SDN-DMM for Intelligent Mobility Management in Heterogeneous Mobile IP Networks

Rodrygo Torres Córdova[(1)], Paulo Roberto de Lira Gondim[(1)], Yarisley P. Llerena[(1)], Jaime Lloret[(2)]

(1) Eletrical Engineering Department – Universidade de Brasília – Brasília - Brazil

rodrygo.cordova@rnp.br, pgondim@unb.br, yarisley@aluno.unb.br

(2) Integrated Management Coastal Research Institute, Universitat Politecnica de Valencia – Valencia – Spain.
jlloret@dcom.upv.es

ABSTRACT

Mobility management applied to the traditional architecture of the Internet has become a great challenge due to the exponential growth in the number of devices that can connect to the network. This article proposes a network architecture based on Software Defined Networking (SDN), which not only provides its intrinsic benefits, but also deals with the distributed mode of mobility management in heterogeneous access networks in a simplified and efficient way, ensuring mainly the continuity of IP session as the user shifts between networks and efficient use of the communication network infrastructure. A comparison with two other recent proposals show that the proposed architecture has a better performance.

Keywords: SDN, DMM, software defined network, distributed mobility management

I. INTRODUCTION

Mobile communication networks have become the main users method of access to the Internet, which has significantly increased the number of mobile devices connected to the global network [1].

As the services offered by operators of mobile communication networks tend towards involving solutions totally based on the IP protocol for both voice and data services [2][3] and the communication sessions must be continued, the IP mobility management has become fundamental for communication networks, so that such an increase can be supported [4].

The number of mobile devices connected to the network has increased exponentially, due to the expansion in both availability and use of applications that require mobility, especially real-time collaborative tools [5], and the current mobility management solutions are not adequate for satisfactorily meeting the requirements imposed on the infrastructures of communication networks.

The IETF IP mobility management standards, as Mobile IPv6 (MIPv6) [6] and Proxy Mobile IP Version 6 (PMIPv6) [7] depend on central units that manage both control and data traffic, elaborated according to the traditional routing of IP packets. They also pose problems, as sub-optimized routing, low scalability, overload of processing in network devices and low granularity in the mobility management service.

Additionally, the use of heterogeneous access networks (HetNets) imposes other difficulties to such a management, as the integrated management of resources and the soft transitions among networks.

The mobility management faces a constant challenge regarding the communication network efficiency with no increase in its complexity. The addition and use of new protocols, signaling messages or processes that cause overhead due to the encapsulation and control traffic are examples of how the mobility management directly impacts on the OPEX (Operational Expenditure) and CAPEX (Capital Expenditure) of a communication network.

An alternative for dealing with the intrinsic problems of centralization and costs in the mobility management is the new concept, called distributed mobility management (DMM) [8] [9]. Its main characteristic is the clear separation between the data plane and the control plane. The data plane is distributed along the equipments on the access network edge for approximating the mobile agent to the end user, implementing a flatter mobility approach in the network [10]. Therefore, the traffic forwarded to the mobile node (MN) does not cross a specific central point in the network, i.e., the mobility anchor, as it occurs in centralized solutions. In DMM, the traffic is forwarded by the mobile agents nearest the user.

The DMM solutions can be categorized into two types, depending on the distribution level of the control plane [1,11,12]:

- Partially distributed: the data plane is completely distributed among the network elements and the control plane is centralized in control points in the network; and
- Fully distributed: both data plane and control plane are completely distributed among the network elements and there exists no central entity of control.

On the other hand, Software Defined Networking (SDN) represents an emerging paradigm, introducing the programmability of the network from a centralized location, and decoupling the network's control logic from the underlying routers and switches, thus improving the ability to manage the network state. In SDN, OpenFlow protocol is a multivendor standard defined by the Open Networking Foundation, for working on top of TCP protocol and allowing an OpenFlow Controller to stablish instructions (structured as flows) to an OpenFlow switch.

This manuscript proposes an architecture that uses the SDN paradigm with the OpenFlow protocol for the implementation of a network-based DMM solution in an environment of heterogeneous networks for dealing with the above-mentioned challenges of IP mobility management.

The remainder of the manuscript is organized as follows: Section II addresses the related work on network-based DMM; Section III describes our proposal for DMM, called SDN-DMM (Software-Defined Network for Distributed Mobility Management); Section IV reports the analytical evaluation of the proposals; Section V provides the results and discussions. Finally, Section VI presents our conclusions, reports on the ongoing studies and suggests future work.

II. RELATED WORK

The literature reports several strategies for mobility management in heterogeneous cellular networks. This section addresses a discussion about the main strategies, focusing on partially and fully distributed solutions for mobility management of IP networks based.

The network-based centralized mobility management comprehends two entities in the network core, namely LMA (Local Mobility Anchor) and MAG (Mobile Access Gateway), used in PMIPv6 [7] (Proxy Mobile IP Version 6) protocol, the most important proposal and the network-based counterpart of MIPv6 protocol.

In PMIPv6, the network is in charge of any update regarding the MN location. The LMA works as a centralized entity which maintains the MN tracking by employing a binding record, and the MAG represents the first-hop router in the domain.

Initially, the mobility management was centralized, causing problems of single point of failure, network congestion by centralized processing, delays due to geographic location (centralized entity far from the MN) and scalability issues, among others. Thus, distributed solutions arose, with performance improvements snared by the breakdown of the features of centralized entities of the CMM (Centralized Mobility Management) in different entities distributed in DMM. As a negative result, an increase in the control signaling and a more complex architecture in terms of entities were observed.

In distributed network-based proposals, all the mobility management is performed on the network side, and the functions of location and encapsulation of data are distributed in entities at network borders. Such a procedure avoids single points of failure and attack and improves scalability and reliability in the network.

Thus, this section contains a description of the two groups of proposals for DMM, namely fully distributed and partially distributed. For each group, and for the sake of comparison with SDN-DMM, specific proposals are chosen to be presented in more detail: the proposals by J. Lee and Y. Kim [13] and by Bernardos, C.; Oliva, A.; Giust, F. [14], respectively partially and fully distributed.

i. Fully Distributed Solutions for DMM

a) The “Draft-Jaewoon” Proposal, by J. Lee and Y. Kim

A network-based fully distributed management solution (“Draft-Jaehwoon”) is presented in [13], in which all control and data plane functions are distributed among the mobile access gateways (MAGs). No central unit, as the local mobility anchor (LMA), is responsible for the processes of signaling and forwarding of traffic. Such responsibilities are distributed among the entities located on the network edges.

The protocol assigns a network prefix, called PREF (in a supernet concept) to the mobility domain, where a different subnetwork that belongs to PREF is allocated for each MAG of the domain. Although each MAG is responsible for a specific subnetwork, all MAGs inform the PREF prefix as the network to which they belong through its mobility service interfaces.

When the mobile node enters the domain for the first time, it is provided with an IP address of the subnet corresponding to the MAG it is connecting to and PREF prefix is indicated as the network address. Once the mobile node change its connection point to a new MAG and receive new router advertisement (RA) messages with the option information field containing the same network prefix PREF, it considers it is still connected to the same network and keeps its addressing according to what has been established in the connection with the first MAG.

The original MAG from the mobile node, i.e., the MAG responsible for the IP address used by the mobile node that performed a handover, named anchor-MAG (A-MAG) from now on for future references, acts as an LMA, performing functions such as creation and maintenance of the binding cache entry (BCE), buffer and packet redirection through the tunnels established.

The signaling process starts when an MAG has identified the movement of a mobile node within the domain. In the association between the mobile node and MAG, the latter sends an RA message indicating PREF prefix as the network to which the mobile node will connect. If MAG receives a DHCP request from the mobile node, it considers the mobile node is entering the domain for the first time, i.e., it is not a handoff.

However, if MAG receives any other packet with an source IP address of PREF prefix that does not belong to the scope of its subnetwork, it detects the mobile node is performing a handoff and starts the signaling process. This MAG is called, for future references, as serving-MAG (S-MAG), which is the MAG that provides connectivity to a mobile node that has performed a handover.

S-MAG sends a distributed proxy binding update (DPBU) message to A-MAG to inform the new location of the mobile node and creates a binding update list (BUL) for the establishment of a tunnel between them. When A-MAG receives the DPBU from S-MAG, it creates a BCE entry for the mobile node and responds with a distributed proxy binding acknowledgement (DPBA) message for the establishment of the bidirectional tunnel to be used for the redirection of all traffic of the mobile node.

