RAMS: A Protocol Extension to PMIPv6 for Improving Handover Performance of Multicast Traffic

Luis M. Contreras
Universidad Carlos III de Madrid
Leganés, Spain
`contreras.uc3m@gmail.com`

Carlos J. Bernardos
Universidad Carlos III de Madrid
Leganés, Spain
`cjbc@it.uc3m.es`

Ignacio Soto
Universidad Politécnica de Madrid
Madrid, Spain
`isoto@dit.upm.es`

Abstract

IP multicast allows the efficient support of group communication services. The increase in the use of multicast has also triggered the need for supporting IP multicast in mobile environments. It is necessary to consider current unicast mobility management solutions in IP-based networks to smoothly extend multicast services to mobile users. Proxy Mobile IPv6 is a network-based mobility management solution, where the functionality to support the terminal movement resides in the network. Recently, a base solution has been adopted for supporting multicast in PMIPv6. Nevertheless, this solution has not special features to optimize performance during handovers and this can cause disruptions in the reception of multicast traffic by the mobile terminals. This paper proposes and evaluates RAMS, an extension to the PMIPv6 protocol for accelerating the multicast service delivery to mobile terminals after handovers. We explore the performance gain obtained by RAMS, compared against the base solution, using two different radio access network scenarios: 3GPP UTRAN and LTE.

1 Introduction

IP multicast allows the efficient support of group communication services (one-to-many or many-to-many) over existing IP networks. Applications like TV distribution take advantage of this extension of the IP protocol which basically facilitates the delivery of a single copy of a data stream to multiple listeners interested in receiving the same content simultaneously. The increasing generalization in the use of multicast has also triggered the need for supporting IP multicast in mobile environments where the receiver is mobile.

Mobility management support in IP-based networks is a topic that has received considerable attention and research effort in recent years thanks to the evolution and increasing popularity of mobile computing technologies and applications, allowing for the "always-on" and the "anywhere and anything" promises to become a reality. A first approach to cover this issue was centered on host-based solutions, mainly Mobile IPv6 (MIPv6) [1], where IP mobile terminals or Mobile Nodes (MNs) are aware of their IP mobility and have to perform operations in order to be able to maintain their ongoing communication sessions. As a further step, a network-based solution has been standardized, called Proxy Mobile IPv6 (PMIPv6) [2], where the functionality to handle the movement of mobile terminals reside in the network, and the terminals only perform the standard IP operations.

Group communication is out of the scope of the PMIPv6 standard specification. Recently, a base solution has been adopted for multicast service delivery in PMIPv6 domains [3]. However, the base solution must be improved to meet some performance requirements [4], especially those referred to the

---

Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, volume: 2, number: 2, pp. 67-82

service quality perceived by the user, which is seriously affected by the disruption of the reception of multicast content during handovers.

This paper presents and evaluates a new proposal, named Rapid Acquisition of Multicast Subscription (RAMS) which extends the PMIPv6 protocol for accelerating the multicast service delivery to the MN after a handover. The structure of the paper is as follows: section 2 introduces some background on PMIPv6, multicast, and the base solution for multicast support in PMIPv6 that has been standardized; section 3 presents the RAMS proposal to decrease the delay for the effective multicast traffic delivery to the MN after a handover; section 4 characterize both the base solution and RAMS proposal for two different radio access network scenarios; section 5 identifies typical values for delay variables in the network under analysis; finally, section 6 compares both solutions and provides some conclusions.

2 Background on PMIPv6 and multicast

2.1 Proxy Mobile IPv6

PMIPv6 is based on MIPv6, reusing much of its concepts and packet formats. In PMIPv6 mobility support is provided by some specific network entities, namely Mobile Access Gateway (MAG) and Local Mobility Anchor (LMA).

The MAG takes care of the mobility signaling on behalf of the MNs attached to its links, tracking the MNs as they move, while the LMA stores all the routing information needed to reach the MNs in the PMIPv6 domain by associating each MN with the MAG that the MN is using. A tunnel between the LMA and the MAG allows the transfer of traffic from and to the MN. Using PMIPv6, the MN can move across a PMIPv6 domain changing its access link, while keeping its IP address.

The MN is registered by the MAG in the LMA by sending a Proxy Binding Update message (PBU, an extension of the MIPv6 Binding Update). The LMA then assigns one or more home network prefixes (HNPs) to the MN. The LMA acknowledges the registration process with a Proxy Binding Acknowledgment (PBA) message that is sent from the LMA to the MAG, containing the HNPs allocated to the MN. The MAG completes the configuration to serve the MN traffic by setting up the appropriate forwarding rules for the downlink/uplink traffic to/from the MN (PMIPv6 assumes the use of point-to-point access links).

The LMA forwards the traffic destined to the HNPs to the correct MAG using the configured tunnel. The MAG decapsulates the packets and forwards them to the MN transparently. In the opposite direction, traffic coming from the MN is encapsulated by the MAG (which is the MN's default router) and decapsulated by the LMA that will route it towards the final destination.

2.2 Multicast basics in access networks

By means of IP multicast, a number of receivers located anywhere in the network can subscribe to a content in the form of a multicast session group. The content is distributed using a particular data stream forming a multicast flow. A single copy of such flow is carried on every link in the network along the multicast path dynamically created to reach the interested receivers. The data stream is replicated on the routers where the multicast path topologically diverges.

The multicast source of the data stream does not maintain any subscription list of interested receivers. The source simply sends the data stream to an arbitrary group of hosts represented by an IP multicast address. The receivers indicate their interest on receiving certain content by explicitly joining the multicast group. The Multicast Listener Discovery (MLD) [5] defines the control messages for managing the group membership process. Multicast protocols distinguish between multicast receiver (host part) and multicast router (network part) functionalities. Basically, the host part is devoted to the group subscription management, while the router part is focused on building and maintaining the multicast distribution tree.

A multicast router in the receiver's sub-network will capture the control messages for joining or leaving a multicast group. In some cases, the router can act as a proxy [6] for the group membership indications of the receivers connected to it, instead of the multicast router role described above. This typically occurs in aggregation networks, where the first-hop router concentrates the traffic of a huge number of receivers. The proxy performs the router part of the group membership protocol on each downstream interface, while it plays the host role on the upstream interface towards the next multicast backbone router. The proxy is in charge of summarizing the subscription demand of the receivers, acting as a unique host towards the upstream multicast router.

2.3 Multicast in PMIPv6

Multicast is out of the scope of the PMIPv6 standard specification. This produces inefficiencies when distributing contents to multiple receivers (individual copies of the same stream per MN) and when supporting MN mobility (slowed dynamic adaptation of multicast distribution tree in case of handover).

Several solutions have been already proposed supporting multicast for Mobile IP networks [7]. However, they are not directly applicable to PMIPv6 because of the specificities of network-localized management environments, where the MN is not aware of network layer changes. A new approach is then needed to provide multicast in PMIPv6 domains. With such aim, the MULTIMOB working group was chartered at the Internet Engineering Task Force (IETF) to specify a solution for multicast listener mobility compatible with current PMIPv6 and multicast protocol standards, that is, without any protocol extensions.

