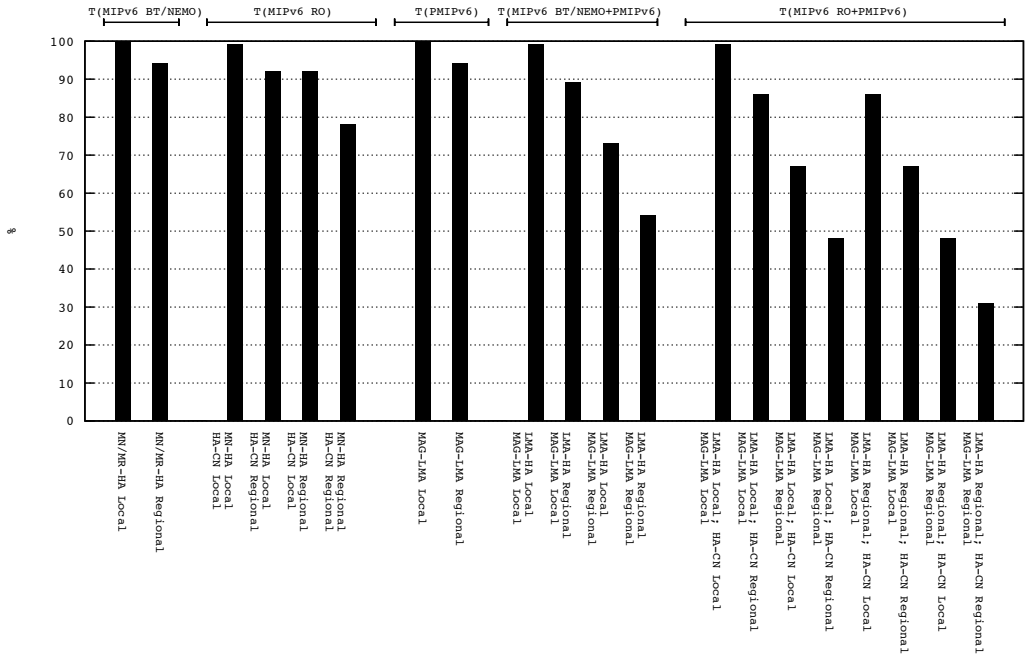

Figure 3: Percentage of handovers below 150 ms

5.1. Testbed description

In order to evaluate experimentally the behavior of some of the combinations of mobility protocols considered, we have designed and deployed a testbed covering the scenarios shown in Figure 4. The role of the HA, MR/MN, LMA, MAG and AR are played by Linux boxes. The MR/MN, MAG and AR are equipped with Atheros wireless cards using the `ath5k` driver[4]. We have used an in-house software implementation of the MIPv6 BT/NEMO B.S. protocol developed under the framework of the POSEIDON project[5] and the OAI PMIPv6 implementation.[6]

The main purpose of these tests is to confirm the findings of our analytic analysis in terms of handover delay. In order to do so, we have measured the handover time, splitting it into several steps: *i)* layer-2 handover delay (T_{L2}^{ho}), *ii)* movement detection delay (T_{MD}), *iii)* IP configuration time (T_{IP}) and *iv)* mobility signaling

[4] `http://linuxwireless.org/en/users/Drivers/ath5k`

[5] `http://enjambre.it.uc3m.es/~poseidon/`

[6] OpenAir Interface PMIPv6: `http://www.openairinterface.org/components/page1103.en.htm`

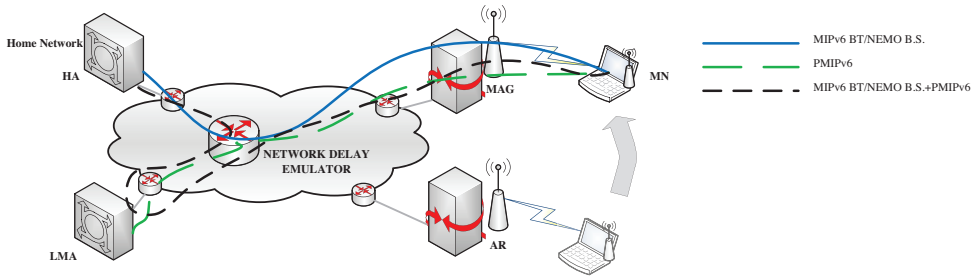

Figure 4: Testbed description

delay ($T_{BU/BA}$ and $T_{PBU/PBA}$) of the mobility protocol.