As A-MAG cannot know when a mobile node will perform another handover, S-MAG sends a distributed proxy binding release update (DPBRU) message to A-MAG informing on the disconnection of the mobile node when it notices the mobile node has been or will be disconnected. After receiving the DPBRU, A-MAG starts to store the packets addressed to the mobile node and responds with a Flush message to S-MAG, which cleans the mobile node registrations from its BUL and retransmits the Flush to A-MAG, which cleans the current registrations of the BCE and waits for a new DPBU from the new S-MAG for the establishment of the new tunnel.

b) Other Fully Distributed Solutions

F. Giust, C., J. Bernardos and A. De La Oliva [11] described a totally distributed solution based on PMIPv6, using the services of the Media Independent Handover specification (IEEE 802.21). The authors proposed a mechanism for multicasting the PBU (Proxy Binding Update) message sent by a S-DAR (Serving-Distributed Anchor Router), in the group formed by all DARs in a domain. The same authors also designed a totally distributed solution [14, 20] in which MAARs (Mobility Anchor and Access Router) deal with both data and control plans. They proposed a P2P architecture or Unicast, Multicast and Broadcast queries for solving the MAARs-discovery problem in the process of distributed control plan update.

Bernardos, Carlos J., Juan Carlos Zuniga, and Alex Reznik [21] also proposed a network-based Fully-DMM. In the proposal, new network entities called D-GW (Distributed Gateways) are distributed at the edge of the network, near the UEs (User Equipment) in the access network. The D-GW interact directly with the UEs and, therefore, they develop the access and routing functionalities. It has the advantage that for PDN (Packet Data Network) connections in the HPLDN (Home Public Land Mobile Network) no extra functionality is used, being transparent to the UE and the rest of the entities of the network. The D-GW also replaces the ePDG (Evolved Packet Data Gateway) providing the functionality of IpSec tunneling for the UE in access untrusted non-3GPP. In 3GPP access cases it works as a simple relay between eNB (EnodeB) and SGW (Serving Gateway). The similarity of D-GW with PGW makes it possible to reuse the implementation of the software stack. The D-GWs behave as AR and mobility-signaling agents and include the LMA functionalities for PMIPv6-based networks.

Finally, Carmona-Murillo, Javier et al. [22] proposed a fully distributed architecture designed to track the mobility of users in MPLS-based network called DM3 (distributed mobility management MPLS). The authors assume that an MPLS domain exists in the access network between the ingress LER/egress LER (Label Edge Router). Both ingress LER (ILER) and egress LER (ELER) are the border MPLS routers that define the limits of the access network. DM3 architecture relies on the distributed mobility agent called the mobility distributed anchor (MDA). This node provides mobility management functions and is an intermediate node between the ILER and the serving ELER. The distributed mobility anchor agent (MDA) is responsible for the LSP (Label Switching Path) redirection when the MN moves to an adjacent network. This way, the LSP will be composed of a set of forwarding paths that adapt to track host mobility and localize signaling in an area close to the location of the MN. The proposed architecture is based on MIPv6 protocol. It has the main disadvantage that the MN needs to be changed, given its active participation in the mobility management.

ii. Partially Distributed Solutions for DMM

In P-DMM (Partially - DMM) the location and handover functions (control plane) are separated from the data traffic routing (data plan). The data plan is distributed and the control plan is centralized. The main goal is the routing optimization.

a) The "Draft-Bernardos" proposal, by Bernardos, C.; Oliva, A. and Giust, F.

A network-based partially distributed DMM solution ("Draft-Bernardos") is presented in [14]. It is based on the PMIPv6 protocol, in which a central unit, called central mobility database (CMD) performs the main functions of control plane, whereas the forwarding of traffic, data plane, is the responsibility of units called mobility anchor and access router (MAAR).

CMD assumes the LMA figure, however, it does not forward traffic to the mobile node, i.e., it does not participate of the data plane operations. It is the central entity of the control plane and manages sessions and registration of IPs, location and binding cache. MAG is replaced by network entity MAAR, which performs LMA functions, as treatment of aspects of control plane, as local binding cache, and aspects of the data plane, as supply of connectivity to the mobile node.

Each MAAR manages a set of nonsuperposed specific IP prefixes assigned to the mobile node when it is connected to MAAR. At each new connection of the mobile node to a new MAAR, a new IP address is allocated in the mobile node, under the responsibility of MAAR. The home network prefix (HNP) address concept of PMIPv6, in which only one IP address is allocated to the mobile node for all mobility domain, is changed to the address concept called local network prefix (LNP).

When the mobile node connects to MAAR for the first time, MAAR creates a unique identification for the mobile node within the mobility domain (MN-ID) and stores an exclusive LNP address to be assigned to the mobile node. MAAR inserts such information in its local binding cache entry (BCE) table and sends it through a standard proxy binding update (PBU) message to CMD. The latter includes the information in its BCE table, whose scope is global for the mobility domain, creates a mobility session for the mobile node and responds to MAAR with a standard proxy binding acknowledgement (PBA) message. After MAAR has received the PBA, it sends an RA message that includes the LNP previously stored for the mobile node, which starts to use the IP address and the current MAAR is referenced as Serving-MAAR (S-MAAR), i.e., the MAAR that is providing connection to the mobile node.

When the mobile node performs a handover for the new MAAR, the latter follows the same procedures described and allocates a new NLP for the mobile node. However, when it sends the information to the CMD, which verifies the mobile node was previously connected to another MAAR using an NLP address allocated by MAAR, called, respectively, anchor-MAAR (A-MAAR) and previous-LNP (pLNP), CMD updates the entries in its BCE table and sends messages to A-MAAR containing the new location of the mobile node, and to the current MAAR, i.e., S-MAAR, about the pLNP used and the address of the responsible A-MAAR. At this moment, S-MAAR and A-MAAR establish an IP-in-IP tunnel for the redirectioning of the traffic related to pLNP, which guarantees the continuity of the IP sessions that were in progress when the mobile node performed the handover for a new MAAR.

Therefore, the mobile node has two IP addresses, i.e., the new address provided by S-MAAR, the LNP, and the previous address provided by A-MAAR, i.e., the pLNP. All traffic related to pLNP is sent through the tunnel between the MAARs and all traffic related to LNP is treated directly by S-MAAR, without the use of encapsulation techniques, through which new IP sessions are established with such address.

b) Other Partially Distributed Solutions

Yi, Li et al. [23] proposed a DMM solution named D-PMIPv6. This solution includes separate data plane and control plane and divides the LMA into two entities, namely CLMA (Control plane Local Mobility Anchor) and DLMA (Data plane Local Mobility Anchor). CLMA manages signaling messages, whereas DLMA establishes the tunnels with the MAGs (Mobile Access Gateway) for the forwarding of the data packets. X. Keqiang et al. [24] added IP flow mobility to this solution by establishing a routing policy to support differentiated services through the marked packets. Both CLMA and DMLA entities are centralized in each plane and must manage all network nodes, thus representing relevant points of possible failure.

Another partially distributed solution was proposed by Luo, W., and J. Liu. [25], named ePMIP (enhanced Proxy Mobile IPv6) solution for supporting DMM. Two logical functions, namely Location Management Function (LMF) and Distributed Anchoring Function (DAF) were introduced. LMF maintains the mapping between the IP address and the in- formation on the MNs location. It can be implemented in the LMA, which constitutes eLMA (evolved LMA). DAF includes the Distributed Routing sub- Function (DRF), which enables the route optimization between MN and CN (Correspondent Node), and the Distributed Mobility sub-Function (DMF), which guarantees the MN mobility when the route optimization has been activated. DAF can be implemented in the PMIPv6 MAG and constitutes an eMAG (evolved MAG). The draft considers mainly route optimization and inter-eLMA communications.

On the other hand, P. Seite, P. Bertin, and J. Lee [26] described a solution in which the tunnels between MARs (Mobility capable Access Routers) are used only for ongoing sessions initiated prior to the handover. The data packets of the new sessions opened at serving MAR will be directly routed. For the optimization of the routes, each previous MAR that still has the ongoing sessions of the MN establishes a tunnel with the serving MAR to maintain the continuity of the sessions. The control plane is centered, however, the interaction between MAR and the database is not specified in the draft.

D. R. Purohith et al. [27] proposed another partially distributed solution considering an architecture, called Seamless Internetwork Flow Mobility (SIFM), which uses the concepts of PMIPv6 and SDN (Software Defined Net- working) and defines a Flow Controller (FC) similar to the OpenFlow controller. PGW on LTE networks and Wireless Access Gateway (WAG) on WiFi networks act as OpenFlowhybrid switches that perform mobility signaling on behalf of the UE. They follow the instructions of the FC when the MN moves from an LTE network to a WiFi in order to provide seamless transition. The FC makes the routing decisions, instructs the switch on how to forward similar packets and performs functions related only to mobility.

The table 1 summarizes the following characteristics of the work discussed in this Section: (i) whether mobility management is partially distributed or fully distributed, (ii) the network entities that make up the architecture, (iii) new messages defined in the architecture, and (iv) the access technologies considered in the solution.

