Flat Access and Mobility Architecture: an IPv6 Distributed Client Mobility Management Solution

Fabio Giust∗†, Antonio de la Oliva†, Carlos J. Bernardos†
∗ Institute IMDEA Networks, Spain
† Universidad Carlos III de Madrid, Spain
E-mail: fabio.giust@imdea.org
E-mail: {aoliva, cjbc}@it.uc3m.es

Abstract—The use of centralized mobility management approaches – such as Mobile IPv6 – poses some difficulties to operators of current and future networks, due to the expected large number of mobile users and their exigent demands. All this has triggered the need for distributed mobility management alternatives, that alleviate operators’ concerns allowing for cheaper and more efficient network deployments.

This paper proposes a distributed mobility solution, based on Mobile IPv6 and the use of Cryptographic Generated Addresses. We analytically compare the solution to Mobile IPv6, and derive in which scenarios it performs best.

I. INTRODUCTION

The increasing demand of mobile data services from users is no longer a threat to operators, but a reality that needs to be tackled. We are witnessing that the number of wireless mobile subscribers accessing data services does not stop increasing. This is motivated by a variety of different reasons: 3G and WLAN accesses are widely available (combined, coverage reaches almost 100% of dense populated areas in developed countries) and affordable by users (most mobile handsets are 3G and WLAN capable, all laptops and netbooks are equipped with WLAN interfaces, 3G USB modems are quite cheap and operators offer flat rates to their customers). Besides, the number and popularity of applications designed for smartphones that make use of Internet connectivity is getting larger every day, contributing to an increase of market penetration of such devices (e.g., iPhone, Android, Blackberry and Windows Mobile phones), which results in growing demands for Internet connectivity everywhere.

Additionally, operators are migrating their networks to full IP based networks – for both voice and data – triggering a real need for IP mobility management solutions, which up to now had shown little or no deployment penetration. Most of the currently standardized IP mobility solutions, like Mobile IPv6 [1], or Proxy Mobile IPv6 [2] rely to a certain extent on a centralized mobility anchor entity. This centralized network node is in charge of both the control of the network entities involved in the mobility management (i.e., it is a central point for the control signalling), and the user data forwarding (i.e., it is also a central point for the user plane). This makes centralized mobility solutions prone to several problems and limitations, as identified in [3]: longer (sub-optimal) routing paths, scalability problems, signaling overhead (and most likely a longer associated handover latencies), more complex network deployment, higher vulnerability due to the existence of a potential single point of failure, and lack of granularity on the mobility management service (i.e., mobility is offered on a per-node basis, not being possible to define finer granularity policies, as for example on a per-application basis).

Because all the aforementioned issues, big operators are now looking for alternative mobility solutions that are more distributed in nature, allowing for a cheaper and more efficient network deployment capable to meet their customers’ requirements. In particular, there is an effort in the IETF, called Distributed Mobility Management, that is currently addressing exactly this particular problem, first starting from a clear definition of the problem statement [3].

There are basically two main approaches being researched now: one aimed at making Mobile IPv6 work in a distributed way, and another one doing the same exercise for Proxy Mobile IPv6. In this paper we present a complete solution for the Mobile IPv6 case, called Flat Access and Mobility Architecture (FAMA). While the general concept of IP distributed mobility management is not new, we argue that – to the best of the authors knowledge – this is the first complete description and evaluation of a concrete solution for client IP mobility.

The rest of the paper is organized as follows. In Section II we briefly summarize how centralized mobility management works, by describing the operation of Mobile IPv6, and highlight the main limitations of this kind of approach. Section III introduces our solution, which is then analyzed and compared with Mobile IPv6 – in terms of overhead, handover and communication delays – in Section IV. Before concluding the paper in Section VI, we compare our solution with other existing distributed approaches in Section V.
II. BACKGROUND AND MOTIVATION

A. Centralized Mobility Management: Mobile IPv6

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 MN, called Home Address (HoA). The HA (see Fig. 1) is in charge of defending the MN’s HoA when the MN is not at home, and redirecting received traffic to the MN’s current location. When away from its home network, the MN acquires a temporal IP address from the visited network - called Care-of Address (CoA) - and informs the HA about its current location by sending a Binding Update (BU) message. An IP bi-directional tunnel between the MN and the HA is then used to redirect traffic from and to the MN. There is also optional support to avoid this suboptimal routing and enable the MN to directly exchange traffic with its communication peers - called Correspondent Nodes (CNs) - without traversing the HA. This additional support is called Route Optimization (RO), and allows the MN to also inform a CN about its current location.