An additional constraint is observed regarding the network from where the MN obtains the multicast stream it subscribes to. Basically, two kinds of solutions can be differentiated: the remote subscription, where the MN gets the multicast data stream from the LMA, and the local subscription, where the MN directly obtains the multicast stream from the access router. According to the original MULTIMOB charter terms, only the remote subscription case was considered for the released standard.

The base solution [3] provides a way to manage multicast traffic delivery to MNs unaware of their mobility. The MN expresses its interest in joining or leaving a multicast group by sending MLD control messages to the MAG, which acts as the first hop at the point-to-point link. The MAG will maintain the individual multicast status of the interface for that link and will handle the multicast traffic towards the MN accordingly to the MLD messages received. In the base solution the MAG incorporates MLD proxy functionality. As a proxy, the MAG is responsible of summarizing the group subscription requests of the MNs connected to it.

With the remote subscription model, the multicast traffic reaches the MNs after going through the corresponding LMA (note that there might be multiple LMAs in the same domain). A distinct MLD proxy instance is then defined per LMA connected to the MAG, in such a way that every MAG-LMA tunnel is part of a separate MLD proxy domain. For every proxy instance in the MAG, the tunnel interface pointing to the LMA becomes the proxy upstream interface, whereas the links towards the MNs are the corresponding downstream interfaces of each instance. The LMA is the entity in charge of interacting with the multicast infrastructure out of the PMIPv6 domain.

The summarization of control messages in upstream that the MAG performs is applied per set of MNs associated with a certain LMA, as the different proxy instances of the same MAG are isolated one from the other. The LMA maintains the multicast state of every tunnel interface. Such status reflects the summarized view offered by the MAG on behalf of the attached MNs bound to the LMA. A multicast stream will be delivered over the tunnel or removed from it according to the aggregated behavior of the

MNs attached to the MAG.

The LMA, the MAG and the tunnel linking them are all part of the multicast tree built to distribute the multicast traffic. This branch will be common to every multicast tree providing any content subscribed by an MN in a MAG and associated to a particular LMA. It makes possible to send a single copy of a data stream per group of MNs demanding the same content.

The handover process involves the set up of a new point-to-point link at the new MAG (nMAG) where the MN attaches, and the release of the old one at the previous MAG (pMAG) where the MN comes from. In such a mobile environment, the MLD proxy instances in the MAGs should be able of adding and removing downstream interfaces dynamically.

Figure 1 shows the complete process. Following the expected standard behavior of MLD for a new link set up, once the MAG has completed the configuration of the interface on the new point-to-point link (that is, after receiving the PBA message from the LMA for this MN), the MAG kernel immediately sends an MLD General Query. At this moment the point-to-point link still has not been configured as part of the downstream set of interfaces of any MLD proxy instance. The decision of what proxy instance corresponds to the new attached MN will be based on the LMA serving the MN, as there is one instance per LMA. Once the link is configured as a new downstream interface of the corresponding proxy instance, the proxy instance sends an additional MLD Query for getting knowledge of any active multicast subscription by the MN.

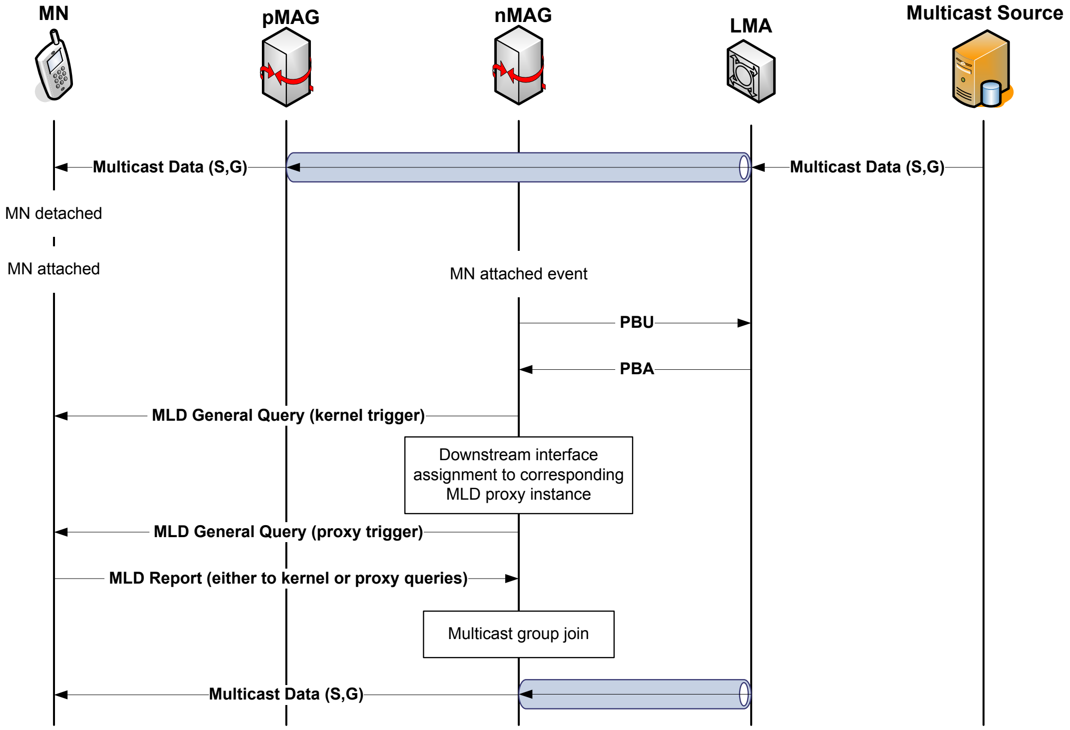


Figure 1: MN handover in base solution

When the MN receives the MLD Query, the MN provides information about the active memberships it maintains, if any, in the form of an MLD Report message. The nMAG, upon receiving the Report message, will pass it to the assigned proxy instance for further processing (if the assignment to a proxy instance has not been completed, the Report message is discarded). Finally, the proxy instance will set up the multicast status of the downstream interface if one or more subscriptions are active, or will not take any further action in other case.

In case of the content requested by the entering MN is currently being received by another MN managed by the same MLD proxy instance, the content is directly delivered to the receiver, without triggering an MLD Report upstream. If not, a change in the group membership database is done, and the nMAG sends an MLD Report upstream requiring such flow. The LMA will update the multicast status of its downstream interfaces and will proceed to forward the stream on the tunnel to the nMAG.

3 Description of RAMS proposal

The method used in the base solution to get knowledge of an existing multicast subscription relies on the intrinsics of the MLD protocol, which sends multicast membership interrogation messages once a new link is up. The answer to that request message will report any active multicast subscription by the MN. Due to this behaviour, despite of being a straightforward method, the MAG can incur in a huge delay in receiving the corresponding MLD Report message due to either the MLD processing time or the radio transfer delays.

