A Hybrid Centralized-Distributed Mobility Management Architecture for Network Mobility

Tien-Thinh Nguyen and Christian Bonnet
Department of Mobile Communications
EURECOM
Sophia-Antipolis, France
Email: {Tien-Thinh.Nguyen, Christian.Bonnet}@eurecom.fr

***Abstract*—In the mobile era, the demand for Internet connection for moving vehicles (such as cars, buses and subways) is growing fast. The Network Mobility (NEMO) basic support protocol (B-NEMO) was introduced to provide Internet access for a group of users in a moving vehicle in an effective manner. As an extension of Mobile IPv6 (MIPv6), B-NEMO inherits the limitations from both the host-based and the centralized mobility management protocol such as sub-optimal routing (especially, in the nested NEMO scenario), reliability and scalability issues. Recently, Distributed Mobility Management (DMM) has been introduced as a new trend to overcome the limitations of the centralized mobility management protocols. However, DMM may not be a suitable scheme for moving vehicles since it faces several challenges such as complex address and tunnel management, high signaling cost, and long handover latency in case of users moving at a high speed. In this document, we propose a hybrid centralized-distributed mobility management architecture in the context of NEMO. Our solution allows the devices to obtain connectivity either from fixed locations or mobile platforms (e.g., a NEMO) and move between them, while keeping their on-going flows. The numerical results showed that the solution helps keeping the advantages of DMM approach in terms of signaling cost, packet delivery cost, handover latency, and end-to-end delay even in the moving vehicle scenario.**

Keywords*—*IP Mobility, Proxy Mobile IPv6, Distributed Mobility Management, Network Mobility.

I. Introduction

Technology has now driven us in the mobile era in which the number of mobile phones accounts for more than 87 percent of worldwide devices shipments (including traditional PC, tablets and mobile phones) in 2014 [1]. Accordingly, estimates say that global mobile broadband subscriptions grew by around 30 percent year-on-year and reached 2.5 billion in the third quarter of 2014 [2]. Additionally, this trend does not show any sign of slowing down. The increasing number of subscriptions has been driven by a variety of reasons such as the increasing number of mobile devices which become more and more powerful and intelligent (especially, in the low- and mid-range price), the enhancement of wireless access technology in terms of coverage, speed and quality, as well as the explosion of mobile applications [2], [3]. As a result, mobile data traffic will increase around 8 times between 2014 and 2020, to reach 17 exabytes per month by 2020 [2]. On the other hand, the mobility of the devices puts new requirements on mobile operators to deliver services anywhere and at any time, for example, providing Internet connection on moving vehicles such as cars, buses and subways.

In order to provide Internet access for a group of users in a moving vehicle in an effective manner, the concept of Network Mobility (NEMO) and the NEMO basic mobility support protocol (B-NEMO) [4] have been introduced. However, as an extension of the well-known protocol - Mobile IPv6 (MIPv6), B-NEMO inherits some drawbacks from MIPv6 such as sub-optimal routing and signaling overhead (especially, when a nested NEMO is considered). Additionally, as a centralized management protocol, the central mobility anchor represents a bottleneck and single point of failure [5].

To tackle the problems of the centralized mobility management (CMM) protocol (particularly, when considering a huge number of traffic demand), a new paradigm, called Distributed Mobility Management (DMM), has been introduced. A lot of work has been done considering different DMM approaches [6]. All of them proved that DMM is generally a promising scheme. The reason is that DMM allows the traffic to be easily offloaded from the core to the network edge due to the fact that the mobility anchor is put at the network edge. In addition, the mobility support is enabled when necessary and the traffic is better distributed among the network entities. Altogether, DMM helps reduce the network congestion and resources waste. However, DMM also leads to several issues such as complex address and tunnel management, high signaling cost and long handover latency as the number of addresses and the number of bi-directional tunnels associated with the MN increase, for example, in case of users moving at a high speed (such as smart phone users on vehicles) and/or with long-lasting flows [7], [8], [9]. As a result, DMM may not be a suitable scheme from the vehicle perspective.

In the context of NEMO, several proposals have been introduced to address the limitations of B-NEMO. However, some of them (e.g., [10], [11]) leverage on the centralized approach, thus inheriting the limitations of the centralized protocol. The others i.e, [12], [13], [14], which are based on the DMM concept, are not yet complete solutions. For instance, when the mobile nodes perform a handover from a fixed infrastructure to a NEMO, their on-going flows will be interrupted. Moreover, they encounter the issues of DMM in

This work has been partially supported by the French project SYSTUF. EURECOM acknowledges the support of its industrial members: BMW Group Research & Technology, IABG, Monaco Telecom, Orange, SAP, SFR, ST Microelectronics, Symantec.

case of vehicle as mentioned earlier.

In this document, we propose a hybrid centralized-distributed mobility management architecture in the NEMO context (called H-NEMO). As a network-based approach, our solution provides mobility support for all legacy devices. That means, these devices can obtain connectivity either from fixed locations or mobile platforms (e.g., a NEMO) and move between them, while keeping their on-going flows (without their involvement in signaling procedure). Furthermore, an appropriate approach (among DMM and PMIPv6) will be applied according to the flow characteristics and the mobility feature of the node, thus mitigating the limitations of both centralized and distributed mobility management approach. For instance, the flow will be routed via the central mobility anchor for a long-lived flow or a high mobility node. Otherwise, it will be routed following the DMM concept. The numerical results showed that H-NEMO generally outperforms the centralized-based schemes (and other distributed-based ones as well) not only in case of low mobility and/or a short-lived flow scenario but also in a high mobility and/or long-lived flow scenario in terms of signaling cost, packet delivery cost, handover latency and end-to-end delay.