B. Limitations of centralized mobility management solutions

Centralized mobility solutions, such as Mobile IPv6, base their operation on the existence of a central entity that anchors the IP address used by the mobile node. This central anchor point is in charge of tracking the location of the mobile and redirecting its traffic towards its current topological location. While this way of addressing mobility management has been fully developed by the Mobile IP protocol family and its many extensions, there are also several limitations that have been identified [3]:

- Sub-optimal routing. Since the home address used by a mobile node is anchored at the home link, traffic always traverses the home agent, which leads to paths that are, in general, longer than the direct one between the mobile node and its communication peer. This is exacerbated with the current trend in which content providers push their data to the edge of the network, as close as possible to the users. With centralized mobility management approaches, user traffic will always need to go first to the home network and then to the actual content, adding unnecessary delay and wasting operator’s resources. With a distributed mobility architecture, as the anchors are located at the very edge of the network, close to the user terminal, data paths tend to be shorter.

- Scalability problems. With current mobility architectures, networks have to be dimensioned to support all the traffic traversing the central anchors. This poses several scalability and network design problems, as the central mobility anchors need to have enough processing and routing capabilities to be able to deal with all the mobile users’ traffic simultaneously. Besides, the operator’s network also needs to be dimensioned to be able to cope with all the users’ traffic. A distributed approach is inherently more scalable, as the mobility management tasks are distributed and shared among several network entities, which therefore do not need to be as powerful as the centralized alternative.

- Reliability. Centralized solutions share the problem of being more prone to reliability problems, as the central entity is a potential single point of failure.

- Lack of fine granularity on the mobility management service. With current centralized mobility management solutions, mobility support is offered at a user granularity. This means that the network can just decide if mobility is provided or not to the user, but cannot offer a finer granularity, for example, to allow part of his/her traffic not to be handled by the mobility solution. There are many scenarios in which part or all the traffic of a user does not really need to be mobility enabled, as for example when the user is not mobile (at least during the lifetime of the communication) or the application itself is able to effectively deal with the change of IP address caused by the user movement. In all these situations, it would be more efficient not to enable mobility.

- Signaling overhead. This is related to the previous limitation. Any mobility management solution involves certain amount of signaling load. By allowing mobility management to be dynamically enabled and disabled on a per application basis, some signaling can be saved, as well as the associated handover latency. Of course, this depends on the particular scenario, as the use of distributed mobility architectures can also lead to a higher signaling load in case of very dynamic scenarios in which all the traffic is required to be mobility enabled.

III. DESCRIPTION OF THE SOLUTION

Distributed Mobility Management approaches try to overcome the limitations of the traditional centralized mobility management, i.e., Mobile IP, by bringing the mobility anchor closer to the MN. Following this idea, in FAMA the MIPv6 centralized home agent is moved to the edge of the network.
Fig. 2. FAMA architecture and example scenarios

being deployed in the default gateway of the mobile node. That is, the first elements that provide IP connectivity to a set of MNs are also the mobility managers for those MNs. In the following we will call these access routers Distributed Anchor Routers (DARs).

Every time a mobile node attaches to a DAR, it gets an IPv6 address which is topologically anchored at the DAR. That means that while attached to this DAR, the mobile can send and receive traffic using that address without using any tunneling nor special packet handling. Every time the mobile node moves to a different DAR, it gets a new IPv6 address from the new access router. In case the MN wants to keep the reachability of the IPv6 address(es) it obtained from the previous DAR (note that this decision is dynamic and can be done on an application basis for example), the mobile has to involve its MIPv6 stack, by sending a Binding Update to the DAR where the IPv6 address is anchored, using the address obtained from the current DAR as care-of-address. In this way, the IPv6 address that the node wants to maintain plays the role of home address, and the DAR from where that address was configured plays the role of home agent (for that particular address). Note that FAMA architecture basically enables a mobile node to simultaneously handle several IPv6 addresses – each of them anchored at a different DAR – ensuring their continuous reachability by using Mobile IPv6 in a distributed fashion (i.e., each access router is a potential home agent for the address it delegates, if required). This distributed address anchoring is enabled on demand and on a per-address granularity, which means that depending on the user needs, it might be the case that all, some or none of the IPv6 addresses that a mobile node configures while moving within a FAMA domain, are kept reachable and used by the mobile.