This paper presents RAMS, a new approach to handle multicast traffic during handovers in PMIPv6 domains, originally described in [8], complementing it with a sound evaluation and comparison with the basic solution. RAMS proposal consists of the extension of the current PMIPv6 signaling protocol by including: (a) a new multicast information option to keep PMIPv6 entities updated during registration and de-registration processes; (b) new messages to trigger the transfer of such multicast information among PMIPv6 entities; and (c) some flags to govern the global process. No extensions are required for the multicast protocols.

The RAMS signaling flows are internal to the network. In this way, the knowledge by the nMAG of the currently active multicast subscription becomes independent of the underlying radio interface dynamics and relaxes the requirement of a rapid response from the MN in processing MLD messages. The independence of the radio access technology is specially relevant when considering a seamless service provision by means of distinct radio access networks with different performance in terms of latency and throughput [9], that will impact on the time in which the nMAG acquires the information relative to the ongoing multicast subscription.

The LMA tracks the MN along the domain pointing to the MAG where it is attached, therefore the LMA is the best element to store and forward the multicast subscription information to the MAGs as the MN moves. The LMA only needs to know the subscription information (the IP addresses of multicast group, G; and source, S) during the handover event. Apart from that event, it is not worthy to continuously inform the LMA about it.

As the MN moves from pMAG to nMAG, the mobility-related signaling due to the handover is carried out independently by both access gateways. Thus, two scenarios should be considered depending on the order in which the LMA receives notification of the MN registration and de-registration in the nMAG and the pMAG respectively.

Predictive handover. In the predictive case, the LMA firstly receives the MN de-registration from the pMAG before receiving the MN registration from the nMAG. Only for those MNs which maintain an active multicast subscription, the pMAG will include as part of the de-registration PBU message a new mobility option carrying the IP addresses of the multicast subscription(s) active in the MN at that moment.

The LMA will store that information in the corresponding binding cache. If, later on, the MN attaches to an nMAG, this information will be sent to the nMAG as part of the (extended) PBA message to complete the registration process. On the other hand, if no further registration happens, the multicast information will be removed together with the rest of the binding database for that MN.

After receiving the IP addresses of the groups subscribed by the MN, and the sources delivering them, the nMAG can subscribe the multicast flows on behalf of the MN, if there is no other MN receiving them already at the nMAG. The multicast status can be also set up in advance for the point-to-point link towards the MN. Figure 2 summarizes this process.

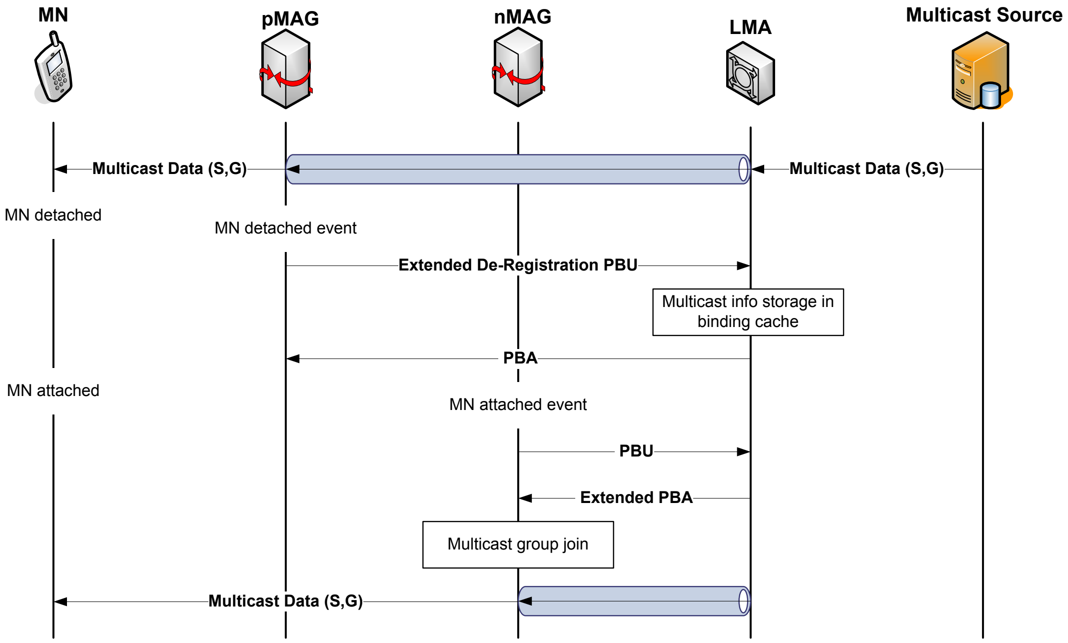


Figure 2: RAMS proposal - predictive handover

Reactive handover. In the reactive case, the LMA receives the MN registration from the nMAG without having previously received the MN de-registration from the pMAG. As the nMAG is not aware of any active multicast subscription of the MN, it will start a conventional registration process sending a PBU to the LMA.

After receiving the PBU message, the LMA will take the decision of interrogating, or not, the pMAG about any existing multicast subscription for that MN. The interrogation process is done by using a new pair of messages named Subscription Information Query and Response, respectively.

Once the multicast subscription information is retrieved from the pMAG, the LMA encapsulates it in the (extended) PBA message forwarded to the nMAG. Then, the nMAG can subscribe the multicast flows on behalf of the MN, if there is no other MN receiving them already at the nMAG. The multicast status can be also set up in advance for the point-to-point link towards the MN. Figure 3 presents the corresponding signaling flow.