The measurements of the different handover steps have been taken by monitoring each of the network interfaces of all the mobility entities involved and time-stamping the transmission and reception of each control message. To emulate different distances between the entities participating in the mobility management (e.g., between LMA and MAG, and between HA and MR), an additional router, capable of adding a variable delay using `netem`[7] (Network delay emulator in Figure 4), has been introduced. In order to extract statistically reliable figures, each test has been repeated between 20 and 40 times.

The time for performing a layer-2 handover (T_{L2}^{ho}) has been measured as the time elapsed between the moment that the wireless client starts trying to associate with a new access point and the moment this association is effectively completed. The movement detection delay (T_{MD}) includes the time required for the layer-2 event notifying of the new link layer connection. The reception of this event at the IP layer triggers the MN/MR to send a Router Solicitation message. As explained in the Section 4.2, the IP configuration time (T_{IP}) is only present on the client-based (e.g., MIPv6 or NEMO B.S.) handovers. In network-based solutions (e.g., PMIPv6), this time is only present when the node attaches for the first time to the localized domain, since roaming within the domain does not require a change in the IP address assigned to the terminal nor its forwarding state. Last but not least, the mobility signaling delay ($T_{BU/BA}$ and $T_{PBU/PBA}$) is inherent to each of the mobility solutions. For MIPv6 BT/NEMO B.S., this time is measured as the time elapsed between the transmission of the BU message and the reception of the BA at the mobile node. For PMIPv6 this time does not only include the

[7]Network Emulator: `http://www.linuxfoundation.org/collaborate/workgroups/networking/netem`

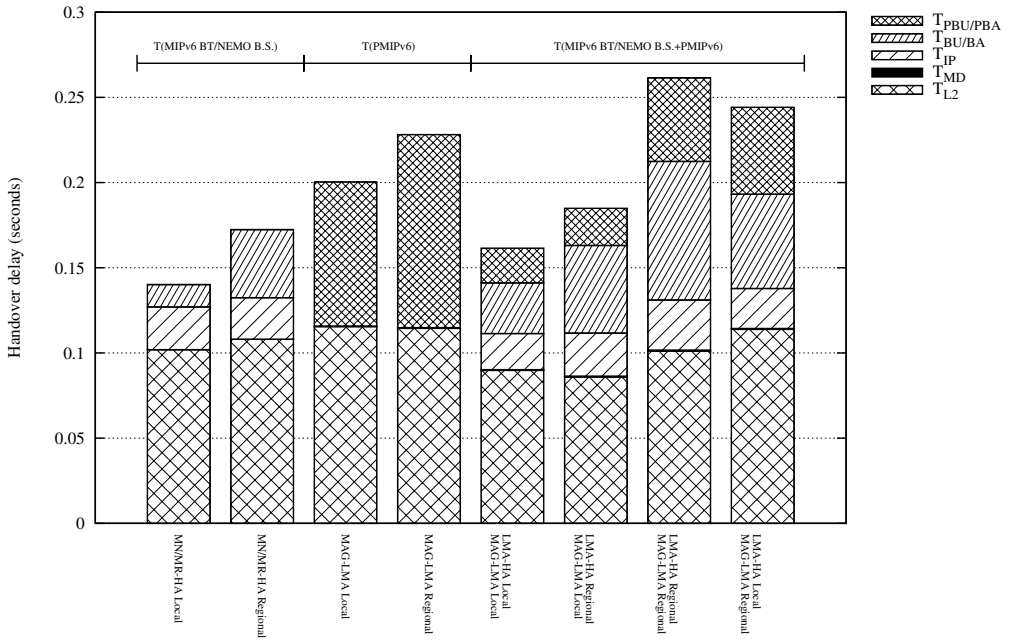

Figure 5: Experimental delay analysis

PBU-PBA exchange between the MAG and the LMA, but also the time required by the MAG to send a Router Advertisement conveying the IP prefix assigned to the mobile node/router in the domain. Finally, the combination of MIPv6 and PMIPv6 incurs in the longest total signaling time, as it comprises the time since the transmission of the PBU by the MAG to the reception of the BA by the mobile node.