In traditional Mobile IPv6, the communication between the MN and the HA is secured through IPsec [4]. Following a similar approach in FAMA is difficult due to the large number of security associations that would be required, since any gateway of the access network can play the role as home agent for any mobile node. In order to overcome this problem and provide authentication between the DAR and the MNs, we propose the use of Cryptographically Generated Addresses (CGAs), as introduced in [6]. Cryptographically Generated Addresses are basically IPv6 addresses for which the interface identifier is generated by computing a cryptographic one-way hash function from a public key and the IPv6 prefix. The binding between the public key and the address can be verified by re-computing the hash function and comparing the result with the interface identifier. To authenticate a message, the packet is signed with the corresponding private key, hence the receiver is able to authenticate the message with the knowledge of the address and the public key. CGAs are a powerful mechanism allowing packet authentication without requiring any public-key infrastructure, and hence it is well-suited for this application.

Following the ideas presented above, every time an MN attaches to a DAR, it configures a CGA from a prefix anchored at the DAR (e.g., by using stateless address auto-configuration mechanisms). This address can then be used by the MN to establish a communication with a remote Correspondent Node (CN) – see Fig. 2-(a) – while attached to that particular DAR. If the mobile then moves to a new DAR (nDAR), the following two cases are possible: i) there is no need for the address that was configured at the previous DAR (pDAR) to survive the movement: in this case there is no further action required; ii) the mobile wants to keep the reachability of the address configured at pDAR: in this case Mobile IPv6 is triggered, and the MN sends a Binding Update message to the pDAR, using the address configured at the previous DAR as home address, and the address configured at the new DAR as care-of address. This BU includes the CGA parameters and signature, which are used by the receiving DAR to identify the MN as the legitimate owner of the address. Although the use of CGAs does not impose a heavy burden in terms of performance, depending on the number of MNs handled by the DAR, the processing of the CGAs can be problematic. To reduce the complexity of the proposed solution, we suggest an alternative mechanism to authenticate any subsequent signaling packets exchanged between the MN and the DAR (in case the mobile performs a new attachment to a different DAR). This alternative method relies on the use of a Permanent Home Keygen Token (PHKT), which will be used

1There are additional parameters that are also used to build a CGA, in order to enhance privacy, recover from address collision and make brute-force attacks unfeasible.
to generate the Authorization node option that the MN has to include in all next Binding Update messages. This token is forwarded to the MN in the Binding Acknowledgment message, sent in reply to the BU. The procedure is depicted in Fig. 3. Once the signaling procedure is completed, a bi-directional tunnel is established between the mobile node and the DAR where the IPv6 address is anchored (the “home” DAR – HDAR – for that particular address), so the mobile can continue using the IPv6 address, as shown in Fig. 2-(b).

In case the MN performs any subsequent movements and it requires to maintain the reachability of an address for which it has already sent a BU, the following BU messages can be secured using the PHKT exchanged before, reducing the computational load at the receiving DAR.

Although this approach is attractive because it reduces the signaling overhead generated by the mobility support, it can be misused in some particular scenarios by malicious nodes that wish to export an incorrect CoA in the BU message, since it does not provide proof of the MN’s reachability at the visited network. Indeed, the CGA approach assures that the BU message has been sent by the legitimate HoA’s owner but it does not guarantee that the same MN is reachable at the provided CoA. In order to provide a more robust solution, we propose a Return Routability procedure similar to the one defined in MIPv6 Route Optimization to mitigate this security issue. The Return Routability procedure starts after the handoff. Instead of sending the BU message, the MN sends a Care-of Test Init message (CoTI). This message is replied by the DAR with a Care-of Test message containing a CoA Keygen Token. The MN can now send a BU using both Home and CoA Keygen tokens to prove its reachability at both the HoA and the CoA. The message and the knowledge of both tokens is a proof that the MN is the legitimate node who has sent the BU and also is reachable at the CoA indicated. As all security improvements, the one proposed incurs in a performance penalty, in this case an increase in the handover delay. Specifically this enhanced security approach requires four messages to be exchanged between the MN and the DAR instead of the two messages of the original solution. In terms of handover delay, it increases it by a factor of two, as the new solution requires two Round Trip Times (RTTs) to conclude, instead of one.