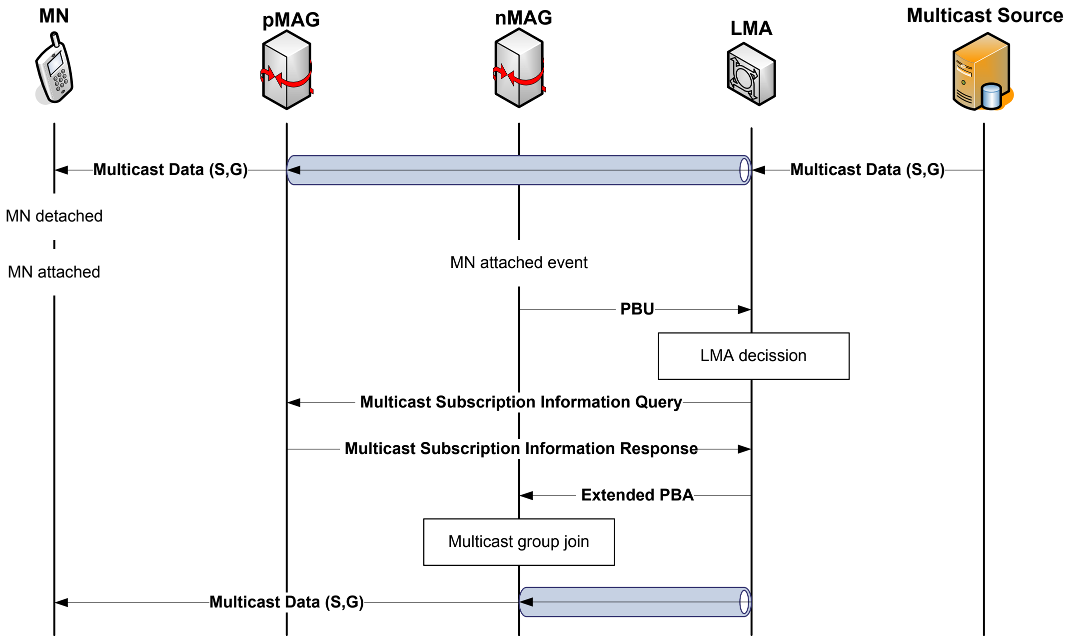


Figure 3: RAMS proposal - reactive handover

4 Base solution and RAMS proposal characterization

4.1 Multicast subscription acquisition delay

A mobile terminal maintaining an active multicast subscription moving from one point-of attachment to another within a PMIPv6 domain will experienced a certain delay in receiving again the multicast content that it was previously receiving at the primary location. Such delay will cause a gap on the content reception. The global delay due to the handover can be split in three main contributions.

- The delay originated by the reattachment process from the pMAG to the nMAG, T_{att}, mainly due to the layer 2 mechanisms for handover support.
- The delay caused by the IP layer location update process, T_{loc}, which has to do with the mobility management process, (including PBU/PBA messages, authentication, routing table update, etc).
- The delay due to the multicast subscription information acquisition by the nMAG, T_{acq}, produced by the fact that the nMAG is not aware of any ongoing multicast subscription (the MN does not re-signal the active subscriptions because it is not aware of the movement).

While the two former delay contributions are common to the unicast case, the last one is specific to multicast, and is the one our proposal tries to minimize. We will focus our analysis on it.

The delay T_{acq}, related to the acquisition of multicast subscription information for both solutions, is roughly split in four main components, as shown in the Figure 4. Table 1 shows the delay parameters under consideration and the solution impacted by each of them.

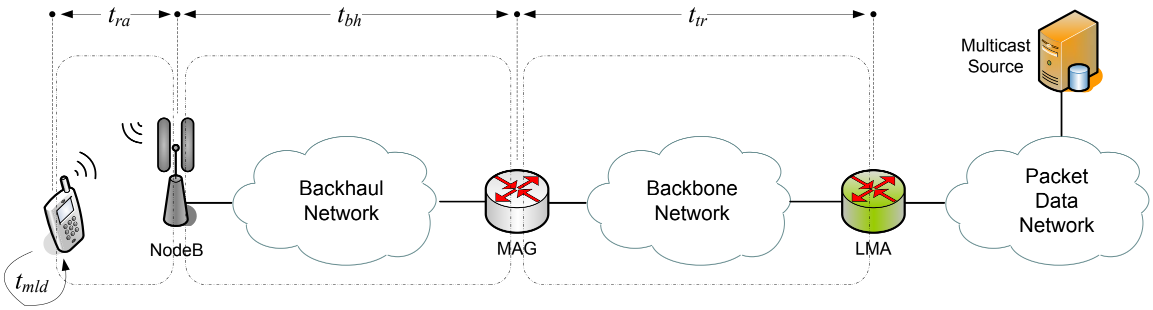


Figure 4: Delay components affecting the acquisition time

4.2 Network scenario under consideration

Mobile operators gradually introduce new technology to cater for customers increasing service demand, driving to the coexistence of several different technologies usually supported by separated networks[10]. Existing cellular architectures allow the deployment of areas served by different wireless technologies. Users equipped with multi-standard terminal will move seamlessly, and will expect to receive similar performance independently of the radio technology they are using on the current point of attachment.

Parameter	**Definition**	**Impact on**
t_{ra}	Radio Access delay	Base Solution
t_{bh}	Backhaul Network delay	Base Solution
t_{tr}	IP Transport Network delay	RAMS
t_{mld}	MLD Processing delay	Base Solution

Table 1: Delay components of interest

In order to compare both solutions we will focus on a scenario based on the Enhanced Packet Core (EPC) of 3GPP Evolved Packet System (EPS) standard, where PMIPv6 is used as mobility management protocol [11]. EPC supports several distinct radio access networks, including 3GPP Long Term Evolution (LTE) and UMTS. A deployment scenario of PMIPv6 in EPC is presented in Fig. 5.

We will center our analysis on the delays introduced by the network under the PMIPv6 operation, highlighting the distinct contributions. This will allow to perform a differential analysis to obtain the performance gain of the RAMS proposal.

4.3 Delay components for the base solution

The timing for multicast subscription information acquisition by the MAG for the base solution is presented in Fig. 6. For the base solution, T_{acq} is due to the transfer of MLD control messages over the radio channel (user plane) between the nMAG and the MN, as well as the time taken by the MN to process and answer the MLD Query. This can be expressed as:

$$T_{acq} = 2 \cdot (t_{ra} + t_{bh}) + t_{mld}. \tag{1}$$

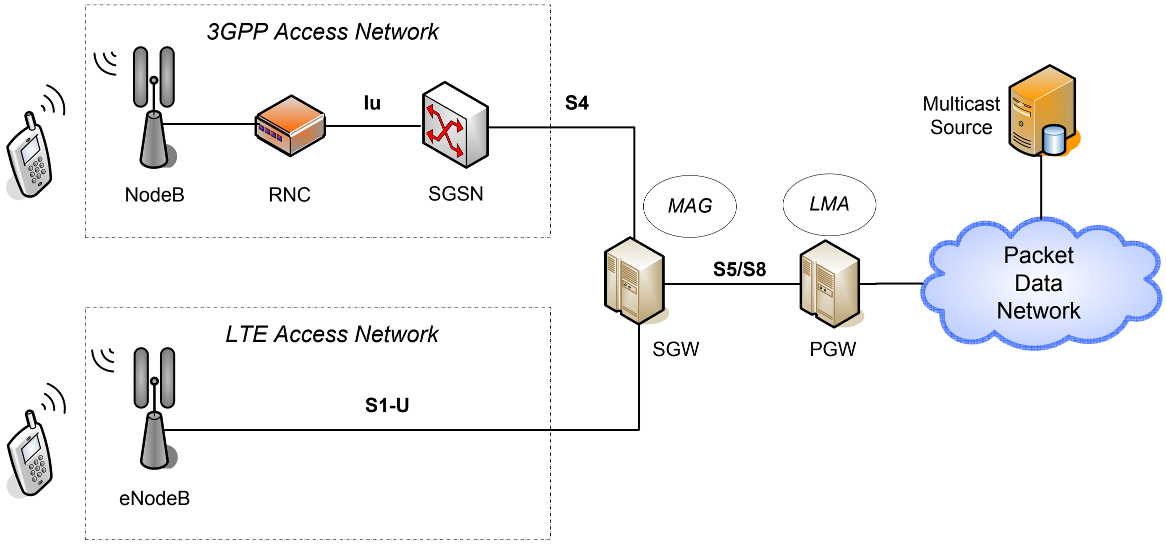


Figure 5: Scenario under analysis

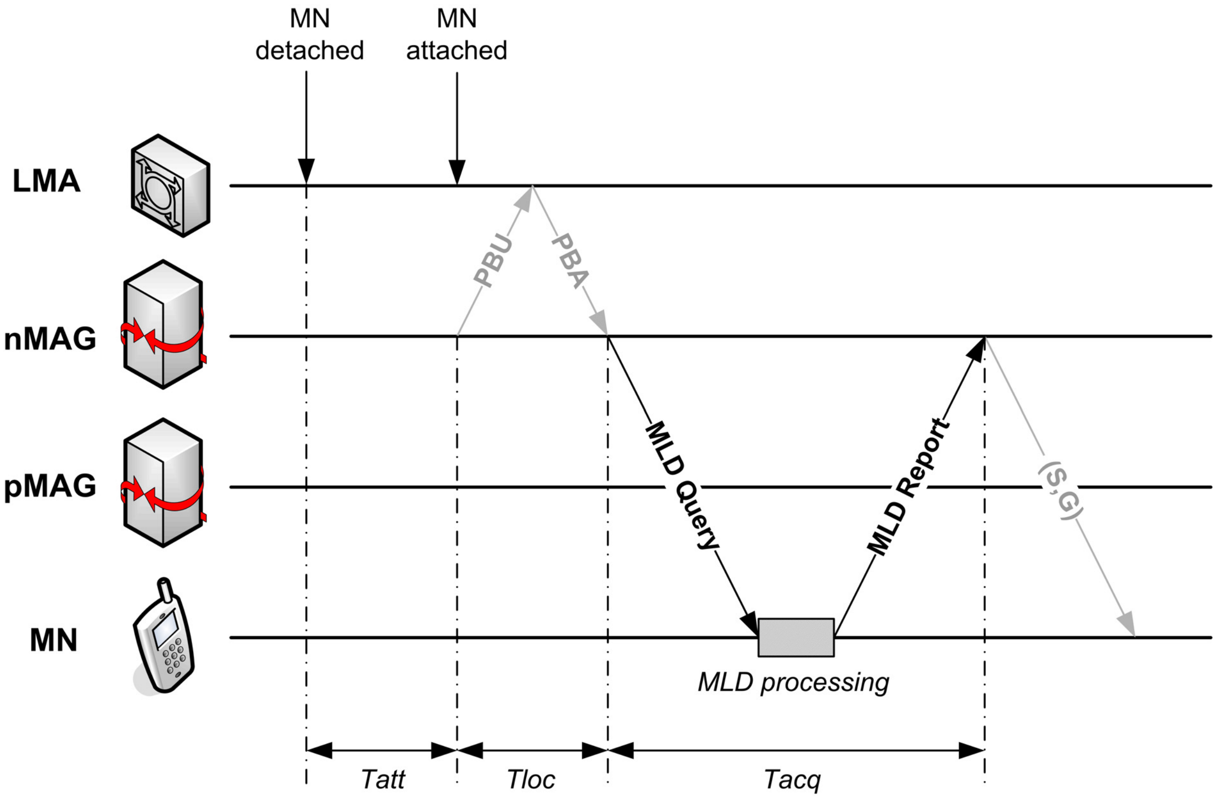


Figure 6: Acquisition delay for the base solution

4.4 Delay components for RAMS

In predictive mode, the registration process (Fig. 7) already integrates the multicast subscription information. As consequence, there is no contribution to T_{acq} (i.e., it is 0 ms).

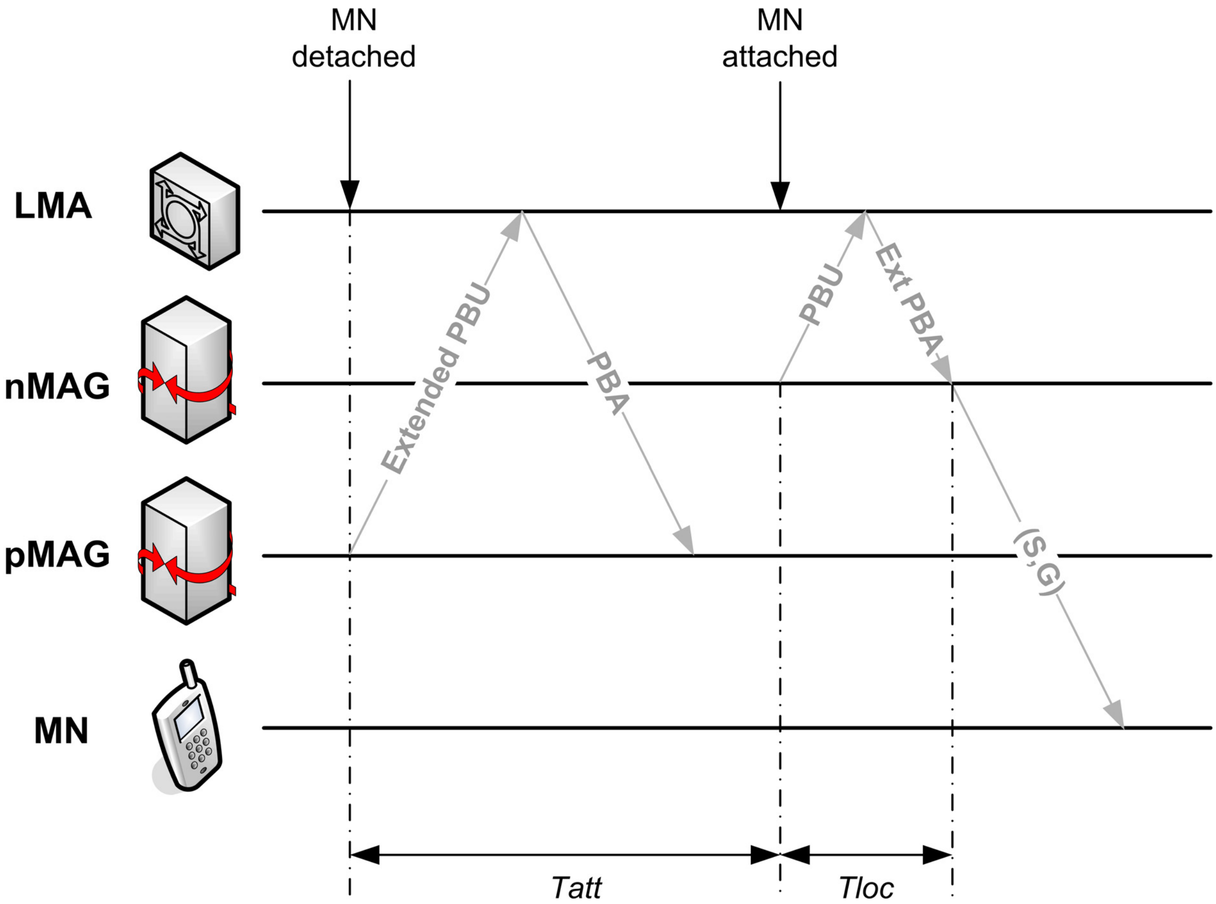


Figure 7: Acquisition delay for RAMS - predictive handover case

For the reactive handover case, there exists an acquisition period by using the new defined messages to interrogate the pMAG about the active subscription of the MN, as shown in Fig. 8. Then, the acquisition time can be formulated as:

$$T_{acq} = 2 \cdot t_{tr}. \tag{2}$$

4.5 Differential analysis of the base solution and RAMS

As can be seen above, the acquisition time differs for every solution. Then, the potential savings, if any, can be directly obtained from the difference of the delay of both solutions, as shown in the following equation:

$$T_{diff} = \begin{cases} 2 \cdot (t_{ra} + t_{bh}) + t_{mld}, & (predictive) \\ 2 \cdot (t_{ra} + t_{bh}) + t_{mld} - 2 \cdot t_{tr}, & (reactive) \end{cases} \tag{3}$$

We will then characterize the delay components to obtain the results of the differential analysis.

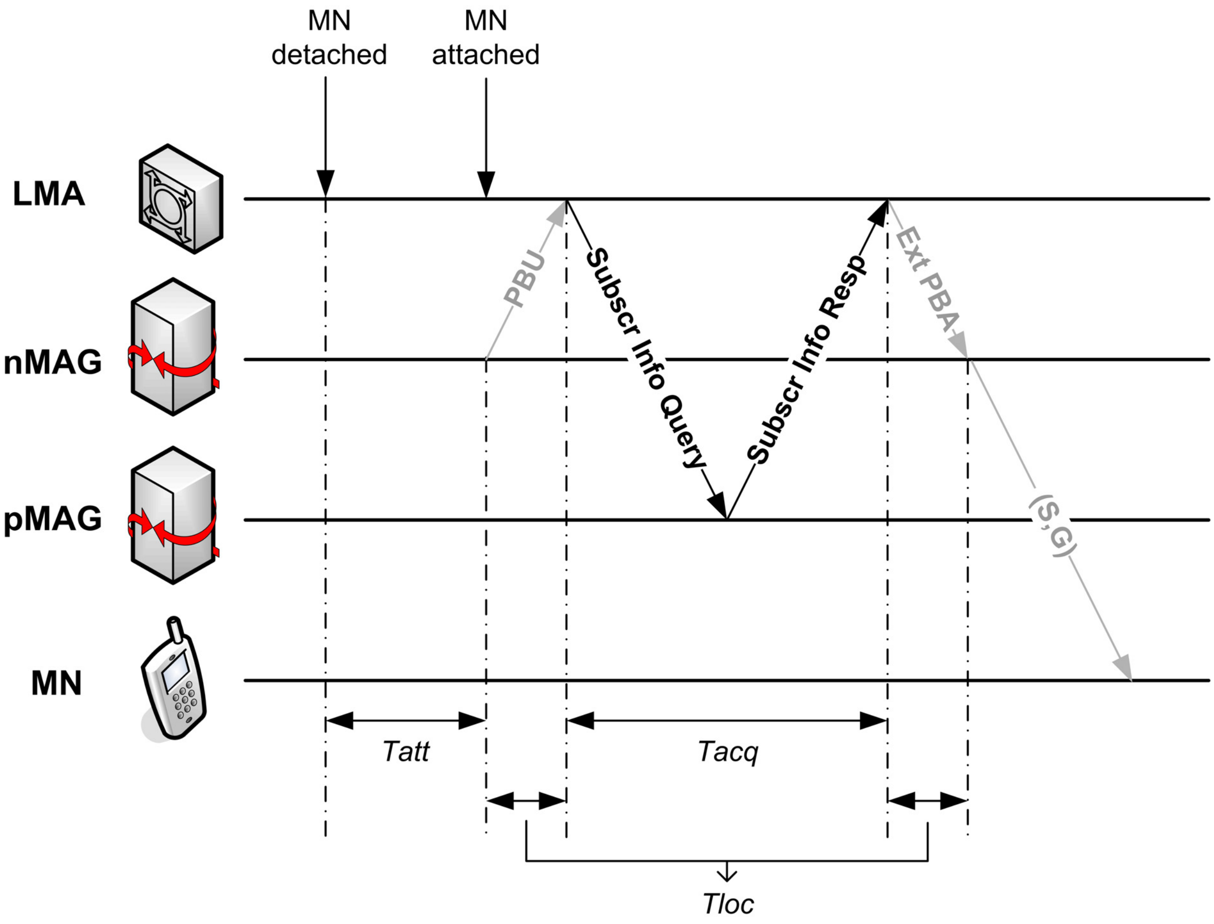


Figure 8: Acquisition delay for RAMS - reactive handover case

5 Delay variables parametrization

The following subsections provide an insight on the parameters contributing to the differential delay in both solutions. Further considerations have to be taken into account depending on the particular system under study. In our case, the specificities of both LTE and 3GPP UTRAN accesses will be taken into account.

5.1 LTE radio access and backhaul network delays

Radio access delay. The expected latency or round trip time for an LTE radio access network is expected to be less than 10 ms, according to [12]. To assess the LTE access delay we refer to the tests reported in [13], [14], and [15]. In [13] a latency of 10 ms between the mobile terminal and the eNodeB is reported considering the transmission in ideal conditions of a little IP packet (a ping command with 64 bytes of payload). In [14], the measurement is extended from the mobile terminal to a co-located Signalling Gateway (SGW), obtaining around 12-13 ms of latency. Finally, the live trial reported in [15] confirms that the LTE latency is in the order of magnitude of the one defined in [12].

The LTE radio access user plane (U-plane) delay consists of node processing delays, Transmission Time Interval (TTI) duration, and radio frame alignment. As stated in [16], under the hypothesis of having no errors in the radio channel the LTE U-plane latency can be written as:

$$t_{ra}(ms) = 3,5 \tag{4}$$

This estimation of the delay assumes unload conditions (i.e., no scheduling delays) and scheduling grant availability (no random access procedure is required), which, together with the ideal radio conditions considered above, leads to a best case assumption. In the rest of the paper we will consider this last value as the contribution of the LTE radio access networl to the total (one-way) delay.

Backhaul delay. The backhaul segment provides the transport resources required to build the network capillarity needed to cover the users' footprint. It spans across areas that are usually poorly equipped, in comparison to high capacity core segments. When choosing backhaul technology, key criteria such as cost, reachability, capacity, and availability of resources (e.g. frequency spectrum, optical fibre) have to be considered.

In mobile networks, the backhaul corresponds to the transport segment between the base stations and the first aggregation node. In the case of LTE, the backhaul extends from the eNodeB to the SGW (S1 interface). There are several solutions for such connection in terms of network topology (ring, star, point-to-point, etc) and technology (optical fiber, microwave transmission, xDSL-based accesses, etc), all of them having distinct properties in terms of performance, reliability and delay. They typically coexist in a real mobile network, in such a way that an MN changing the point of attachment can pass smoothly from one solution to another.