5.2. Experimental Results and Evaluation

Figure 5 shows the different components of the total handover delay obtained from our experimental tests. These tests confirm our initial findings about the performance of the combinations analyzed, and bring some additional interesting conclusions as well.

It is remarkable that the main contribution to the overall handover time is the layer-2 handover delay. Our measured average layer-2 handover delay is about 100 ms, although we should highlight that this layer-2 handover has been performed without any further optimization, and therefore a lower delay could be

obtained by, for example, appropriately configuring the channels to scan and fine tuning the layer-2 mechanisms involved in the association process.

The movement detection time is confirmed to be negligible, with an average measured time of less than a millisecond.

On the other hand, the experimentation has shown that the process of configuration and management of the IP addresses and routes is very time consuming. While these procedures are not present for the case of PMIPv6, the IP configuration time becomes relevant in the case of MIPv6 BT/NEMO B.S. and the combination MIPv6 BT/NEMO B.S+PMIPv6, as the mobile node/router cannot send the Binding Update message until the egress interface is configured properly with the IP addresses assigned by the access router. In fact, our experiments show that the configuration of an IP address in the terminal requires an average of 20 ms[8], which seems to be an implementation-specific aspect that can be improved.

The signaling time of the mobility protocols depends strongly on the distance between the entities involved in the mobility management (i.e., MAG and LMA, or HA and MR and their combinations). When the distance is local, practically all the handovers require less than 150 ms to be completed. When the delay is regional, the percentage of handovers requiring less than 150 ms decreases. The fact of combining protocols also has an impact, being the distance between the MAG and LMA the most significant factor. A longer delay between these two elements impacts significantly in the overall performance of the combination, as reflected by the comparison of the *LMA-HA Local MAG-LMA Local* with the *LMA-HA Local MAG-LMA Regional* cases in Figure 5.

If we look at the mobility signaling delay ($T_{PBU/PBA}$ and $T_{BU/BA}$), it is worthwhile mentioning that our experimental results show that, besides the delay due to the RTT between the involved mobility entities, there is also a non negligible delay caused by the processing time of the software implementation of the mobility protocols. Note that the implementations used in our experiments are research software, with code not optimized for performance.

This results in a measured signaling time noticeably higher in our experiments than the one considered in our theoretical analysis. One particularly unfortunate case is the PMIPv6 signaling delay resulting from the used implementation, as it requires the LMA to remove the tunnel that was being used before doing any further processing of a received PBU. This is a bug of the implementation that

[8]Note that Duplicate Address Detection (DAD) was disabled in our tests. Had it been enabled, it would have added an average of 1 second to the IP configuration time.

introduces a considerable additional delay[9]. If we compare the handover delays of PMIPv6 and MIPv6 BT/NEMO B.S.+PMIPv6 in Figure 5, we can see that the PMIPv6 signaling delay $T_{PBU/PBA}$ of the latter is considerably lower because, in that case, the MN/MR just arrives to the localized domain, so no previous tunnel has to be removed. Based on this, although we have included the results of the PMIPv6 case for completeness, we do not consider them representative of the real performance of the protocol, for the reasons explained above.