The rest of this paper is organized as follows. Section II presents the background information related to IP mobility management in the context of NEMO. Section III describes the proposed solution regarding its design principles, architecture and operations. Section IV provides the qualitative and quantitative analysis including the numerical results. Finally, Section V concludes the paper and provides perspectives for future work.

II. Related Work

A. IP Mobility Management: From Centralized to Distributed Mobility Management

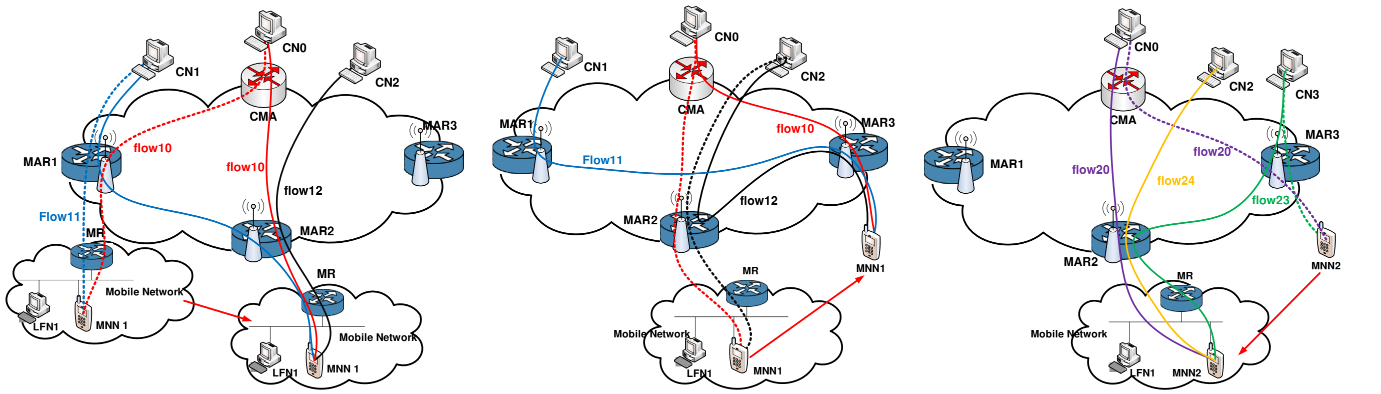
Mobile IPv6 (MIPv6), the first mobility protocol for IPv6, maintains the mobile node's (MN) reachability when it is away from home by introducing a central entity (namely Home Agent - HA) which is a topological anchor point of the permanent MN's IP address. However, as a host-based protocol, requiring MIPv6 protocol stack at the MN (to perform the mobility-related signaling by means of location update procedure) is the main obstacle for the deployment of MIPv6 in the reality. Therefore, Proxy Mobile IPv6 (PMIPv6) [15], was introduced as a network-based protocol to avoid the additional deployment at the MN. While moving inside a PMIP domain, the MN remains its IPv6 address. Thus, from IP layer point of view the MN is unaware of mobility. In other words, mobility can be transparently provided to all legacy devices. This is achieved thanks to the network entity - Mobile Access Gateway (MAG), which performs the mobility-related signaling on behalf of the MNs attached to its access links. In PMIPv6, the Local Mobility Anchor (LMA), similar to the HA in MIPv6, is responsible for maintaining the MN's reachability state and forwarding traffic from/to the current location of the MN. Accordingly, MN's traffic is always encapsulated and tunneled between the MN's LMA and the corresponding MAG.

Today's mobility management protocols (e.g., MIPv6 and PMIPv6) have several major limitations from their centralized and hierarchical nature. Centralizing both the control and data plane functions at the central mobility anchor introduces scalability issues [5]. Also, it leads to sub-optimal paths between the mobile nodes and their corresponding nodes (CNs). To address the limitations of the current mobility management protocol, Distributed Mobility Management (DMM) solutions have been proposed [5]. The key concept of DMM is that instead of having a centralized mobility anchor, the mobility anchors are distributed among network entities and placed as close as possible to the MN e.g., at the router edge of the access network. Also, mobility support is dynamically provided for the sessions when it is really needed.

As DMM is currently a hot topic which gains much interest from both academia and industry, a lot of research publications [6], [7], [8] have carried out the analysis on different DMM approaches, compared them with the conventional mobility management protocols in terms of signaling cost, packet delivery cost, handover delay, packet loss, and end-to-end delay. The results from these analysis showed that DMM is generally a promising scheme to deal with a huge number of data traffic over mobile networks. It is because DMM helps to save the resources in the network in some scenarios since the mobility support is enabled when necessary and the traffic is better distributed among the network entities, thus improving the scalability and reliability of the network. In DMM, the MN obtains a new prefix after each handover, while keeping the old prefixes as long as their on-going flows are still alive. The on-going flow will be routed via the tunnel between the MAR where the flow was initiated (called anchor MAR or aMAR) and the current one (cMAR). As a result, in the situation when the MN is running the long-lasting flows and/or a node with high mobility, the DMM deployment encounters several challenges such as a complex address and tunnel management, a high signaling cost and a long handover latency [7], [8], [9]. Therefore, DMM may not be a suitable scheme for a vehicle or a user on a moving vehicle.