All this variety of technologies and architectures makes difficult to determine an standard delay value for the mobile backhaul segment. In [16] a rough indication of the delay for the S1 interface (at user plane) is provided in the form of the range (1ms, 15ms). A further recommendation [17] establishes in 10 ms the maximum end-to-end two-way delay for the backhaul network. Finally, a more precise value for the S1 (user plane) delay is assigned in [18], where it is considered a value of 5 ms. This last value will be the one considered in this paper, as an average measure of the delays found in real network deployments.

5.2 3GPP UTRAN radio access and backhaul network delays

In coherence with the LTE case, we will assume also here ideal conditions for the radio channel (best case scenario).

To estimate the delay in the access part of an UMTS network we take into account the budget delay estimation described in [19], considering delays due to the processing in different nodes (UE, NodeB, RNC), the radio propagation delay, and the interleaving and queueing.

Additionally, it is needed to add the contribution of the interconnection among the UTRAN and the EPC. Assuming no delay on the S4 interface (this could be the case of a SGSN co-located to the SGW), we will focus on the contribution due to the SGSN itself. In [20] a survey of the delay found in a real mobile operator network is reported. We take as reference the bounding values at 99.8% -percentile for the SGSN (in UMTS), in Down-Link (DL) and Up-Link (UL) directions. Table 2 summarizes the contributions considered here for UMTS.

$t_{ra} + t_{bh}$ **delay (ms)**	**UTRAN Access [19]**	**SGSN [20]**
Uplink	65,3	2
Downlink	52,3	3

Table 2: Delay components of interest

5.3 MLD Query processing delay

The MLD protocol specifies a field in the MLD header, *Maximum Response Delay*, which is used in the MLD Query messages to separate in time the hosts' MLD Report responses in a sub-network, as a way of avoiding an avalanche of Report messages arriving to the router at the same time. Its default value is 10 ms. A host receiving an MLD Query message will wait a random time in the range (0, *Maximum Response Delay*) to send the MLD Report message.

The value of the Maximum Response Delay field is configurable through the *Query Response Interval* parameter. This variable is expressed in ms, and its default value is 10 ms. Nevertheless, on current widely deployed commercial equipments the minimum value for this parameter is 1 s (Cisco IOS 12.2SR, Juniper Junos 9.5). Without a particular tuning, it can be expected that the response delay would be among 0 and 1 s.

In this paper we will assume an immediate response of the MN to the received MLD Query. This would be the case of a MAG with a pre-configured value of 0 ms (best case) for the *Maximum Response Delay*.

5.4 IP transport network delay

The different elements in a network are typically connected through an IP transport network. An operational network as the one considered in our analysis will be part of a controlled environment, with an engineered parameterization able to ensure the performance objectives of a carrier network. The delay of an IP backbone is proportional to the number of nodes (routers) traversed. In [21] a commercial operator network is analysed to determine the delay due to a single hop. According to the results obtained there, the single hop delay is bounded to 1 ms as per 99th percentile. It should be noted that the study was carried out on a network implementing STM-1 and STM-4 capacity links. If we assume that new deployments using high capacity Ethernet links (1 Gb to 10 Gb) will become the standard for future backbone networks, including PMIPv6 ones, the maximum contribution of 1 ms can be seen as an upper limit. Similar delay boundaries (on average) have been reported for network tests under traffic congestion applying *diffserv* Quality of Service (QoS) mechanisms for traffic prioritization, which is a common procedure in real networks [22].

To account for delay of multiple hops in the network we will approximate the total contribution as if every router in the path adds the maximum delay reported in [21]. This is equivalent to consider that all the datagrams of a certain flow experiment the maximum possible delay. It has been demonstrated [23], [24] that this deterministic approximation to the total delay (proportional to the number of hops) is less accurate and exceeds an statistical delay calculation, leading in this case to a worst case scenario. Then it can be established that

$$t_{tr}(ms) = N \cdot 1 \tag{5}$$

being N the number of routers traversed by the message.

The study in [21] is based on an IPv4-based network. However, the performance of current network equipment is similar for IPv4, IPv6, or mixed traffics [25].

6 Comparison and conclusion

Table 3 recaps the distinct parameters and shows the expected savings obtained by the RAMS proposal.

Attending to the results in Table 3, it can be noted that the performance of RAMS reduces significantly (in the order of hundred of ms) the acquisition delay in networks with high to moderate radio

	Delay components (ms)				**Differentail delay T_{diff} (ms)**	
Radio Technology	t_{ra}	t_{bh}	t_{mld}	t_{tr} (5 hops)	Predictive case	Reactive case
UTRAN	55,3 (DL) / 67,3 (UL)		0	5	122,6	112,6
LTE	3,5	5	0	5	18	8

Table 3: Delay savings due to RAMS (in ms) per radio access technology

access delay. Additionally, it slightly improves the delay in networks with very low radio access delay. This is under the assumption of ideal conditions (no errors and no load in the radio access network, instant response of the MN to MLD Query messages, etc), a real characterization of the radio access could highlight a more significant gain by using RAMS.

Furthermore, the RAMS proposal makes the acquisition delay independent of the radio technology in use, allowing the mobile multicast listeners to experience a similar performance independently of the radio access network they connect while moving across a PMIPv6 domain.

7 Acknowledgements

The work of Carlos J. Bernardos has been partially supported by the Ministry of Science and Innovation of Spain under the QUARTET project (TIN2009-13992-C02-01) and the European Community's Seventh Framework Programme (FP7-ICT-2009-5) under grant agreement n. 258053 (MEDIEVAL project).

References

[1] D. B. Johnson, C. E. Perkins, and J. Arkko, "Mobility Support in IPv6," IETF RFC 3775, June 2004, http://www.ietf.org/rfc/rfc3775.txt.

[2] S. Gundavelli, K. Leung, V. Devarapalli, K. Chowdhury, and B. Patil, "Proxy Mobile IPv6," IETF RFC 5213, August 2008, http://www.ietf.org/rfc/rfc5213.txt.

[3] T. C. Schmidt, M. Waehlisch, and S. Krishnan, "Base Deployment for Multicast Listener Support in Proxy Mobile IPv6 (PMIPv6) Domains," IETF RFC 6224, April 2011, http://www.ietf.org/rfc/rfc6224.txt.

[4] D. von Hugo, H. Asaeda, B. Sarikaya, and P. Seite, "Evaluation of further issues on Multicast Mobility: Potential future work for WG MultiMob," IETF Internet-draft (work in progress), June 2010, http://tools.ietf.org/html/draft-von-hugo-multimob-future-work-02.

[5] R. Vida and L. Costa, "Multicast Listener Discovery Version 2 (MLDv2) for IPv6," IETF RFC 3810, June 2004, http://www.ietf.org/rfc/rfc3810.txt.