Table 1 – Comparison among Proposals for DMM

Proposal	**Mobility Management**		**Network Entities**	**New Messages**	**Access Technologies**
	F-DMM	P-DMM			
J. Lee and Y. Kim [13]	X		MAGs	*Distributed Proxy Binding Update (DPBU)* *Distributed Proxy Binding Acknowledgement (DPBA)*	Generic
F. Giust et al. [11]	X		DAR	*MIH messages*	Generic
C. J. Bernardos et al. [21]	X		D-GW	*Update D-GW address*	Cellular (LTE) WiFi
Carmona-Murillo, Javier, et al. [22]	X		MDA ILER ELER	*Handover Notification*	Cellular (Generic)
C. J. Bernardos et al. [14]		X	CMD MAAR	*PBU/PBA extended for route update*	Generic
L. Yi et al. [23]		X	DLMA, CLMA, MAG	*Proxy Binding Query (PBQ) Proxy Query ACK (PQA)*	Generic
K. Xie et al. [24]		X	DLMA, CLMA, MAG	*Flow Move Update* (FMU) *RS message is extended containing information about the flows*	Cellular WLAN
W. Luo and J. Liu [25]		X	eLMA eMAG	*PMIP Binding Query Request, PBQR* *PMIP Binding Query Answer, PBQA* *PMIP Binding Change Inform, PBCI* *PMIP Binding Change Ack, PBCA*	Generic
P. Seite et al. [26]		X	MAR	*(messages aren't altered)*	Generic
D. R. Purohith et al. [27]		X	FC MA	*Flow Modification Message* *Port Status Update*	Cellular (LTE) WLAN
SDN-DMM		X	OpenFlow switch SDN Controller	*OpenFlow messages*	3G LTE WiFi

III. SDN-DMM: AN ARCHITECTURE BASED ON SDN FOR THE MANAGEMENT OF DISTRIBUTED MOBILITY

Traditional IP mobility solutions, as those standardized by IETF, are somehow supported on the traditional routing of IP packets and specific network devices for playing key roles in the mobility and treatment of communication processes, as registration processes and techniques of packet encapsulation.

Such solutions cause overhead in the infrastructure, regarding the network devices involved and the data transportation, which increases the complexity for the operation and maintenance of the network. These network devices must perform not only their main function, e.g., packet routing in routers, but also extra functions, as tunneling, proxy actions, packet analysis, among others.

The use of techniques of data encapsulation and sub-optimized routing based on the address of the IP packet is inappropriate for the mobility scenario and causes the inefficient use of the communication network infrastructure, reduction in the effective rate of data transmitted and overload of certain communication links, whereas others are underutilized. Furthermore, such negative points are potentialized by the expected increase in the number of mobile devices connected to the network (due, for example, to the Internet of Things - IoT).

This section proposes a network-based partially Distributed Mobility Management solution based on the SDN paradigm for dealing with the challenges addressed. It uses the OpenFlow protocol and ONOS controller to provide IP mobility management in heterogeneous access networks focusing on the continuity of IPs sessions.

The proposal aims at a reduction in the general complexity of the system and the operational (OPEX) and investment (CAPEX) costs through the separation between the control and data planes of the communication and the treatment of packets according to the IP flows. The following aspects must be satisfied:

- Continuity of the IP session;
- Optimized routing;
- Distributed architecture;
- Transparency for the mobile node;
- Scalability;
- Performance;
- Low Complexity.

Figure 1 shows a basic reference SDN architecture composed of an ONOS (C1) controller, which can be implemented in a distributed hierarchical cluster mode for the elimination of single failure points and treatment of the scalability aspect of the control plane, OpenFlow switches, a media service in the Internet (MD), a transport backbone network and heterogeneous access networks.

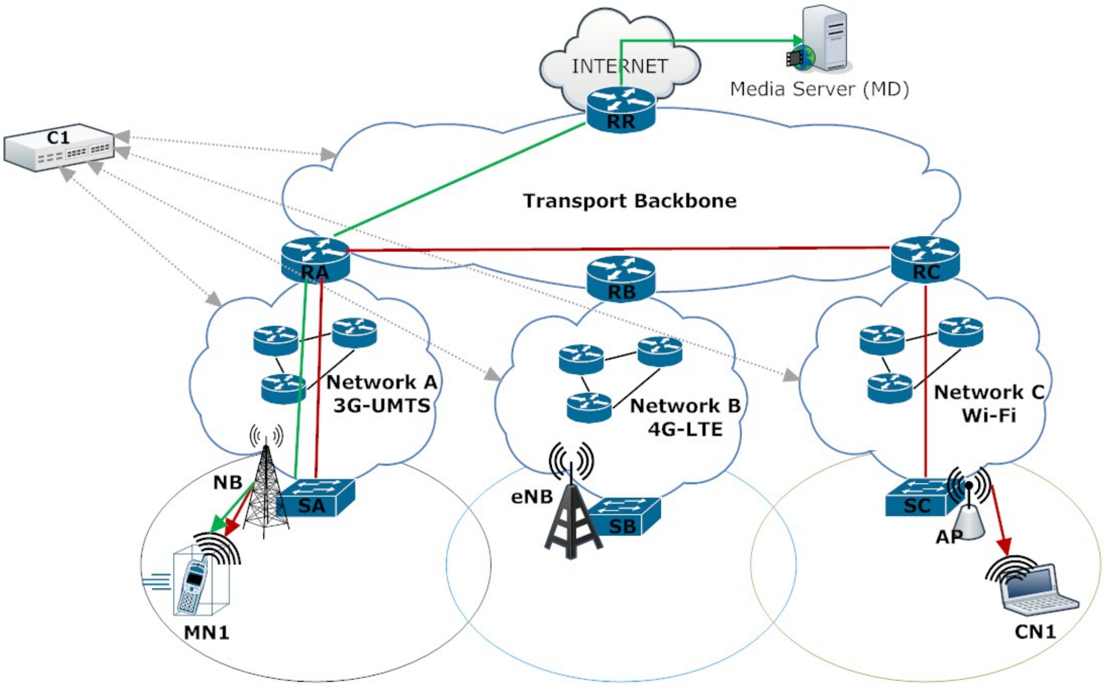

Figure 1. SDN-DMM architecture.

The general scenario analyzed involves the mobility performed by mobile node MN1, which moves from access network A (3G-UMTS) to access network B (4G-LTE) and then to access network C (WiFi), while keeping its communications with the Media Server and correspondent node CN1. The specific situations analyzed in this mobility process are shown in Figure 2, where:

- s_1 = mobile node MN1, media Server MD and corresponding node CN1 are connected to their origin networks and establish communication normally through the network infrastructure (no mobility has occurred);
- s_2 = as the communications established in the previous situation are in progress, mobile node MN1 changes its connection point from network A (3G-UMTS) to network B (4G-LTE), which causes the mobility process;
- s_3 = mobile node MN1 performs another handover, changing its connection from network B (4G-LTE) to network C (WiFi), while the sessions are still in progress.

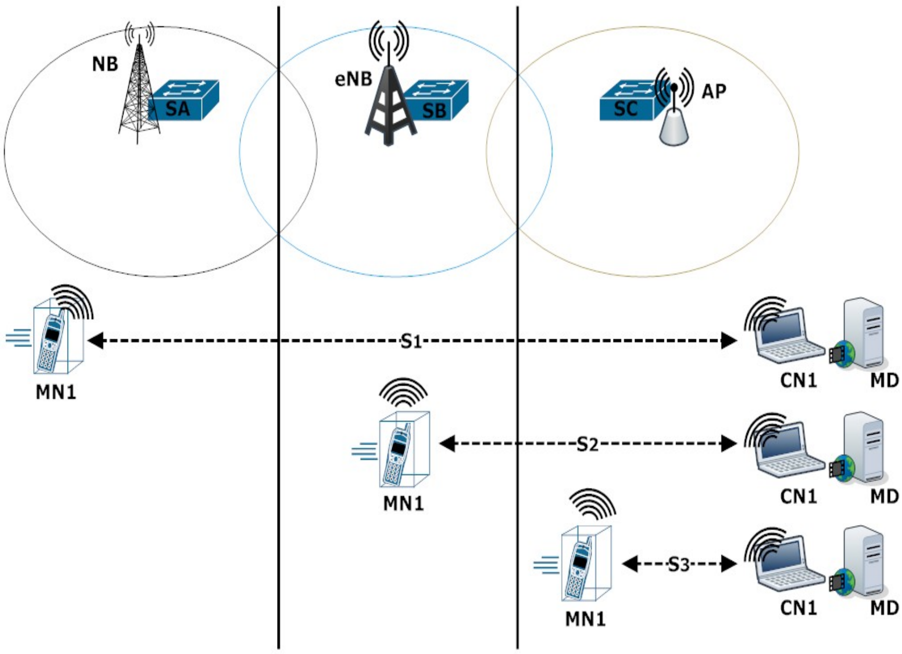

Figure 02. Connection scenarios.