B. Network Mobility (NEMO)

Network Mobility (NEMO) refers to the mobility of an entire network which changes its point of attachment to the Internet. Thus, the main purpose of NEMO support is that it allows every node in the mobile network to be reachable while moving around. Moreover, the mobility should be transparent to the nodes inside the mobile network. Following these concepts, the basic network mobility support (B-NEMO) [4] is introduced based on MIPv6. In B-NEMO, a specific gateway called Mobile Router (MR) is presented. The MR will be connected to the fixed infrastructure and provides connectivity to the nodes inside the mobile network. Like the mobility support of a mobile node (host-based approach), B-NEMO is also based on the bi-directional tunnel between the MR and its HA to enable mobility support when the MR (and its associated MNs) is away from home. Thus, as a topological anchor point of MR/MN address, the data packets destined to the mobile network are delivered to the HA, which then tunnels them towards the MR. The MR, after removing the tunnel headers, forwards the packets to the destination inside the mobile network. However, as an extension of MIPv6, B-NEMO inherits the limitations from both the host-based and the centralized mobility management protocol. As a result, several NEMO solutions have been proposed to overcome these limitations.



(a) MN attachment and MR handover (b) MN moves from a NEMO to a fixed network (c) MN moves from a fixed network to a NEMO

Fig. 1: Architecture and operations of the proposed solution.

Following the centralized manner, in [10] (NC-Soto), the authors suggest that the MR plays the role of a moving mobile MAG, thus allowing the MN to move between a NEMO and a fixed MAG. In [11], a cost-efficient NEMO scheme (NC-Jeon) is proposed. On one hand, both of them are based on PMIPv6 which is a network-based approach, thus, addressing the limitations of the host-based protocol. However, they do not consider the nested scenario. Also, they inherit the limitations of the centralized mobility management approach.

Regarding the distributed-based solution, several papers have been done to bring the DMM concept into the context of NEMO. Yet, they are not complete solutions. Specifically, in [12] a distributed mobility scheme for NEMO based on MIPv6 (HD-NEMO) is proposed. The MR's HAs are distributed among the routers at the network edge. Again, as a host-based protocol, the MN needs an additional deployment to participate into the mobility support. In [13], the authors propose a network-based DMM scheme for NEMO (ND-Ernest) including the nested scenario. Yet, it is not clear how the distributed DMM works. In [14], the authors introduce a network-based DMM for NEMO (ND-Do), however, this scheme does not address the nested NEMO scenario. Additionally, the two last proposals do not support the mobility of a node from a fixed infrastructure to a NEMO. That means, in this case, the on-going flows can not be maintained.

III. Hybrid Centralized-Distributed Mobility Management Architecture for Network Mobility

A. Design Principles

In order to overcome the limitations from the host-based, the centralized, as well as the distributed mobility management protocol, we proposed a solution guided by a set of design principles:

- Transparency: The mobile node can be totally free to roam between a NEMO and a fixed infrastructure without signaling involvement (it means the network-based mobility is used);
- Dynamicity: The dynamic use of an appropriate mobility scheme (among PMIPv6 and DMM) allows mitigating the limitations of a centralized/distributed approach;
- Addressing the additional issues related to mobility support for NEMO such as pinball issue;
- Supporting nested NEMO scenario.

B. Architecture and Operation

The architecture and operations of the solution are illustrated in Fig. 1. The proposed solution works in a similar way to the one described in [16]; however, the solution in [16] does not consider the NEMO scenario. Our solution relies on the network-based DMM scheme proposed in [17]. In this case, the Central Mobility Database (CMD) is extended to the Central Mobility Anchor (CMA) which behaves as an LMA.

In more details, when an MR attaches to an MAR's subnet, it obtains a prefix (and a set of prefixes for its associated nodes) from this MAR (called MAR-prefix) following the DMM behaviour. It can also get a set of prefixes from the CMA (called LMA-prefix) as in a PMIPv6 domain. Each time the MR performs a handover, it keeps the LMA-prefix while obtaining a new prefix from the new MAR. Accordingly, the MR can configure two different addresses from the allocated prefixes. The MR then can select between the two addresses to start new flows. Similarly, when an MN attaches to the MR, it obtains two different prefixes and configures two corresponding addresses. The MN then can start new flows using one of the two addresses (e.g., flow10 using LMA-address and flow11 using MAR-address as depicted in Fig. 1). At this stage, for the long-lived flow, the MR/MN should select the LMA-address, while for the short-lived flow, the current MAR-address should be used. By doing so, the average number of active prefixes of the MR/MN is kept lower than a threshold value (N_0). Our solution also allows maintaining the on-going flows of the MN in case of MR handover (see Fig. 1(a)) as well as the MN handover between a NEMO and a fixed infrastructure (see Fig. 1(b) and Fig. 1(c)).