If we compare the results of the handover delay of our theoretical analysis performed in Section 4.4 with the experimental results here, we can see that measured values from the conducted tests are higher than the ones resulting from the analytic evaluation. This is due mainly to higher layer-2 handover delays and some processing times not considered in the theoretical analysis (most are implementation-specific, therefore not invalidating our analysis). Nevertheless, our general finding that combining IP mobility protocols tends to increase the overall handover delay is confirmed as can be seen by comparing the MIPv6 BT/NEMO B.S. results with the MIPv6 BT/NEMO B.S.+PMIPv6 results in Figure 5.

6. Facing the shortcomings of combining IP mobility solutions

We have identified, described and analyzed different combinations of IP mobility protocols standardized by the IETF and the functionality that they provide. An important result of the analysis is that, although the combination is needed to be able to obtain a mix of the properties of the different protocols, it also involves a cost, both from the point of view of additional overhead and from the point of view of handover delay.

Regarding handover delay, combining different mobility solutions has an impact on the performance, even when some of the steps required by each of the mobility protocols – such as movement detection – are shared, resulting in a longer handover interruption. We have also showed that this performance penalty can be significant in certain cases.

These results raise the need to develop mechanisms to alleviate the costs of combining different mobility solutions, while keeping the advantages brought by the combination. The experience teaches us that we cannot expect a single mobility solution with all the functionalities to be universally adopted. Instead we advocate for solutions that include in their design the flexibility to activate and deactivate the use of their supported mobility functionalities (as already supported

[9]This bug has already been reported to the software developers.

in a very limited way by the latest 3GPP specifications [5]). For example, it would be interesting to design a mechanism so that the mobile nodes and the network can negotiate about the mobility support functions required in a particular situation, choosing a particular set of functions from those available as required. This implies placing a very active role in the mobile nodes regarding the use of mobility solutions. This is in contradiction with part of the motivation for the network-based mobility approach that has been favored by some operators. However, in fact, the heterogeneity of access networks and the current trend towards environments with several options for network access from different operators make unavoidable to place more responsibility in the mobile nodes regarding the mobility solutions to use, at least if they want to enjoy an optimized performance adapted to their needs but compatible with network requirements.

Recently there has been an interest in mobility solutions that push the mobility anchors to the network edges (closer to the mobile nodes), what is called Distributed Mobility Management (DMM) [14]. This work is being carried out at the IETF, and at this stage, is first properly scoping the problem of DMM as well as analyzing how current standardized protocols could be used to achieve a DMM-alike behavior. The reasons for this approach are distributing the functionality to avoid bottlenecks and single points of failure, keeping mobility signaling close to the MN, improving handover efficiency because mobility signaling is local (and therefore faster), and optimizing data path routing (the anchor is close to the MN so, in general, long diversions through a far-off anchor are avoided). But, interestingly from our point of view, the DMM approach also assumes a more active role in the MN, that chooses how to use mobility support functionality on a per-communication basis. We are going to use a particular instantiation of this approach to show how building this flexibility in mobility mechanisms can help in minimizing the negative effects of combining different IP mobility schemes.

6.1. Distributed Mobility Management

The mobility management approaches described in Section 2 have in common the use of a centralized anchor point in charge of defending the MN's address and forwarding the packets addressed to the mobile node towards its current destination. The use of this central anchor point has several drawbacks, such as longer (sub-optimal) routing paths, scalability issues or longer handover latencies. Additionally, not every IP application requires IP address continuity (i.e., to keep the IP address across movements), and therefore it is preferable to enable mobility on a per-application basis. These are the issues that have motivated the development

of the Distributed Mobility Management (DMM) approach. As part of this effort, needs for new mechanisms and protocols are being identified.