The end-to-end communication between hosts in scenario s_1 is established in the network through bidirectional IP flows, i.e., a flow from MN1 to CN1 and another from CN1 to MN1 (represented in Figure 1 by the red flow) and a flow from MN1 to MD and another from MD to MN1 (represented in Figure 1 by the green flow). Such flows are established by OpenFlow rules that insert corresponding entries in the forwarding table of the network devices involved in the formation of the flow path. The process of establishment of the bidirectional IP flows through OpenFlow rules is shown in Figure 3.

In such a scenario, switch SA, which has received an IP flow from mobile node MN1 to be forwarded to correspondent node CN1 and media server MD, identifies corresponding entries to such flows in its forwarding table and takes the appropriate actions. It forwards the packets that belong to both flows to router RA, which analyzes its forwarding table and sends the flow addressed to MD to router RR and the flow addressed to CN1 to router RC.

In scenario s_2, when mobile node MN1 perfoms the handover process from NodeB (NB), inserted in the context of network A, to eNodeB (eNB), inserted in the context of network B, the mobility process is started.

For the continuity of the IP sessions, mobile node MN1 keeps its IP address during the handover process from network A to network B for avoiding changes in the end-to-end communication, mainly in relation to layer 3 and higher layers, as the replacement of TCP/IP sockets. Due to its new location and use of an IP address different from the addressing scope of the network assigned to its new connection point, i.e., network B, the packets addressed to the MN1 IP address must be correctly forwarded by the network infrastructure to its new position.

For a transparent process for the final hosts, the communication network must self-adapt, i.e., provide a new communication channel that enables the end-to-end communication among the hosts for a new scenario. However, it must not affect the other ongoing communications in the network and their routing.

For a more efficient use of the available infrastructure and avoidance of overloads due to techniques of registration or packet encapsulation, the solution uses a modified abstraction of the Northbound layer, called mobility intent. It readjusts the data plane of the network, so that the infrastructure can forward the packets addressed to the mobile node based on the bidirectional IP flow, rather than on the traditional packet routing by the destination IP address.

An intent can be considered a communication intention performed at a high level; it informs the controller about which network points should be established a communication path [15].

After controller C1 has been known on both intention (mobility intent) and physical topology, it chooses the best way for redefining the communication paths by sending low level control rules through the Southbound interface (OpenFlow rules) to the network equipments involved and readjusts its forwarding tables for new bidirectional IP flows for the mobile node in its new location.

In the solution when switch SB has identified the mobile node or an IP flow whose original address is different from network B, it informs such an event to controller C1, which, after having detected a change in the topology involving the location of mobile node MN1, readjusts the data plane of the communication network changing the OpenFlow rules for inserting new entries in the forwarding tables of the network equipments involved in the changes. Then a new set of links for the creation of a new bidirectional IP flow between the hosts is reestablished. Figure 3 shows the main stages of the mobility management process, where:

1. Mobile node MN1 performs handover from NB to eNB, keeping its original IP address;
2. Switch SB identifies the presence of the mobile node with a IP network address different from its scope and informs the event to controller C1;
3. Controller C1 detects a topology change in the mobility of mobile node MN1, therefore, it recalculates a new path and sends new OpenFlow rules for readjusting the communication links used;
4. The new bidirectional IP flows are established in the network, which enables the packets addressed to mobile node MN1 to be correctly forwarded according to the IP flow, rather than to the traditional IP routing. Therefore, the general routing and the other ongoing communications in the network are not affected.

In this scenario, the traffic between hosts is routed according to the new bidirectional IP flow established by the change in the data plane performed by the controller. The packets sent to mobile node MN1 with an address of network A, can be directly delivered to MN1 even when it is connected to network B, in an optimized and transparent way and with no overhead addition for the data transportation in the network.

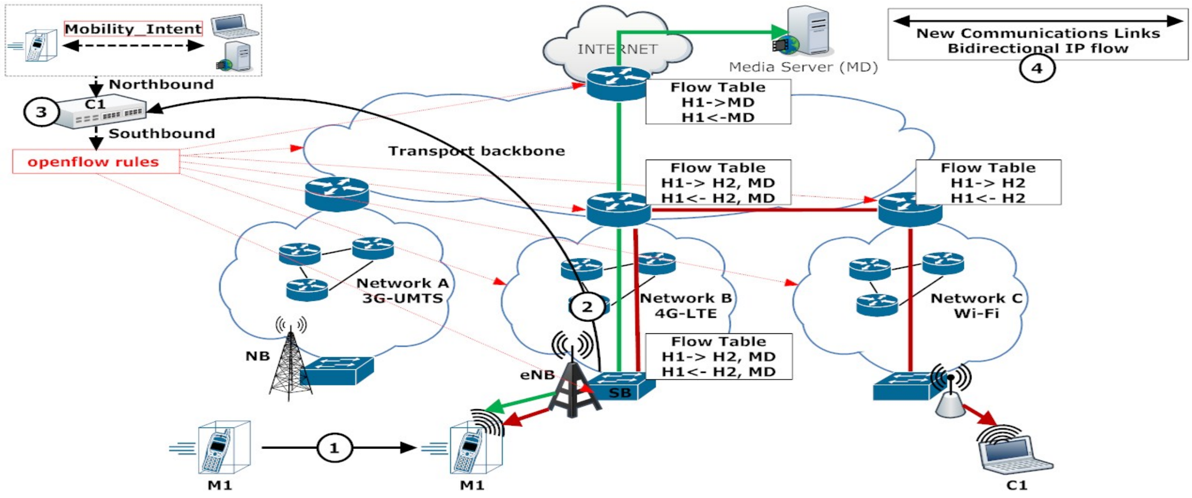

Figure 3. Handoff and establishment of the bidirectional IP flow in s_2.

The above-mentioned process is repeated when mobile node MN1 performs another handover by connecting to the access point AP (Fig. 3), inserted in the context of network C. Switch SC informs controller C1 on the new change in topology and controller C1 reestablishes the bidirectional IP flow between the hosts. The communication between mobile node MN1 and correspondent node CN1 now occurs directly in network C.

In SDN-DMM, scalability for the data plane is obtained as a central node is not dependent or overloaded for the mobility of the mobile node. The network equipment on the edge of the access network only forward the traffic to the mobile nodes connected to it and do not redirect the traffic to the mobile nodes connected to other network equipments.

Regarding the control plane, the controller is the central point of the network intelligence responsible for the mobility of the domain. It can be implemented hierarchically in cluster mode and a certain controller becomes responsible for part of the infrastructure, whereas the others act as backup and responsible for other parts of the network, which enables issues on scalability and the unique failure point to be addressed for the control plane.

As the mobile nodes are not involved and do not conduct any other additional process regarding mobility and the network infrastructure does not need to support any other new protocol, the benefit of the proposal regards the maintenance of the same level of complexity for both the network edge and final devices and the network core, according to the SDN-based architecture. The processing and energetic consumption in the mobile node are not affected and the network equipments do not need to create or maintain specific tables and processes, as encapsulation and binding relations to provide mobility.

IV. ANALYTICAL EVALUATION

This section addresses the analytical evaluation of the draft-Jaehwoon [13], draft-Bernardos [14], and SDN-DMM proposals conducted with metrics of signaling and handover latency costs. A comparative table of the aspects of routing, performance, requirements for the end hosts and architecture/complexity is also provided.

i. Signaling cost

The signaling cost is related to the number of control messages generated specifically for supporting the mobility process and keeping the continuity of the mobile node sessions. Its three main components are ([16]):

i. C_{update}= Cost of the location update of the MN with binding update messages;
ii. $C_{refresh}$ = Cost of periodical updates;
iii. C_{de-reg}= Cost of deregistration due to the execution of a new mobility within the domain.

The total signaling cost C_{SIG}, provided in [15], is given by:

$$C_{SIG} = \mu_{SN}\left(C_{update} + C_{refresh} + C_{de-reg}\right) \quad (1)$$

where μ_{SN} represents the subnet crossing rate performed by the mobile node. The individual cost of each of the three components is given by the sum of C_{X-Y} cost, which is the symetrical cost for the transmission of a control packet between two, X and Y, network units, and PC_Xcost, which is the cost involved in the processing of the control packet by network unit X [11].