[6] B. Fenner, H. He, B. Haberman, and H. Sandick, "Internet Group Management Protocol (IGMP) / Multicast Listener Discovery (MLD) - Based Multicast Forwarding ("IGMP/MLD Proxying")," IETF RFC 4605, August 2006, http://www.ietf.org/rfc/rfc4605.txt.

[7] I. Romdhani, M. Kellil, H.-Y. Lach, A. Bouabdallah, and H. Bettahar, "IP mobile multicast: Challenges and solutions," *IEEE Communications Surveys and Tutorials*, vol. 6, no. 1, pp. 18–41, First Quarter 2004.

[8] L. M. Contreras, C. J. Bernardos, and I. Soto, "Rapid acquisition of the MN multicast subscription after handover," IETF Internet-draft (work-in-progress), June 2010, http://tools.ietf.org/html/draft-contreras-multimob-rams-00.

[9] B. Orlandi, E. Bizouarn, F. Taburet, F. Scahill, J.-L. Lafragette, R. Evenden, S. Ringland, S. Johnson, T. Melia, T. Twell, and Y. E. Mghazli, "Improving the Customer Experience for Heterogeneous Wireless Access," *Bell Labs Technical Journal*, vol. 15, no. 4, pp. 23–44, March 2011.

[10] S. Landström, A. Furuskär, K. Johansson, L. Falconetti, and F. Kronestedt, "Heterogeneous networks – increasing cellular capacity," *Ericsson Review*, no. 1, 2011,

http://www.ericsson.com/res/thecompany/docs/publications/ericsson_review/2011/heterogeneous_networks.pdf , last viewed June 2011.

[11] I. Ali, A. Casati, K. Chowdhury, K. Nishida, E. Parsons, S. Schmid, and R. Vaidya, "Network-based mobility management in the evolved 3GPP core network," *IEEE Communications Magazine*, vol. 47, pp. 58–66, February 2009.

[12] NGMN Alliance, "Next Generation Mobile Networks Field Trial Requirements," November 2010, http://www.ngmn.org/uploads/media/NGMN_Field_Trial_Requirements.pdf, last viewed June 2011.

[13] R. Irmer, H.-P. Mayer, A. Weber, V. Braun, M. Schmidt, M. Ohm, N. Ahr, A. Zoch, C. Jandura, P. Marsch, and G. Fettweis, "Multisite field trial for LTE and advanced concepts," *IEEE Communications Magazine*, vol. 47, pp. 92–98, February 2009.

[14] A. Harada, Y. Ofuji, and N. Okubo, "Overview of Super 3G (LTE) System and Experimental Results," *NTT Technical Review*, vol. 6, no. 11, pp. 1–9, November 2008.

[15] L. Bollea, M. Caretti, and V. Torrasi, "The Telecom Italia LTE trials: radiomobile system evolution towards 4G (In Italian)," *Notiziario Tecnico Telecom Italia*, no. 3, pp. 25–34, 2010.

[16] 3GPP, "Feasibility study for evolved Universal Terrestrial Radio Access (UTRA) and Universal Terrestrial Radio Access Network (UTRAN)," 3GPP TR 25.912 V8.0.0, December 2008.

[17] NGMN Alliance, "Next Generation Mobile Networks Optimised Backhaul Requirements," August 2008, http://www.ngmn.org/uploads/media/NGMN_Optimised_Backhaul_Requirements.pdf, last viewed June 2011.

[18] 3GPP, "Evolved UTRA and UTRAN; Radio Access Architecture and Interfaces (Release 7)," 3GPP TR R3.018 V1.0.0, October 2007.

[19] ——, "Delay Budget within the Access Stratum (Release 4)," 3GPP TS 25.853 V4.0.0, March 2001.

[20] P. Romirer-Maierhofer and F. Ricciato, "Towards anomaly detection in one-way delay measurements for 3G mobile networks: A preliminary study," in *Proc. of the 8th IEEE international workshop on IP Operations and Management (IPOM'08), Samos Island, Greece, LNCS*, vol. 5275. Springer-Verlag, September 2008, pp. 1–14.

[21] K. Papagiannaki, S. Moon, C. Fraleigh, P. Thiran, and C. Diot, "Measurement and Analysis of Single-Hop Delay on an IP Backbone Network," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 6, pp. 908–921, August 2003.

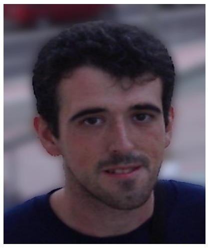
[22] C. Filsfils and J. Evans, "Deploying diffserv in backbone networks for tight SLA control," *IEEE Internet Computing*, vol. 9, no. 1, pp. 58–65, January 2005.

[23] D. D. Vleeschauwer, M. J. C. Büchli, A. V. Moffaert, and R. E. Kooij, "End-to-End Queuing Delay Assessment in Multi-service IP Networks," *Journal of Statistical Computation and Simulation*, vol. 72, no. 10, pp. 803–824, October 2002.

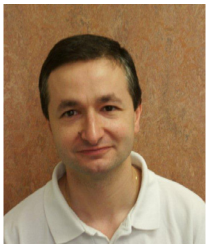
[24] M. Mandjes, K. van der Wal, R. Kooij, and H. Bastiaansen, "End-to-end delay models for interactive services on a large-scale IP network," in *Proc. of the 7th IFIP workshop on performance modelling and evaluation of ATM & IP networks, Antwerp, Belgium*, June 1999.

[25] Cisco White Paper, "Performance-comparison testing of IPv4 and IPv6 throughput and latency on key cisco router platforms," http://www.cisco.com/web/strategy/docs/gov/IPv6perf_wp1f.pdf, last viewed June 2011.

Luis M. Contreras is a Ph.D. student at the Universidad Carlos III de Madrid. He holds a Telecom Engineer degree from the Polytechnic University of Madrid (1997). He firstly joined Alcatel Spain taking several positions (R&D, standardization, systems design and customer engineering) in both wireless and wired network fields. Since 2006 he is with Orange Spain being part of the Network Strategy & Planning group, and serves as Peering manager for Orange Spain. .



Carlos J. Bernardos received a Telecommunication Engineering degree in 2003, and a PhD in Telematics in 2006, both from the University Carlos III of Madrid (UC3M), where he worked as a research and teaching assistant from 2003 to 2008 and, since then, has worked as an Associate Professor. His current work focuses on vehicular networks and IP-based mobile communication protocols. He has published has published over 30 scientific papers in prestigious international journals and conferences, and he is also an active contributor to the Internet Engineering Task Force (IETF).



Ignacio Soto received a telecommunication engineering degree in 1993, and a Ph.D. in telecommunications in 2000, both from the University of Vigo, Spain. He was a research and teaching assistant in telematics engineering at the University of Valladolid from 1993 to 1999. In 1999 he joined University Carlos III of Madrid, where he was an associate professor since 2001 until 2010. In 2010, he joined Universidad Politécnica de Madrid as associate professor. His research activities focus on mobility support in packet networks, heterogeneous wireless access networks, and automatic network management and operation.