To allow the MR/MN to select the appropriate prefix to start a new flow (upon several metrics such as the duration of the flow and the mobility features of the node), a connection manager (CM) application [18] which is located on the MR/MN can be used. The CM can be typically used to simplify the configuration and selection among the available interfaces, as well as to store the user's preferences, security profiles, and manage handoff process. Moreover, in case of

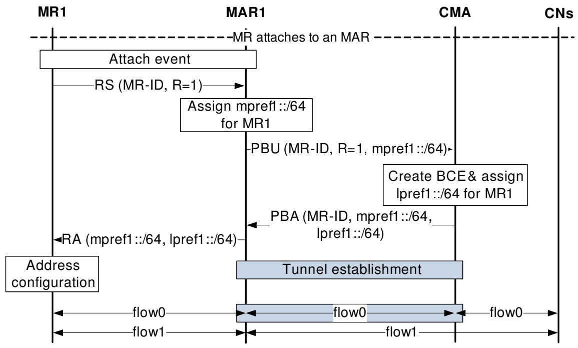


Fig. 2: Initial attachment of the MR.

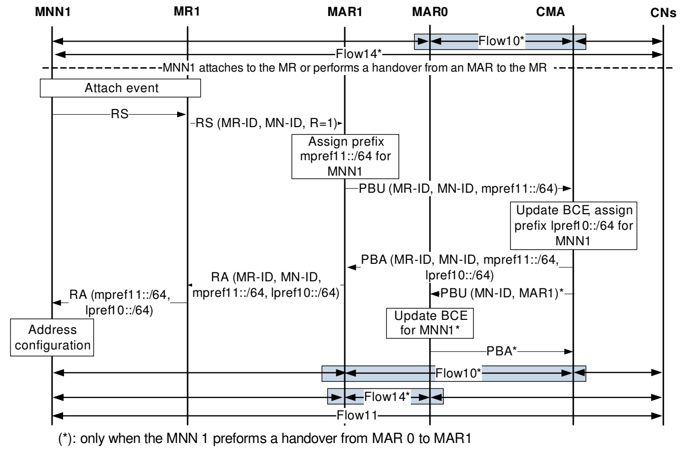


Fig. 3: Initial attachment of the MN.

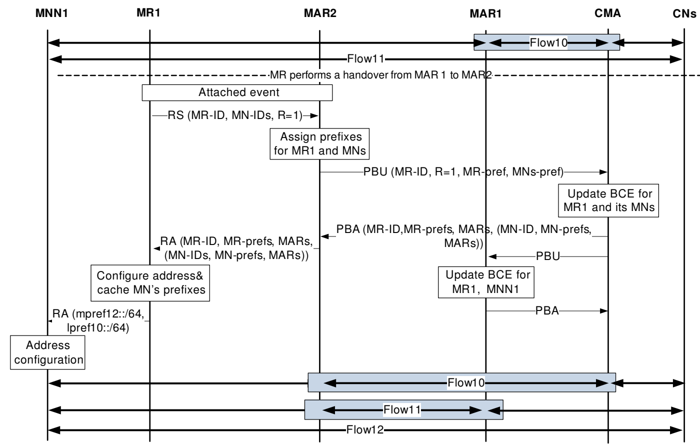


Fig. 4: Signaling flow for MR handover between two MARs.

multiple interfaces, one possible solution is that one interface is dedicated for LMA-prefix while the other is for MAR-prefix. However, the detail of the selection process is out of scope of this document.

1) Initial Registration: The signaling flow when an MR (MR1) attaches to an MAR (MAR1) is illustrated in Fig. 2. After receiving the Router Solicitation (RS) message from MR1 which includes its identifier (MR1-ID) as well as a flag (R) set to value of 1 (indicating that it is acting as an MR), MAR1 allocates a prefix for MR1 (called mpref1::/64). MAR1 then sends a Proxy Binding Update (PBU) to the CMA. As MR1 enters the domain for the first time, the CMA creates a Binding Cache Entry (BCE) for it. Additionally, acting as an LMA, the CMA allocates a prefix for MR1 (called LMA-prefix, lpref1::/64) and updates its entry with the new prefix. The CMA then replies by a Proxy Binding Acknowledgement (PBA) consisting of the allocated prefixes. Upon receiving the PBA, MAR1 sends these prefixes to MR1 by means of a Router Advertisement (RA) message. MR1 then, based on these prefixes, configures its IPv6 addresses (mpref1::MR1/64 and lpref1::MR1/64) and can start new communication flows using these addresses (e.g., flow0 and flow1 using lpref1::/64 and mpref1::/64, respectively). It is noted that flow0 is routed via the mobility tunnel between MAR1 and CMA following PMIPv6 behaviour while flow1, following DMM concept, is delivered using a standard routing manner.

When a mobile node (in this case called mobile network node or MNN, let says MNN1) attaches to MR1, MR1 sends a RS message including the MR-ID, MN-ID, flag (R) to MAR1 to obtain a prefix for MNN1. After assigning a prefix for MNN1 (mpref11::/64), MAR1 exchanges PBU/PBA with the CMA for registration process as well as obtaining the LMA-prefix for this MNN (lpref10::/64). MAR1 then sends a RA message including two allocated prefixes to MR1 as similar to that in the previous paragraph. MR1, after caching these prefixes, includes them in a RA sent to MNN1. MNN1 then configures two addresses (mpref11::MNN1/64 and lpref10::MNN1/64) and can use these addresses to start new flows as depicted in Fig. 3 (e.g., flow11). If the mobile node moves from a fixed infrastructure (another MAR) to the MR, the similar operations are executed. The CMA also sends PBUs to the MNN's anchor MARs to update its location. The CMA also adds the list of MNN's anchor MARs and the associated prefixes in the PBA sent to MAR1. The tunnel is then established between MAR1 and each MNN's anchor MAR to redirect the traffic destined to the previous addresses of MNN1 (e.g., flow10 and flow14 in Fig. 3).