The signaling costs for each proposal are evaluated below:

- Draft-Jaehwoon

When a mobile node performs a handover, the new S-MAG sends a DPBU message to A-MAG, which responds with a DPBA message. Such messages are exchanged for informing the new location of the mobile node and establishing the tunnel for the redirectioning of the traffic, therefore, C_{update} cost is defined by:

$$C_{update}^{\text{Draft-Jaehwoon}} = 2C_{SMAG-AMAG} + PC_{AMAG} + PC_{SMAG} \quad (2)$$

As the periodical updates are performed as in PMIPv6, the same messages described above are exchanged between the MAGs at an$R_{BCE} = \frac{1}{\mu_{SN}T_{BCE}}$ rate, where T_{BCE} is the lifetime of a BCE entry [15]. Therefore, $C_{refresh}$ is given by:

$$C_{refresh}^{\text{Draft-Jaehwoon}} = R_{BCE}(2C_{SMAG-AMAG} + PC_{AMAG} + PC_{SMAG}) \quad (3)$$

Regarding C_{de-reg} cost, when the mobile node moves to a new MAG, DPBRU/DPBRA and Flush messages are exchanged for the removal of the old registrations from the tables. Therefore, the deregistration cost due to the new mobile node location is defined by:

$$C_{de-reg}^{\text{Draft-Jaehwoon}} = 4C_{SMAG-AMAG} + 2PC_{AMAG} + 2PC_{SMAG} \quad (4)$$

- Draft-Bernardos

The following expressions are used for the signaling cost components [16]:

$$C_{update}^{\text{Draft-Bernardos}} = (N_{PR} + 1)(2C_{CMD-MAAR} + PC_{CMD} + PC_{MAAR}) \quad (5)$$

$$C_{refresh}^{\text{Draft-Bernardos}} = R_{BCE}(2C_{CMD-MAAR} + PC_{CMD} + PC_{MAAR}) \quad (6)$$

$$C_{de-reg}^{\text{Draft-Bernardos}} = 4C_{CMD-MAAR} + 2PC_{CMD} + 2PC_{MAAR} \quad (7)$$

- SDN-DMM

In the SDN-DMM proposal the OpenFlow switch sends a message to the controller after detecting the mobile node has performed a handover. The controller verifies the change that occurred in the topology, recalculates a new path and sends OpenFlow messages, so that the switches involved establish a new communication path through bidirectional IP flows.

Similarly to the scenario between the movement of the mobile node in draft-Bernardos, which results in the allocation of new LNP addresses, and the switches affected in the data plane readjustment, the number of switches involved (except the one that identified the mobility) can be represented by N_{PR}, where C_{update} cost is represented by

$$C_{update}^{\text{SDN-DMM}} = 2C_{CTR-OFS} + PC_{CTR} + PC_{OFS} + (N_{PR})(C_{CTR-OFS} + PC_{OFS}) \quad (8)$$

where CTR indicates the cost related to the controller and OFS indicates the cost related to the OpenFlow switch. As the switch that received a message from the controller for adjusting the data-plane responds to the controller only if a problem has occurred for the establishment of the IP flow requested, the processing of the controller and the cost for the transmission of this related message are not added to C_{update} cost.

Regarding the periodical updates and the deregistration processes, the SDN-DMM approach does not address the concepts of registration and binding cache used for storing information on location for the establishment of tunnels that redirect packets when the mobile node performs a handover.

Through the knowledge of physical topology, the controller only adjusts the data plane of the network on demand according to the movement of the mobile node. The entries in the forwarding tables for the IP flows, due the handover of the mobile node, in the OpenFlow switches are inserted and removed in two ways, i.e., during the update process, when the controller sends messages to the switches involved for supporting mobility, by inserting or removing entries, and in the expiration of entries in the switches that no longer forward traffic related to the bidirectional IP flow established. Therefore, signaling cost $C_{SIG}^{\text{SDN-DMM}}$ is equivalent to $C_{update}^{\text{SDN-DMM}}$.

ii. Handover latency cost

This cost is given by the time spent from the disconnection of the old network to the ending of the mobility procedure that enables the mobile node to obtain global connectivity of the ongoing IP sessions again.

In general, the handover operation can be divided into three stages, according to [16]:

i. t_{L2} = time interval between the disconnection of the previous link and the connection to the new link;
ii. t_{auth} = time interval spent on the authentication and authorization processes due to the new association;
iii. $t_{binding}$ = time interval related to the mobility phase conducted according to the responsible protocol, through which functions, as bindings, routing adjustment and establishment of tunnels are performed.

As t_{L2} and t_{auth} are independent of the mobility protocol employed, only component $t_{binding}$ is considered for an analysis of the handover latency cost ($C_{handover}$) generated for each protocol, therefore,

$$C_{handover} = t_{binding} \tag{9}$$

$t_{binding}$ (Eq.10) depends on both operation and approach for the mobility in each proposal. However, it can be understood as the sum of the times necessary for the transmission and processing of the control messages.

$$t_{binding} = \sum RTT_{X-Y} + T_{proc}^{X} \tag{10}$$

where RTT_{X-Y} is the time for the packet exchange between the network entities X and Y; and T_{proc}^{X} is the processing time of a packet by entity X.

The handover latency costs for each proposal are evaluated below:

- Draft-Jaehwoon

The $t_{binding}$ process starts by S-MAG announcing the PREF network prefix through an RA message for the recently connected mobile node, which, after having visualized the PREF prefix, considers it is still connected to the same network, maintains its IP address and normally transmits its packets.

After S-MAG has received the packets, it verifies the mobile node has performed a handover process, therefore, it sends a DPBU message, so that A-MAG is informed about the new location of the mobile node. A-MAG responds to S-MAG with a DBPA message and establishes a tunnel between them for the addressing of the mobile node packets. At this moment, the mobile node is globally reconnected and can send and receive the packets through the tunnel established between the MAGs.

Therefore, the handover latency cost is given by:

$$C_{handover}^{Draft-Jaehwoon} = RTT_{MN-MAG} + RTT_{MAG-MAG} + T_{proc}^{MN} + 3T_{proc}^{MAG} \tag{11}$$

where RTT_{MN-MAG} is the time of the sending of the RA message from MAG to MN and return of the IP traffic from MN to MAG, $RTT_{MAG-MAG}$ is the time for the DPBU/DPBA message exchange between the MAGs, T_{proc}^{MN} is the processing time of the RA message by the MN, and $T_{proc}^{MAG}$is the MAG processing time performed once, when S-MAG receives the traffic, and twice, in the communication between the MAGs.

- Draft-Bernardos

The mobility phase starts when the mobile node is connected to the new S-MAAR and requests a new IP address. After S-MAAR has registered such an IP with CMD, it is informed about the existence of other prefixes that are still being used by the mobile node and have active sessions. Therefore, they must be treated by the mobility process, so that such IP sessions can be continued. S-MAAR establishes a tunnel with each A-MAAR for addressing the traffic corresponding to those prefixes.

Therefore, the handover latency cost comprises the sending of the RS message from the mobile node to S-MAAR, which sends a PBU message to CMD, which responds to S-MAAR with an extended PBA message and sends an extended PBU message to the other A-MAAR, and finally the message exchange between MAARs for the establishment of the tunnel.

As all MAARs simultaneously receive and process the extended PBA/PBU messages sent by CMD, the handover latency cost is given by:

$$C_{handover}^{Draft-Bernardos} = \frac{1}{2}RTT_{MN-MAAR} + RTT_{CMD-MAAR} + RTT_{MAAR-MAAR} + 3T_{proc}^{MAAR} + T_{proc}^{CMD} \quad (12)$$

where ½ $RTT_{MN-MAAR}$ is the time of the RS message from the mobile node to MAAR, $RTT_{CMD-MAAR}$ is the time of the PBU/PBA message exchange between CMD and MAAR, $RTT_{MAAR-MAAR}$ is the time between the message exchange between S-MAAR and A-MAAR for the tunnel establishment, T_{proc}^{MAAR} is the MAAR processing time, which occurs when it receives the RS message from mobile node, the extended PBA message from CMD, and the message for tunnel establishment from MAAR, and T_{proc}^{CMD} is the CMD processing time when it receives the PBU message from MAAR.

- SDN-DMM

Soon after the mobile node has connected to the new wireless network, maintaining its original IP address, it starts to normally transmit the packets. When the OpenFlow switch identifies the mobile node has performed a handover, it sends an OF message to the controller about the event and awaits instructions.

The controller calculates a new path through which the new bidirectional IP flow is established for the correct addressing of the mobile node packets due to its new location and sends to affected switches an OF message, so that they insert the corresponding IP flow in the forwarding table.

At this moment, the packets related to the mobile node are correctly forwarded by the network infrastructure according to the bidirectional IP flow established and global connectivity is again provided, which enables the continuity of the IP sessions.