Note that on every attachment of a node to a DAR, the terminal also obtains a new IPv6 address which is topologically anchored at that DAR, and that this address can be used for new communications (avoiding in this way the tunneling required when using an address anchored at a different DAR), as shown in Fig. 2-(c). A mobile can keep multiple IPv6 addresses active and reachable at a given time, and that requires to send – every time the MN moves – a BU message to all the previous DARs that are anchoring the IP flows that the MN wish to maintain. For instance, in the example depicted in Fig. 2-(d), the MN sends a BU to the first DAR containing CGA-HoA as home address, while the BU it sends to the second DAR contains CGA-CoA1 as home address.

IV. ANALYSIS OF THE SOLUTION

In the following section we focus on conducting a simple analysis of the performance achievable by FAMA, comparing it with the one that would be obtained with plain Mobile IPv6.

The comparison is performed considering the three most important characteristics of a mobility protocol: i) the packet and signaling overhead, ii) the handover delay, and iii) the delay between both communication endpoints.

A. Overhead Analysis

As explained in Section III, our proposed Distributed Mobility Management solution – FAMA – is based on Mobile IPv6, pushing the home agent functionality to the edge of the network. Once an MN moves, the “home” DAR is in charge of tunneling the packets to the new MN’s location, identified by its care-of address. In terms of packet’s overhead, FAMA and Mobile IPv6 (without Route Optimization) share the same overhead since both use a bi-directional tunnel between the MN and an anchor point, hence both incur in a 40-byte overhead due to the packet encapsulation.

However, as compared to Mobile IPv6, in FAMA there may be more than one mobility anchor involved, and therefore this introduces a higher signaling load, since the number of BU/BA messages is increased. In particular, in plain Mobile IPv6 there is a single BU/BA exchange at each handoff – as only one home address is maintained by the MN – while in FAMA we have the BU/BA (and the CoTI/CoT in case additional security is required) exchange multiplied by the number of IPv6 addresses that need to be kept reachable.

B. Handover delay

The handover delay corresponds to the time during which an IPv6 address is not usable because of a change of the point of attachment. During this process there are multiple operations performed like the L2 attachment, the movement detection, the
clearly the main advantage of a distributed mobility management approach as compared with classical centralized mobility solutions, because the delay between the mobile node and its anchor is lower in the distributed approach as the anchor in this case resides at the edge of the network, instead of at the core of the operator. It is also worth noting how as the MN gets farther away from its HA/DAR, the handover delay increases, hence FAMA is better suited for flows with short duration or mobile nodes with low mobility. This characteristic is explored in more detail in the next section.

C. Communication delay

We next analyze of the delay experienced by packets exchanged between an MN and its communication peer (i.e., a CN). In Mobile IPv6, user data traffic always traverses the HA, although this path may not be the shortest one between the MN and the CN. This way of forwarding packets is known as triangular routing and is characterized by delays that might get to be large, since the packets must go through the MN’s home network, which can be located at a long distance from the MN. Due to the large delays introduced by the triangular routing, MIPv6 [1] already includes a procedure called Route Optimization that basically builds a secure direct path between the MN and the CN. Hence packets exchanged between MN and CN flow directly through the shortest path between the two nodes, without passing through the HA. This mechanism needs additional support from the CN, required to enable the route optimization of packets. In the case of FAMA, packets flow between the MN and the CN through the HA/DAR as in the case of Mobile IPv6 without RO. The difference between both approaches is that in the case of FAMA, DARs are expected to be located near the MN, hence the effect of triangular routing is highly minimized, obtaining delays of the order of RO-enabled Mobile IPv6. In the previous section, it was mentioned that the use of FAMA is better suited for flows with short duration or low mobility MNs. This is due to the fact that as the MN moves away from the HA/DAR handling a flow, the inefficiency introduced by the triangular routing increases.