2) Handover Operation for a Non-nested NEMO: After attaching to a new MAR (see Fig. 4), MR1 sends an extended RS message including the MR-ID, flag (R), and the list of MN-IDs associated to its subnet to the current MAR (MAR2). Based on these information elements, MAR2 allocates a set of prefixes for the MR as well as for its MNs. Again, MAR2 exchanges PBU/PBA messages with the CMA to update the location of MR1/MNs and obtain the address of the anchor MARs where MR/MNs was attached (MAR1). In our solution, the LMA-prefixes are included in the PBA sent to MAR2. After that, MAR2 unicasts a RA including the prefixes allocated to MR1. MR1 in turn configures its addresses similar to the previous section, caches the prefixes for the MNs, and advertises RA messages to each attached node. Based on this information, MNN1 configures a new MAR-address (mpref12::MNN1/64) while keeping the LMA-one (lpref10::MNN1/64). In parallel, CMA informs the MR1/MN's previous MARs the current location of MR1/MNs. The tunnels are then established between MAR2 and MAR1; and between MAR2 and CMA to route the traffic for flow11 and flow10, respectively. Following the DMM concept, MNN1 can start a new flow (let say flow12) which is delivered in a standard routing manner without any tunneling mechanism.

3) Handover Operation of a Mobile Node from a NEMO to a Fixed Infrastructure (MAR): When MNN1 performs a

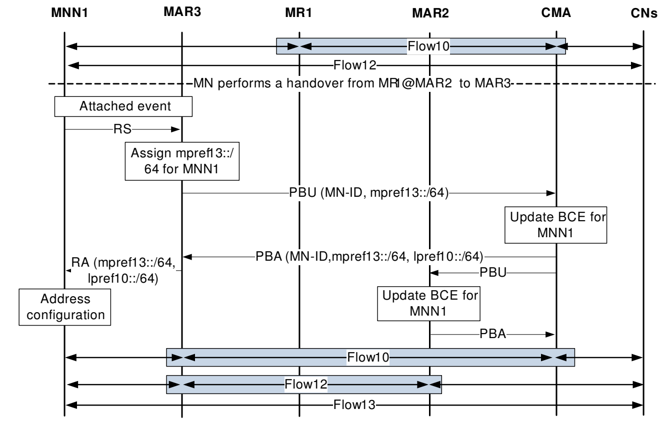


Fig. 5: Handover signaling when an MN moves from a NEMO to an MAR.

handover from MR1 (attached to MAR2) to an MAR (MAR3), the similar operations as described in the previous section will be executed to update the location of MNN1, and obtain a new MAR-prefix (mpref13::/64). Based on this prefix, MNN1 configures its new MAR-address while remaining the LMA-one (lpref10::MNN1/64). It then can use these addresses to start a new flow (e.g., flow13 using mpref13::/64). A bi-directional tunnel is also established between MAR3 and CMA for the traffic destined to lpref10::/64 (flow10) similar as in PMIPv6, and between MAR2 and MAR3 for the traffic destined to mpref12::/64 (flow12) as well (see Fig. 5).

4) Mobility Support for a Nested NEMO: Now we consider the mobility support in a nested scenario: MR2 (with MNN2) moves from MAR3 to MR1 (attached to MAR2). Our solution ensures the continuity of the on-going flows, for example, flow20 and flow23 which were started when MNN2 at MAR3 and using lpref20::/64 and mpref23::/64, respectively. The detailed operations for this scenario are described in Fig. 6. When attaching to MR1, MR2 sends a RS message including its identifier, and the list of MN-IDs associated to its network to MR1. MR1 then requests a set of prefixes of MAR2 for MR2. The same operations as in the previous section will be executed at MAR1, CMA to allocate the MAR-prefixes and LMA-prefixes, update the corresponding BCE for MR2 and its MNs, and establish the tunnel MAR2-MAR3 (if necessary). After receiving the prefixes allocated to MR2 and its attached nodes, MR1 unicasts a RA to MR2. Based on these information, MR2 configures its addresses and advertises a RA to each attached MNN, which then configures a new MAR-address while keeping the LMA-one. Finally, the on-going flows are delivered to the nodes as depicted in Fig .6.

IV. Performance Analysis

A. Qualitative Analysis

Table I provides a summary of the proposed approach characteristics compared to the existing proposals. A detail of each proposal is provided in [4], [10], [11], [12], [13], [14].