As all OpenFlow switches simultaneously receive and process the messages sent by the controller, equation (13) gives the handover latency cost for the SDN-DMM approach:

$$C_{handover}^{SDN-DMM} = \frac{1}{2}RTT_{MN-OFS} + RTT_{CTR-OFS} + 2T_{proc}^{OFS} + T_{proc}^{CTR} \quad (13)$$

where $\frac{1}{2}RTT_{MN-OFS}$ is the time of the packet-delivery from the mobile node to the OpenFlow switch, $RTT_{CTR-OFS}$ is the time of the packet-exchange between the controller and the OpenFlow switch, T_{proc}^{OFS} is the processing time of the OpenFlow switch performed when it receives the packet from the mobile node and the OpenFlow message from the controller, and T_{proc}^{CTR} is the processing time of the controller.

V. RESULTS AND DISCUSSION

This section provides the results of the analytical evaluation of the proposals that evidence the benefits of an SDN-DMM approach, regarding costs of signaling and handover latency.

i - Signaling

For a general comparative result in the signaling cost evaluation, as the network topology affects the cost according to the DMM approach used, the cost for the transmission of a control packet among the network entities, cost C_{X-Y}, is the same for all solutions analyzed, i.e., $C_{MAG-MAG} = C_{CMD-MAAR} = C_{CTR-OFS} = C$ [15].

Moreover, for the cost calculation of the control packet processing by entity X, PC_X does not affect the signaling cost significantly, hence, the evaluation of a DMM approach perspective. The same processing cost was considered for all entities in the solutions, where $PC_X = PC$.

The comparison of the signaling cost between SDN-DMM and draft-Jaehwoon is shown in equation (14). It was obtained by the ratio between the SDN-DMM signaling cost, equation (8), and the draft-Jaehwoon signaling cost, which is the sum of equations (2), (3) and (4).

$$\frac{C_{SIG}^{SDN-DMM}}{C_{SIG}^{\text{Draft-Jaehwoon}}} = \frac{2 + N_{PR}}{6 + 2R_{BCE}} \quad (14)$$

As the comparison between SDN-DMM and draft-Jaehwoon uses the HNO address concept and the signaling process always involves at most two MAG pairs, N_{PR} can be considered the number of open flow switches involved in the signaling process for a bidirectional IP flow update and support to a mobile node handover.

Figure 4 shows the comparison of the signaling cost between SDN-DMM and draft-Jaehwoon with the dwell time variation in the subnet for different numbers of switches involved in the process.

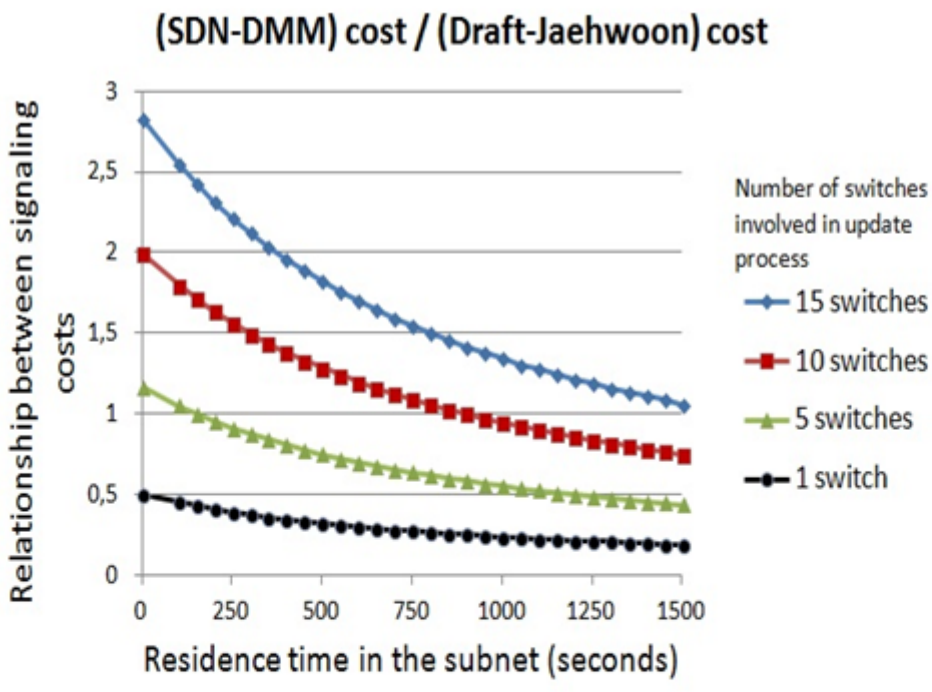

Figure 4 – Signaling cost: SDN-DMM vs Draft-Jaehwoon

As shown in Figure 4, as the bidirectional IP flow must be established between the switches affected by the handover process of the mobile node, the larger the number of OpenFlow switches affected by the update process of a new bidirectional IP flow, the higher the signaling cost of the SDN-DMM solution in comparison with draft-Jaehwoon, because the signaling change in draft-Jaehwoon is limited to at most two MAG pairs.

As the bidirectional IP flow alterations occur on the network edge, the number of switches affected by the update process probably does not exceed 10 switches, which may quickly decrease the cost of the signaling cost of SDN-DMM to below that of draft-Jaehwoon in function of the time of permanence of the mobile node in the subnet.

The signaling cost of SDN-DMM becomes lower as the time of permanence of the mobile node in the subnet increases, because the cost of periodical updates has a higher weight in the determination of the total signaling cost. As SDN-DMM does not involve such a periodical update cost, differently from draft-Jaehwoon, SDN-DMM requires a lower signaling cost.

The comparison between SDN-DMM and draft-Bernardos provides equation (15), which represents the ratio between the SDN-DMM signaling cost, according to equation (8), and the draft-Bernardos signaling cost, which is the sum of equations (5), (6) and (7).

$$\frac{C_{SIG}^{SDN-DMM}}{C_{SIG}^{\text{Draft-Bernardos}}} = \frac{2 + N_{PR}}{6 + 2N_{PR} + 2R_{BCE}} \quad (15)$$

Figure 5 shows the comparison between the signaling cost of SDN-DMM and draft-Bernardos, where the relation between signaling costs is shown according to the variation in the time of permanence of the mobile node in the subnet, which can be understood as the variation of its handover rate for four different numbers of active prefixes with current sessions.

The signaling cost of SDN-DMM is, at least, 50% lower in comparison with that of draft-Bernados because the OpenFlow switches that receive the messages from the controller in an update process for the establishment of a new bidirectional IP flow do not need to answer them. Therefore, the C_{update} cost of the SDN-DMM solution, which increases in function of the increase in the number of LNP prefixes, becomes lower.

Moreover, as SDN-DMM does not use messages for periodical updates and deregistration, when the mobile node remains longer in the subnet, which is a situation of higher weight of periodical updates in the calculation of the signaling cost, the relative cost of SDN decreases, as it does not have such a component.

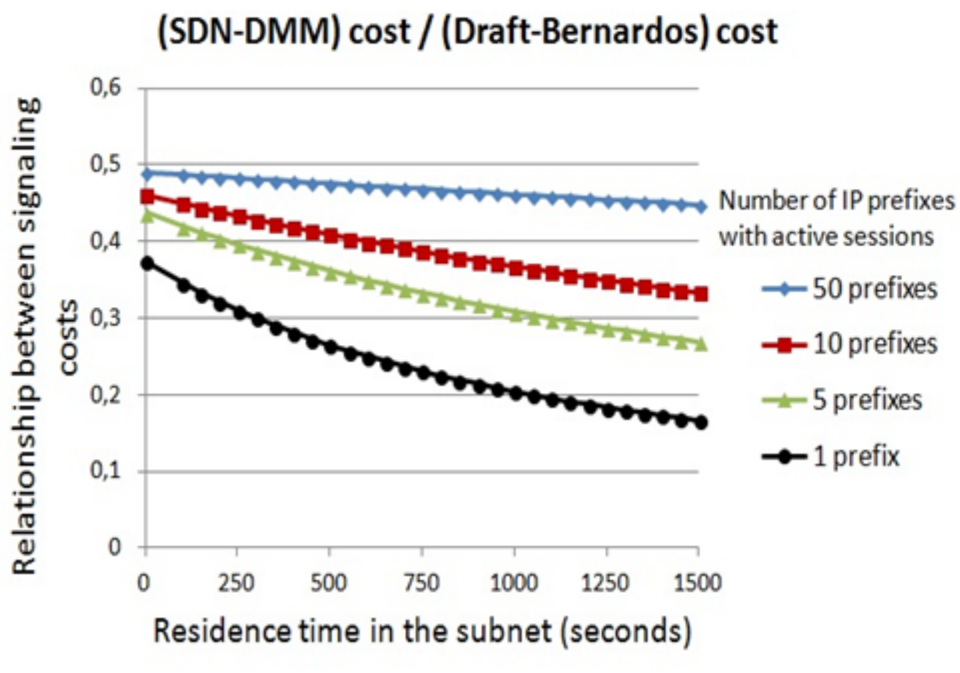

Figure 5 – Signaling cost: SDN-DMM vs Draft-Bernardos

ii- Handover latency

The time spent on the t_{L2} and t_{auth}handover operation stages and the packet exchange between the mobile node and its access point to the network, RTT_{MN-AP}, will be considered the same in all proposals analyzed, for the evaluation of the impact of stage $t_{binding}$ on the handover latency cost. Therefore, the variation of the components that comprise $t_{binding}$ becomes the most impacting factor on the total handover latency cost.