Although the DMM paradigm is relatively new, there have been already proposals for distributed mobility management solutions. These can be classified into two main groups: those that aim at making client IP mobility approaches (such as MIPv6) work in a distributed way [15], and those that aim at doing that for proxy network-based mobility (like PMIPv6) [16]. Both approaches follow the same basic principle: access routers (i.e., the first-hop router the mobile node attaches to) perform two main functionalities: *i)* provide attaching terminals with valid IP addresses that are topologically valid at the access router, and *ii)* anchor these addresses and keep their reachability even when the mobile node moves. In order to do so, the access router may also have to implement the HA or MAG/LMA functionality (depending on the approach followed) for certain IP addresses. It is important to note that the activation of the IP address mobility management is performed on demand on an IP address basis, depending on the requirements of the applications running on the user terminal that are using that given IP address. We provide next an example: an MN attaches to an access router (*AR1*), and obtains an IP address/prefix topologically correct (ip_1) at that location. If the MN starts any applications, they will make use of ip_1 as source IP address for the new communications, thus benefiting from direct (optimal) routing to the correspondent node. We consider next the situation in which the MN moves to a new access router (*AR2*): after attaching, the MN obtains a new valid address/prefix (ip_2), which is used by any new connections. Besides, in case the MN requires IP address continuity (e.g., because it has ongoing communications using ip_1 that cannot survive an IP address change), the previous access router (*AR1*) implements HA/LMA functionality, and keeps defending the IP address ip_1, and forwarding packets addressed to the new location of the MN.

6.2. Combining IP mobility protocols in a dynamic and distributed fashion

In order to clarify the requirements and operation of the new DMM paradigm, this section presents a DMM example. In [17] a particular DMM solution is proposed, which is being developed within the context of the EU MEDIEVAL project[10]. This solution, as any DMM solution, allows mobile nodes to choose when to use, or not, mobility support for each of their applications. But, additionally, it also offers two different types of mobility support solutions: a global and

[10] `http://www.ict-medieval.eu/`

a local one. The mobile node can choose to use one solution, the other or none, for each of their applications, which provides the advantages of both solutions without incurring in the inefficiencies identified in Section 4.

We are going to give a brief description of the solution described in [17]. Note that a detailed description is out of the scope of this article, we are just going to highlight the properties in this solution that enable an efficient combination of different mobility mechanisms, in this case a global and a local mobility solution.

The MEDIEVAL mobility architecture leverages on the concept of Distributed Mobility Management [18], for the development of both network-based and client-based mobility management. The access network is organized in Localized Mobility Domains in which the network-based scheme inspired by [16] is applied. Users are expected to be most of the time roaming within a single LMD, but, for those cases where this is not possible (e.g., roaming to a network owned by a different operator or running a different mobility support scheme), a client-based DMM approach is followed (e.g., based on [15]). In order to integrate both approaches, so a mobile node can simultaneously have sessions managed by a network-based ("PMIPv6 alike") approach and a client-based ("MIPv6 alike") approach, a novel architectural element called Mobile Access Router (MAR) is introduced. The MAR is a network entity implementing all the functionalities of its counterparts in the standard mobility protocols (MIPv6 and PMIPv6), so it is able to play the role of plain access router, home agent, local mobility anchor and mobile access gateway on a per address basis.

A mobile node obtains a locally anchored IP address every time it attaches to a MAR. While moving between MARs of the same LMD, the reachability of these addresses is maintained – if needed by an application – by using a PMIPv6 based DMM mode of operation. Note that in this mode, and for each IP address, the anchoring MAR plays the role of LMA, and the currently serving MAR plays the role of MAG. On the other hand, if the MN moves out of an LMD, it enables its MIPv6 stack, obtains a new care-of address and signals its current location to the relevant MARs (i.e., those that anchor IP addresses which reachability should be maintained). Note that in this case, these MARs play the role of HA for those addresses.

We are not arguing here that DMM (or the MEDIEVAL DMM approach) is the perfect mobility solution. In many scenarios it has the advantage of reducing the distance between the MN and its anchor point in use, which improves efficiency. But it may also have drawbacks. It pushes functionality/complexity to the edge of the network. It can also be very inefficient for MNs moving and opening several long-term communications, because that can result in tunnels from the

MN to several network nodes, tunnels that can become quite long with time, as the MN moves away from its initial position. These long tunnels mean losing the performance advantage of the DMM solution.