The proposed solution, similar to NC-Soto, NC-Jeon, ND-Enerst, and ND-Do is a network-based approach. That means the MN does not need to modify its IP stack. Also, the mobility scope is limited to a local domain to reduce the signaling overhead of the MN. On the contrary, the host-based

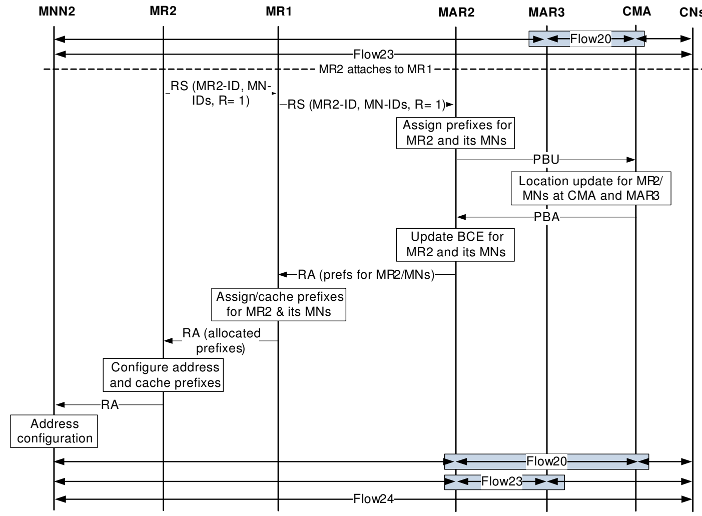


Fig. 6: Signaling flow for a Nested NEMO scenario.

approaches (B-NEMO and HD-NEMO) allow the MN to roam globally at a cost of additional deployment at the MN.

Unlike the previous proposals, which are either centralized or distributed ones, H-NEMO allows a dynamic selection between the two approaches, thus providing a flexible way to mitigate the limitations of each approach. Also, in our solution, the node can move between a fixed infrastructure and a NEMO while maintaining its on-going flows, making our solution different from the network-based DMM proposals (ND-Enerst and ND-Do). It is because the information related to the MNN is also stored at the CMA.

B. Quantitative Analysis

For the purpose of quantitative analysis, we compare H-NEMO with the existing proposals including B-NEMO, one network-based centralized (NC-Soto), one host-based distributed (HD-NEMO), and one network-based distributed protocol (ND-Do) regarding different metrics including signaling cost, packet delivery cost, handover latency, and end-to-end delay.

1) Reference Model: Fig. 7 shows a reference topology for performance analysis. The hop-count distances between the entities for performance analysis are defined as follows: h_{mc} - the average number of hops between MAR and CMA, between MAG and LMA; h_{ac} - the average number of hops between the MAR where the prefix is allocated (aMAR) and the current MAR (cMAR); h_{lc}, h_{cn} and h_{hc} - the average number of hops between the CN and the LMA/CMA, MAR, HA, respectively; h_{hl} and h_{mh} - the average number of hops between the HA and the LMA/CMA and MAR, respectively; h_{mm} - the average number of hops between two adjacent MARs; h_{hh} - the average number of hops between two HAs (in the host-based approaches). It is noted that the number of hops between the MR and its MAR (h_{mr}) or between the MN and its MR/MAR (h_{mn}) are assumed to be one (wireless link). And, m represents the number of MNNs attached to the MR.

2) Mobility Model: In this paper, we assume that the MAR/MAG subnet residence time is a random variable which follows an exponential distribution with mean value $1/\mu$ and

TABLE I: Qualitative analysis: Comparison between the existing proposals for NEMO and H-NEMO.

Metrics/Scheme	**B-NEMO**	**NC-Soto**	**NC-Jeon**	**HD-NEMO**	**ND-Enerst**	**ND-Do**	**H-NEMO**
Mobility management type	Host-based	Network-based	Network-based	Host-based	Network-based	Network-based	Network-based
Centralized or Distributed?	Centralized	Centralized	Centralized	Distributed	Distributed	Distributed	Dynamic*
Mobility scope	Global	Local	Local	Global	Local	Local	Local
MN modification	Required	Not required	Not required	Required	Not required	Not required	Not required
Roaming from a NEMO to a MAR/MAG	Yes	Yes	No	Yes	Yes	Yes	Yes
Roaming from a MAR/MAG to a NEMO	Yes	Yes	No	Yes	No	No	Yes
Support nested NEMO	Yes	-	-	Yes	Yes	-	Yes
Pinball issue	Yes	-	-	No	No	-	No
Tunnelling over wireless link (MAR/HA-MR)	Yes	Yes	No	Yes	No	No	No
Corresponding Node	Anywhere	Anywhere	Anywhere	Anywhere	Inside domain	Anywhere	Anywhere
(-): Not specified, (*): Centralized/Distributed							

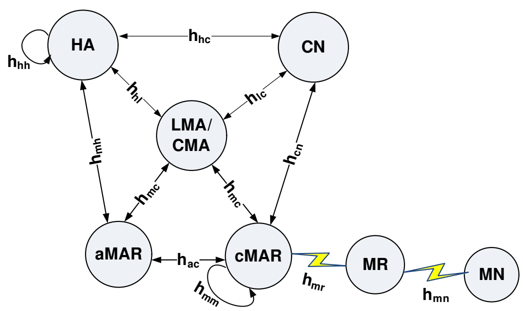


Fig. 7: Reference network topology.

the MAR/MAG coverage area is circular with radius R. According to [19], the subnet border crossing rate μ is calculated as:

$$\mu = \frac{2\upsilon}{\pi R}, \tag{1}$$

where υ is the average velocity of the MR/MN.