Based on [16], we consider the relation in (16) an equal packet processing performed by the network units, shown in (17), and a fixed distance between the central control element and each access point to the network in the access network in DMM partially distributed, expression (18).

$$t_{l2} + t_{auth} + RTT_{MN-MAG} + 2T_{proc}^{MAG} = t_{l2} + t_{auth} + RTT_{MN-MAAR} + T_{proc}^{MAAR} + T_{proc}^{CMD} = t_{l2} + t_{auth} + RTT_{MN-OFS} + T_{proc}^{OFS} + T_{proc}^{CTR} = T_{commun} \quad (16)$$

$$T_{proc}^{MN} = T_{proc}^{MAAR} = T_{proc}^{CMD} = T_{proc}^{OFS} = T_{proc}^{CTR} = T_{proc}^{Unit} \quad (17)$$

$$RTT_{CMD-MAAR} = RTT_{CTR-OFS} = RTT_{control} \quad (18)$$

A 50ms value, obtained by the experiments conducted in [15], is attributed to T_{commun}, whereas 1 ms is assigned to T_{proc}^{Unit}, as the time of packet processing by the unit is equal to the latency inserted by the packet processing in a hop between routers, which is typically lower than 1 ms [13] in a backbone network. 5 ms, 10ms and 20 ms, which are common values for the communication between the edge network equipment and the concentrator located in a Point of Presence along a backhall network [14] are use for $RTT_{control}$.

The equation (19), obtained by the ratio between equations (13) and (11), represents the relation between handover latency costs of SDN-DMM and draft- Jaehwoon.

$$\frac{C_{handover}^{SDN-DMM}}{C_{handover}^{\text{Draft-Jaehwoon}}} = \frac{T_{commun} + 4T_{proc}^{Unit} + 2RTT_{CTR-OFS}}{2T_{commun} + 4T_{proc}^{Unit} + 2RTT_{MAG-MAG}} \quad (19)$$

Figure 6 shows the impact on the relation between handover latency costs between SDN-DMM and draft-Jaehwoon, according to the latency variation as the mobile node conducts new handover processes in the network, which can be understood as the increase in the distance between MAGs that establish the tunnel for the addressing of the packets of the mobile node, so that it can be provided with global connectivity.

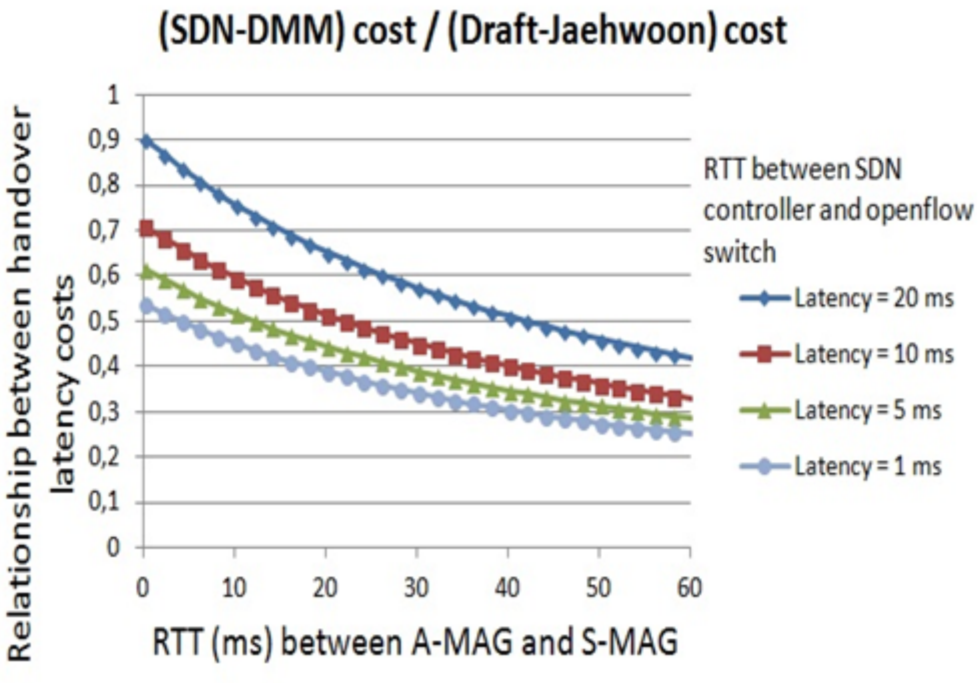

Figure 6 – Handover latency cost: SDN-DMM vs draft-Jaehwoon

Due to the approach employed for the supply of mobility to IP sessions in draft-Jaehwoon, in which process T_{commun} is required twice, in comparison with SDN-DMM, the handover latency cost is always lower when the distance between the concentrator and the OpenFlow switch is equal to the distance between the MAGs that perform tunneling. If a typical 5ms RTT is used for both, the cost reduction is approximately 44%.

Another important aspect is the implementation of the controller in relation to the OpenFlow switches in the access networks is generally performed in a way its location results in the shortest collective latency between them, i.e., the controller assumes a central position, so that as the mobile node performs a handover between the access networks, the latency between the OpenFlow switch and the controller is not significantly changed. Another implementation that reduces latency is the implementation of the controller in a hierarchical cluster mode, in which the controller responsible for certain access networks is implemented near them.

As MAGs are located on the edge of the access network, the movement of the mobile node in the access networks increases the latency between S-MAG and A-MAG, which does not occur in SDN-DMM and contributes to its lower handover latency cost during the movement of the mobile node, as shown by the curves in Fig. 6.

The relation betwen handover latency costs between SDN-DMM and draft-Bernardos is shown in equation (20), obtained by the ratio between equations (13) and (12).

$$\frac{C_{handover}^{SDN-DMM}}{C_{handover}^{\text{Draft-Bernardos}}} = \frac{T_{commun}+4T_{proc}^{Unit}+2RTT_{CTR-OFS}}{T_{commun}+6T_{proc}^{Unit}+2RTT_{CMD-MAAR}+2RTT_{MAAR-MAAR}} \quad (20)$$

Figure 7 shows the results provided by equation (20).

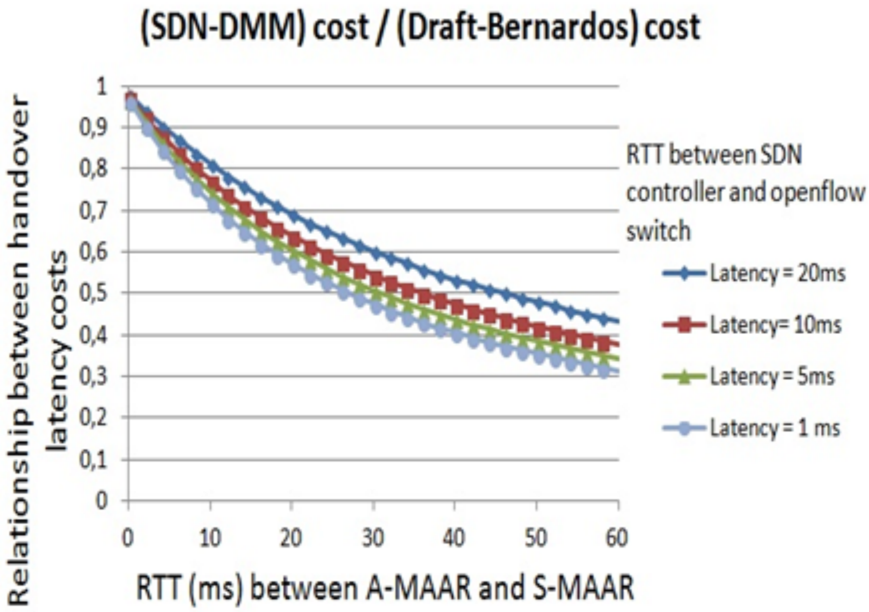

Figure 7 – Handover latency cost: SDN-DMM vs draft-Bernardos

The handover latency cost is higher in draft-Bernardos than in SDN-DMM due to the larger number of interactions in the $t_{binding}$ stage. First, the registration and update of the prefixes that are still active in the mobile node during the handover performed between S-MAAR and CMD are required, which results in component $RTT_{CMD-MAAR}$. The tunnel between A-MAAR and S-MAAR must then be established through the addition of component $RTT_{MAAR-MAAR}$, so that the traffic of the prefixes that are still active can be forwarded, which also results in a larger amount of T_{proc}^{Unit} processing.