The interesting point for us is that we can identify key elements in the MEDIEVAL architecture that enable an efficient combination of mobility mechanisms:

- In MEDIEVAL mobile nodes are provided with the ability to choose, for each of their applications, the mobility mechanisms that are used to support their mobility. This is a feasible feature for client-based mobility approaches, but it is usually not present in other situations such as network-based mobility approaches or for hosts connected to mobile networks.

- An information system based on IEEE 802.21 [19] is defined so network nodes and mobile nodes can exchange information about the mobility support requirements of the mobile node, and to request the activation or deactivation of certain functions in the network, such as the network mobility support for a particular prefix.

- Mobile nodes not aware of the MEDIEVAL approach (e.g., unable to communicate with the network through IEEE 802.21) still receive, transparently for them, a default mobility support service based on PMIPv6 in this case.

Generalizing from this example, to achieve an efficient coexistence of mobility approaches that offers the advantages of their different functionalities without having performance penalties, we need the following:

- Flexibility for activating/deactivating particular mobility functionalities.

- A system for exchanging the required information among the mobile nodes and the network.

- A default behavior to provide mobility support to legacy nodes.

Ideally, the flexibility to activate/deactivate functions should be a built-in characteristic in any (new) mobility solution developed, and the information system should be standardized independently of the mobility solutions. Legacy nodes in this context means nodes that cannot exchange information and negotiate with the network the mobility support mechanisms to be activated. These legacy nodes must be able to benefit from a default mobility support functionality, even if it

is with some inefficiencies. The default mobility support functionality will depend on the situation, examples are a MIPv6 service, a PMIPv6 service, or a MIPv6+PMIPv6 combination. The important point is that being unaware of the information system should not make a mobile node unable to use the basic mobility functionality, even if it is with some inefficiencies.

7. Conclusions

Many IP mobility support protocols have been and still are being developed by several standardization bodies and, in particular, by the IETF. Each protocol provides a different functionality (terminal mobility or network mobility, for example) and/or require operations in different nodes (mobility support based on the terminal or on the network). The current trend in the evolution of mobile communication networks is towards terminals with several network interfaces (these are already a market reality) that get ubiquitous Internet access by dynamically changing access network to the most appropriate one. Handovers between different access networks are going to be usual. In this situation the different mobility solutions developed by the IETF are going to co-exist. There is no winner solution because each of them addresses different requirements.

In some scenarios we need a combination of different mobility solutions because a mix of some of their properties is required. This article has analyzed the different combinations and the functionality that they provide. But the combinations also have a cost. The article evaluates this cost for the different combinations of mobility solutions, both from the point of view of the overhead and from the point of view of the handover delay. The conclusion is that the combination of solutions may have an important impact on the overhead and handover delay of the communications, leading to performance penalties which can be significant in certain cases. A first contribution of this paper is an analytic study of these penalties so designers of mobility solutions can consider them when addressing an scenario that requires a combination of different mobility protocols. The findings of this analytic study have been confirmed, for the case of the handover delay, by an experimental evaluation conducted in a Linux-based testbed.

The second contribution of this article is giving an outline of how future mobility protocols can be designed, or previous ones adapted, to facilitate the combination of different mobility protocols. The key properties required are: flexibility to allow the activation and deactivation of mobility functionalities according to mobile and network requirements, the existence of a system so the network and the mobile nodes can exchange the information needed to support that activation

and deactivation, and the provision of a default mobility support service for legacy terminals. We have briefly summarized in this article a solution that integrates a PMIPv6-based with an MIPv6-based distributed mobility management solution using this approach.

Future work includes the definition of the information system in an independent way from the mobility solutions. The proposed behavior –mobile nodes negotiating with the network which mobility support functionalities to use– has to be achieved in an automatic way. It would be completely unrealistic to have users involved in choosing mobility mechanisms, but their preferences must be taken into account. We intend to explore the use of cognitive networks approaches, such us pattern-based agreement techniques, to implement this automatic negotiation.

Acknowledgments

We want to thank the anonymous reviewers for their useful comments which have helped to improve this article.

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7-ICT-2009-5) under grant agreement n. 258053 (MEDIEVAL project) and also from the Spanish MICINN through the I-MOVING project (TEC2010-18907).