3) Cost Analysis: In this paragraph, the signaling cost ($SC_{(.)}$) and the packet delivery cost ($PC_{(.)}$) are investigated. The signaling cost (per handover) is the signaling overhead for updating the location ($C^u(.)$) as well as for refreshing the bindings ($C^r_{(.)}$) for the MR itself and its attached MNs. Although different signaling messages have different sizes, we assume that they have the same size for simplicity. Also, the cost for transmitting a signaling message is supposed to be proportional to the distance between the source and the destination. The proportion is α for wired and β for wireless link. We have:

$$SC_{(.)} = \mu\left(C^u_{(.)} + C^r_{(.)}\right). \tag{2}$$

In our analysis, we only consider the case where the MNNs are always with their MR (do not consider the case where the MNN moves from/to the fixed infrastructure). Accordingly, the MR needs only one PBU/PBA (or BU/BA) message to update/refresh the bindings for all the attached nodes. It is noted that in case of HD-NEMO, the functionality of HA is deployed at the MAR. $C^u_{(.)}$ can be therefore given by:

$$C^u_{B-NEMO} = (5+m)\beta + 2\alpha h_{mh}, \tag{3}$$

$$C^u_{NC-Soto} = (4+m)\beta + 4\alpha h_{mc}, \tag{4}$$

$$C^u_{HD-NEMO} = (2+m+2\overline{N}_p)\beta + 2\overline{N}_p\alpha h_{ac}, \tag{5}$$

$$C^u_{NC-Do} = (2+m)\beta + 2\alpha h_{mc} + 2\overline{N}_p\alpha h_{ac}, \tag{6}$$

where $\overline{N}_p$ is the average number of used active prefixes of the MR. It is worth noting that a prefix of the MR should be kept alive in case there exist at least one MR's on-going flow using this prefix or at least one prefix of the MNN allocated at the same MAR is still valid. For the sake of simplicity, we assume that the mean value of active lifetime of the MR and its attached nodes when they are visiting a foreign network are the same (with the value of $1/\delta$). According to [7], $\overline{N}_p$ is calculated as:

$$\overline{N}_p = 1 + \frac{\mu}{\delta}. \tag{7}$$

In the context of this document, the low value of $\overline{N}_p$ represents the low mobility and/or the short-lived flow scenarios. The higher value of $\overline{N}_p$ corresponds to the high mobility and long-lived flow scenarios.

In case of H-NEMO, since the number of active prefixes (excluding LMA-prefix) is kept below the threshold value (N_0), we have:

$$C^u_{H-NEMO} = (m+2)\beta + 2\alpha h_{mc} + 2\overline{N}_m\alpha h_{mc}, \tag{8}$$

where $\overline{N}_m = min\left(\overline{N}_p, N_0\right)$.

Since even the MR remains at the same subnet, the signaling for refreshing the bindings is sent periodically when the binding timer expires (for both MR and MNNs). This procedure is executed on average $R_B = 1/(\mu T_{BCE})$ times when the MR attaches to an MAR, where T_{BCE} is the binding cache entry lifetime. Consequently, $C^r_{(.)}$ is given by:

$$C^r_{B-NEMO} = R_B\left(2\beta + 2\alpha h_{mh}\right), \tag{9}$$

$$C^r_{NC-Soto} = R_B\left(2\beta + 4\alpha h_{mc}\right), \tag{10}$$

$$C^r_{HD-NEMO} = R_B\left[2\overline{N}_p\beta + 2\alpha(\overline{N}_p - 1)h_{ac}\right], \tag{11}$$

$$C^r_{NC-Do} = C^r_{H-NEMO} = 2R_B\alpha h_{mc}. \tag{12}$$

Regarding the packet delivery cost ($PC_{(.)}$), it represents the accumulative cost to deliver packets between an MNN and a CN per unit of time. It is proportional to the distance between the MNN and the CN and the number of packets transmitted. Let λ_{pkt} denote the packet transfer rate and $\lambda_{pkt} = \lambda_p\overline{N}_p$ where λ_p is the packet rate per active prefix. In a centralized-based solution, since only one prefix is active, the total packet rate is λ_{pkt} and the packet is routed via the route MNN-MR-MAR(MAG)-LMA-CN. As in H-NEMO, the MNN can select either the LMA- or MAR-prefix to start a new flow, the packet delivery cost thus depends on the PMIP- or DMM-like behaviour (PC_{PMIP} and PC_{DMM}, accordingly). The packet rate for the LMA-prefix is $\lambda_{lp} = \left(\overline{N}_p - \overline{N}_m\right)\lambda_p$. In a distributed-based solution, the route is MNN-MR-cMAR-aMAR-CN (for the tunneled packet) and MNN-MR-cMAR-CN (for the packet of the flow started at cMAR). Note that the tunneled packet (cMAR-aMAR) belongs to $\overline{N}_p - 1$ prefixes (except the prefix allocated in the current MAR). Thus, the $PC_{(.)}$ is calculated as:

$$PC_{B-NEMO} = \lambda_{pkt}\left(2\beta + \alpha h_{mh} + \alpha h_{hc}\right), \quad (13)$$

$$PC_{NC-Soto} = \lambda_{pkt}\left(2\beta + \alpha h_{mc} + \alpha h_{lc}\right), \quad (14)$$

$$PC_{HD-NEMO} = PC_{NC-Do} = \lambda_p\left(2\beta + \alpha h_{cn}\right) + \lambda_p\left(\overline{N}_p - 1\right)\left(2\beta + \alpha h_{ac} + \alpha h_{cn}\right), \quad (15)$$

$$PC_{H-NEMO} = \frac{\overline{N}_m}{\overline{N}_p}PC_{DMM} + \frac{\left(\overline{N}_p - \overline{N}_m\right)}{\overline{N}_p}PC_{PMIP}, \quad (16)$$

where

$$PC_{DMM} = \lambda_p\left(2\beta + \alpha h_{cn}\right) + \lambda_p\left(\overline{N}_m - 1\right)\left(2\beta + \alpha h_{ac} + \alpha h_{cn}\right),$$
$$PC_{PMIP} = \lambda_{pkt}\left(2\beta + \alpha h_{mc} + \alpha h_{lc}\right).$$