In order to assess how far and how fast an MN can move, we perform the following analysis. Let’s suppose a VoIP communication between two peers, being one of them an MN using FAMA to handle its mobility. Considering the maximum mouth-to-ear delay as specified in [7] of 150 ms, we can assume that Eq. (2) holds:

\[
T_{CN\rightarrow HA/DAR} + T_{HA/DAR\rightarrow MN} \leq 150\text{ms}. \tag{2}
\]

Let’s assume the CN and MN to be in the same geographical region or even city. In order to model this delay, we took average values from the PingER (Ping end-to-end reporting) project\(^2\), between several client-server pairs located in the same regional area. The average delay obtained corresponds to roughly 20 ms, hence from Eq. (2) the delay between the HA/DAR and the MN is upper bounded by 130ms. Assuming that the FAMA domain has a good internal connectivity and is all managed by the same provider, we can conclude that the delay between two DARs is similar to a local delay between two servers located in the same organization from the PingER project (which is on average equal to 5 ms). To simplify, we suppose the access network is deployed in such a way that going farther away from the original DAR increases the delay in a linear way (note that this is a worst case scenario). The maximum number of hops allowed for the VoIP communication can then be derived from Eq. (2), resulting in a maximum distance of 26 hops. This number represents a limit.

\(^2\)http://www-iepm.slac.stanford.edu/pinger/
on the diameter of the FAMA domain, which depends on the access technology used. In the case of a WAN technology such as WiMAX or 3G, one access router can serve a cell of about 50 Km of radius, while in the case of a LAN technology such as IEEE 802.11, the cell radius is reduced to less than 100m. Now let’s look at a typical use case, where a user starts a VoIP conversation and walks across a FAMA domain using IEEE 802.11. The typical speed for pedestrians is 4-5 Km/h [8] and the average call duration is roughly 3 minutes [9]. This means that during the call, the user will walk around 250m, hence performing two handovers and adding a delay of roughly 10ms more than the direct path between the CN and MN. This simple example shows two of the benefits of FAMA: simplicity and low added end-to-end communications delay.

V. COMPARISON WITH PREVIOUS WORK

The design of flat mobile architectures is becoming a quite hot topic in the IETF and 3GPP, with several solutions already proposed. We next compare FAMA with some of them.

According to [10] and its implementation [11], mobility support is provided on demand, that is, only for those MNs that change access point with ongoing connections. The MN configures and maintains an IP address for each access network it visits, and the access router in that access network is the anchor for the communications established using the IP address assigned by the router. This means that an access router acts as a standard router when the MN is attached to it, otherwise it tunnels the packets to the access router where the MN is currently attached. Respect to FAMA, this solution reduces the caches at the mobility agents and it does not require the MN to implement any mobility client, except for the source selection mechanism. In both schemes the IP flows are anchored to the access router that advertised the prefix used in the communication so they both offer the same path delay. As a difference, FAMA uses the address configuration described in [6], while these papers leave an open issue regarding the source address selection: [12] proposes a similar approach with some changes in the Linux source selection algorithm in order to achieve the expected behavior. Moreover, FAMA replicates the signaling, data structures and message format designed in [1] and [13], while [12] is supposed to use the Proxy Mobile IPv6 scheme. However in [10] no exhaustive explanation or suggestion in provided, for instance about how the access routers set up the tunnel between them.

A description on how to distribute Proxy Mobile IPv6 is given in [14], in which small Proxy Mobile IPv6 domains form the whole mobility domain. This draft provides a solution to achieve route optimization in several scenarios, at the cost of excessive control messages exchange. Also, the architecture deployment requires a big effort since every small Proxy Mobile IPv6 domain is made of the complete equipment.

Both [15] and [16] propose to use a Distributed Hash Table (DHT) to store the mobility information of the MNs (in MIPv6 this table is called Binding Cache). The former focuses on how to efficiently manage the DHT and other related aspects providing a simulated evaluation, while the latter focuses on a technique to perform handover using unicasting. Both approaches suffer from requiring a lot of new support and not reusing existing legacy standards and solutions.

VI. SUMMARY AND FUTURE WORK

In this paper we have proposed and analyzed FAMA, a distributed mobility management solution based on Mobile IPv6. The solution brings the advantages of a distributed solution, namely shorter data paths, better scalability and reliability, lower signaling overhead and shorter handover latencies, and better control on the mobility granularity offered by the network. Two different levels of security protection are proposed, not requiring the use of IPsec, and therefore allowing for a faster and easier deployment.

Compared with a pure centralized mobility solution, such as Mobile IPv6, FAMA exhibits a better performance in terms of handover delay and proves to be a better solution for mobile nodes with low mobility patterns.

Future work includes the experimental evaluation of FAMA and its comparison with Mobile IPv6 in a real testbed. Other ongoing work consists of the design and evaluation of a distributed mobility solution based on Proxy Mobile IPv6 that fully considers the address management issues on the terminal.

REFERENCES