[1] C. Perkins, D. Johnson, J. Arkko, Mobility Support in IPv6, RFC 6275, 2011.

[2] V. Devarapalli, R. Wakikawa, A. Petrescu, P. Thubert, Network Mobility (NEMO) Basic Support Protocol, RFC 3963, 2005.

[3] S. Gundavelli, K. Leung, V. Devarapalli, K. Chowdhury, B. Patil, Proxy Mobile IPv6, RFC 5213, 2008.

[4] I. Soto, C. J. Bernardos, M. Calderon, A. Banchs, A. Azcorra, NEMO-enabled localized mobility support for internet access in automotive scenarios, IEEE Communications Magazine 47 (2009) 152–159.

[5] 3GPP, Architecture enhancements for non-3GPP accesses, TS 23.402, v11.2.0, 3rd Generation Partnership Project (3GPP), 2012.

[6] G. Giaretta, Interactions between Proxy Mobile IPv6 (PMIPv6) and Mobile IPv6 (MIPv6): Scenarios and Related Issues, RFC 6612 (Informational), 2012.

[7] P. Serrano, A. de la Oliva, C. J. Bernardos, I. Soto, A. Banchs, A. Azcorra, A CARMEN mesh experience: deployment and results, in: IEEE Workshop on Hot Topics in Mesh Networking, HotMESH'09.

[8] T. Pagtzis, Advanced IPv6 mobility management for next generation wireless access networks, Ph.D. thesis, University College London, 2005.

[9] R. Aguiar, A. Banchs, C. J. Bernardos, M. Calderon, M. Liebsch, T. Melia, P. Pacyna, S. Sargento, I. Soto, Scalable QoS-aware Mobility for Future Mobile Operators, IEEE Communications Magazine 44 (2006) 95 – 102.

[10] Y. Han, J. Choi, S. H. Hwang, Reactive handover optimization in IPv6-based mobile networks, IEEE Journal on Selected Areas in Communications 24 (2006) 1758–1772.

[11] R. G. Clegg, A practical guide to measuring the Hurst parameter, Arxiv preprint math/0610756 (2006).

[12] J. A. Hernandez, I. W. Phillips, Weibull mixture model to characterise end-to-end Internet delay at coarse time-scales, IEE Proceedings-Communications 153 (2006) 295–304.

[13] C. Vogt, M. Zitterbart, Efficient and scalable, end-to-end mobility support for reactive and proactive handoffs in IPv6, Communications Magazine 44 (2006) 74 – 82.

[14] R. Kuntz, D. Sudhakar, R. Wakikawa, L. Zhang, A Summary of Distributed Mobility Management, Internet-Draft (work in progress), draft-kuntz-dmm-summary-00, 2011.

[15] F. Giust, A. de la Oliva, C. J. Bernardos, Flat Access and Mobility Architecture: an IPv6 Distributed Client Mobility Management Solution, in: 3rd IEEE International Workshop on Mobility Management in the Networks of the Future World (Mobiworld 2011), in conjunction with The 30th IEEE International Conference on Computer Communications (IEEE INFOCOM 2011).

[16] P. Seite, Dynamic Mobility Anchoring, Internet-Draft (work in progress), draft-seite-netext-dma-00.txt, 2010.

[17] F. Giust, C. J. Bernardos, S. Figueiredo, P. Neves, T. Melia, A Hybrid MIPv6 and PMIPv6 Distributed Mobility Management: the MEDIEVAL approach, in: Sixth Workshop on multiMedia Applications over Wireless Networks.

[18] H. Chan, Problem statement for distributed and dynamic mobility management, Internet-Draft (work in progress), draft-chan-distributed-mobility-ps-02.txt, 2011.

[19] A. de la Oliva, A. Banchs, I. Soto, T. Melia, A. Vidal., An Overview of IEEE 802.21: Media-Independent Handover Services, IEEE Wireless Communications 15 (2008) 96 – 103.