As the distance between the control element and the point of access to the network are the same in the two proposals, i.e., $RTT_{CTR-OFS} = RTT_{CMD-MAAR}$, the predominant factor for the handover latency cost is $RTT_{MAAR-MAAR}$.

As in draft-Jaehwoon, as MAAR are located on the edge of the access networks, the latency between S-MAAR and A-MAAR may increase, as the mobile node performs handover, which increases the handover latency cost in comparison with SDN-DMM, as shown by the behavior of the curves in Fig. 7 with the latency increase in the x axis.

VI. CONCLUSIONS AND FUTURE WORK

This paper discussed aspects and proposals related to mobility management functions in an IP-based heterogeneous communications network, and proposed an architecture based on Software Defined Networking for Distributed Mobility Management (SDN-DMM), considering the need for session continuity, scalability and performance. The results show SDN-DMM satisfactorily deals with the main challenges of mobility management for an exponential increase in the number of mobile devices.

The proposal provides an efficient and scalable use of the available resources of the communication network infrastructure for the mobile node and the heterogeneous access networks, from both operational (OPEX) and evolutive and investment-based viewpoints through a network-based partially DMM approach associated with SDN and incorporating the intrinsic benefits of the paradigm to the solution.

We observed that SDN-DMM adequately addresses issues of scalability, performance and complexity involved in the distributed mobility management, allowing to ensure IP session continuity and using infrastructure resources related to mobility management tasks in an efficient way.

As a future work about SDN-DMM architecture, an extended evaluation should consider metrics as routing cost and packet delivery cost (treated as the cost for packets to be delivered between the mobile node (MN) and the correspondent node (CN) within the mobility domain [11]), aiming to discuss latency reduction of end-to-end communication and treating the formation of bottlenecks for the traffic forwarding, with the traffic being routed directly between the hosts without use of tunneling techniques.

Another future work involves routing optimization, considering that packets could be addressed directly to the mobile node based on the bidirectional IP flow and without the use of encapsulation techniques (which would add an overhead for both processing and transport of the payload through the communication network), allowing to improve the user quality of experience and efficiently using the available communication resources in a domain mobility scenario.

Future work also involves the validation of the model in a scenario of real experimentation, considering the need for data offloading of IP flows and extensive statistical analysis. Data from operating networks should be considered (for example, based on [17] and [18]). The support to a possible hybrid model (client-based + network-based) with SDN extension and the mobility management involving different autonomous systems has also been considered.

Finally, the impact of the use of SDN-DMM for dealing with multimedia traffic, specially due to need for meeting real time requirements of video streaming applications ([19]) and considering realistic mobility models represents a theme of research to be addressed.

Acknowledgements

This work has been supported by:

- CNPq ("Conselho Nacional para o Desenvolvimento Científico e Tecnológico") – Brazil;
- FAP-DF ("Fundação de Apoio `a Pesquisa do Distrito Federal") – Brazil;
- "Ministerio de Economía y Competitividad"/Spain, through the "Convocatoria 2017 - Proyectos I+D+I - Programa Estatal de Investigación, Desarrollo e Innovación, convocatoria excelencia" (Project TIN2017-84802-C2-1-P).

REFERENCES

[1] Karimzadeh, M., Sperotto, A, Pras, A. *Software Defined Networking to Improve Mobility Management Performance.* AIMS 2014: 118-122.

[2] Karimzadeh, M.; Valtulina, L; Karagianis, G. *Applying SDN/OpenFlow in virtualized LTE to support Distributed Mobility Management (DMM).* In: Proceedings of the 4th International Conference on Cloud Computing and Services Science, CLOSER 2014, 3-5 Barcelona, Spain. 86. SciTePress, Abril 2014.

[3] Cánovas, A., Taha, M., Lloret, J., Tomás, J. *Smart resource allocation for improving QoE in IP Multimedia Subsystems*, Journal of Network and Computer Applications, Volume 104, 15 February 2018, Pages 107-116.

[4] Bernardos, C., Oliva, A., Serrano, P., et al. *An Architecture for Software Defined Wireless Networking*. In IEEE Wireless Communications, vol. 21, no.3, pag 52-61, 2014.

[5] Jimenez, J.M., Martínez, J.O.R., Rego, A., Dilendra, A., Lloret,J. *Study of multimedia delivery over software defined networks*, Network Protocols and Algorithms 7 (4), 37-62. 2015.

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

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

[8] Chan,H.A.,Yokota,H., Xie, J., Seite, P. & Liu,D. *Distributed and dynamic mobility management in mobile internet: Current approaches and issues.* Journal of Communications, 6(1), 4–15, 2011.

[9] Y. Peña Llerena, P.R.L. Gondim, J. Lloret, *Improving throughput in DMM with mobile-assisted flow mobility*, Transactions on Emerging Telecommunications Technologies 29 (3). 2018.

[10] Chan, H. (2014). *Requirements for Distributed Mobility Management*. RFC 7333.

[11] Giust, F; Bernardos, C; Oliva, A. *HDMM: deploying client and network-based distributed mobility management*. Journal Telecommunications Systems, Volume 59 Issue 2, Pages 247-270, Junho 2015.

[12] Yokota, H., Seite, P., Demaria, E., & Cao, Z. Use case scenarios for Distributed Mobility Management. Internet-Draft (work in progress), draft-yokota-dmm-scenario-00.txt. 2010.

[13] J. Lee and Y. Kim, *PMIPv6-based Distributed Mobility Management.* IETF draft, draft-jaehwoon-dmm-pmipv6-04, June 18 2015.

[14] Bernardos, C.; Oliva, A.; Giust, F. *A PMIPv6-based solution for Distributed Mobility Management*. In IETF draft, draft-bernardos-dmm-pmip-06, Março 2016.

[15] ONLAB. *Introducing ONOS - a SDN network operating system for Service Providers.* Whitepaper http://onosproject.org/wp-content/uploads/2014/11/Whitepaper-ONOS-final.pdf [accessed Nov-2015].

[16] Giust F., Bernardos C., Oliva A. (2014). *Analytic Evaluation and Experimental Validation of a Network-Based IPv6 Distributed Mobility Management Solution.* In IEEE Transactions on Mobile Computing, vol. 13, no. 11, November 2014.

[17] Rede Nacional de Pesquisas (RNP). *Estatísticas de operação da Rede Ipê – Backbone da Rede Nacional de Ensino e Pesquisa (RNP),* 2016.

[18] Rede Nacional de Pesquisas (RNP). *Estatísticas de operação do Ponto de Presença da RNP no Distrito Federal (PoP-DF),* 2016.

[19] Taha, M., Garcia, L. Jimenez, J.M., Lloret, J. *SDN-based throughput allocation in wireless networks for heterogeneous adaptive video streaming applications,* The 13th International Wireless Communications and Mobile Computing Conference (IWCMC 2017), June 26-30, 2017, Valencia, Spain.

[20] F. Giust, A. de la Oliva, C. J. Bernardos, and R. P. F. Da Costa, *A network-based localized mobility solution for Distributed Mobility Management,* in Wireless Personal Multimedia Communications (WPMC), 2011 14th International Symposium on, pp. 1-5, 2011.

[21] C. J. Bernardos, J. C. Zunniga, and A. Reznik, *Towards flat and distributed mobility management: A 3GPP evolved network design,* in Communications (ICC), 2012 IEEE International Conference on, pp. 6855-6861, 2012.
[22] Carmona-Murillo, Javier, et al. *Performance Evaluation of Distributed Mobility Management Protocols: Limitations and Solutions for Future Mobile Networks.* Mobile Information Systems 2017 (2017).
[23] L. Yi, H. Zhou, D. Huang, H. Zhang, *D-PMIPv6: A distributed mobility management scheme supported by data and control plane separation*, Mathematical and Computer Modelling 58 (56) (2013) 1415 – 1426.
[24] K. Xie, J. Lin, L. Wu, *Design and Implementation of Flow Mobility Based on D-PMIPv6*, in: Computational Science and Engineering (CSE), 2014 IEEE 17th International Conference on, 2014, pp. 1344–1349
[25] W. Luo, J. Liu, *PMIP Based DMM Approaches*, Internet-Draft draft-luo-dmm-pmip-based-dmm-approach-02, Internet Engineering Task Force (July 2013).
[26] P. Seite, P. Bertin, J.-H. Lee, *Distributed Mobility Anchoring*, Internet- Draft (Work in Progress) draft-seite-dmm-dma-07, Internet Engineering Task Force (February 2014).
[27] D. R. Purohith, A. Hegde, K. M. Sivalingam, *Network architecture supporting seamless flow mobility between LTE and WiFi networks*, in: World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2015 IEEE 16th International Symposium on a, 2015, pp. 1–9.