4) Handover Latency: Since in DMM environment, a mobile node, after handover, can use the new prefix to start a new flow while using the old one for the on-going flows. As a result, we consider the handover latency and the service disruption time. The former ($HO_{(.)}$) is the time needed for the MNN, after handover, to start a new flow, while the latter ($SD_{(.)}$) represents a period when the MNN cannot receive/send the packet of the on-going flows. The service disruption is calculated from the moment the MNN/MR leaves the previous MAR/MAG until the moment the MNN continues receiving the packet of the on-going flow from the CN (for the sake of simplicity, only the flow with the downlink direction is considered). Since the delay between two nodes depends on the bandwidth, the propagation delay and the distance between them, for simplicity, we suppose that the delay is proportional to the distance. The proportion is τ for wired link and κ for wireless link. Thus, we have:

$$HO_{B-NEMO} = t_{L2} + 5\kappa + 2\tau h_{mh}, \quad (17)$$

$$HO_{NC-Soto} = t_{L2} + 5\kappa + 4\tau h_{mc}, \quad (18)$$

$$HO_{HD-NEMO} = t_{L2} + 5\kappa + 2\tau h_{ac}, \quad (19)$$

$$HO_{NC-Do} = t_{L2} + 3\kappa + 2\tau h_{mc}, \quad (20)$$

$$HO_{H-NEMO} = t_{L2} + 3\kappa + 2\tau h_{mc}, \quad (21)$$

where t_{L2} is the layer 2 handover duration.

Regarding the service disruption time, it is calculated as:

$$SD_{B-NEMO} = HO_{B-NEMO} + \tau h_{mh} + 2\kappa, \quad (22)$$

$$SD_{NC-Soto} = HO_{NC-Soto} + \tau h_{mc} + 2\kappa, \quad (23)$$

$$SD_{HD-NEMO} = HO_{HD-NEMO} + \tau h_{ac} + 2\kappa, \quad (24)$$

$$SD_{NC-Do} = HO_{NC-Do} + 3\tau h_{ac} + 2\kappa. \quad (25)$$

$$SD_{H-NEMO} = \frac{\overline{N}_m}{\overline{N}_p}SD_{DMM} + \frac{\left(\overline{N}_p - \overline{N}_m\right)}{\overline{N}_p}SD_{PMIP}, \quad (26)$$

where

$$SD_{DMM} = HO_{H-NEMO} + \tau h_{mc} + \tau h_{ac} + 2\kappa,$$
$$SD_{PMIP} = HO_{H-NEMO} + \tau h_{mc} + 2\kappa.$$

5) End-to-End delay: End-to-end delay ($E2E(.)$) is the packet transmission delay from an MNN to a CN. In a distributed-based solution, the new traffic is routed directly from the current MAR while the old traffic is routed via the anchor one. We have:

$$E2E_{B-NEMO} = 2\kappa + \tau h_{mh} + \tau h_{hc}, \quad (27)$$

$$E2E_{NC-Soto} = 2\kappa + \tau h_{mc} + \tau h_{lc}, \quad (28)$$

$$E2E_{HD-NEMO} = \frac{1}{\overline{N}_p}E2E_{DMM}^{new} + \frac{\overline{N}_p - 1}{\overline{N}_p}E2E_{DMM}^{old}, \quad (29)$$

where

$$E2E_{DMM}^{new} = 2\kappa + \tau h_{cn},$$
$$E2E_{DMM}^{old} = 2\kappa + \tau h_{ac} + \tau h_{cn},$$
$$E2E_{NC-Do} = E2E_{HD-NEMO}. \quad (30)$$

Similar to the packet delivery cost, $E2E_{(.)}$ in case of H-NEMO is given by:

$$E2E_{H-NEMO} = \frac{\overline{N}_m}{\overline{N}_p}E2E_{DMM} + \frac{\left(\overline{N}_p - \overline{N}_m\right)}{\overline{N}_p}E2E_{PMIP}, \quad (31)$$

where

$$E2E_{DMM} = \frac{1}{\overline{N}_m}E2E_{DMM}^{new} + \frac{\overline{N}_m - 1}{\overline{N}_m}E2E_{DMM}^{old},$$
$$E2E_{PMIP} = E2E_{NC-Soto}.$$

6) Nested NEMO: Now we investigate the nested NEMO scenario with n-nesting levels and m associated MNNs. The performance metrics, as similar to the previous paragraphs, are given by:

$$C_{B-NEMO}^{u}(n) = (2n + m + 2)\beta + (2n + 2)\beta + 2\alpha h_{mh},$$
$$C_{HD-NEMO}^{u}(n) = (2n + m + 2)\beta + (2n + 2)\overline{N}_p\beta + 2\overline{N}_p\alpha h_{ac},$$
$$C_{H-NEMO}^{u}(n) = (2n + m + 2)\beta + 2\alpha h_{mc} + 2\overline{N}_m\alpha h_{mc},$$
$$C_{B-NEMO}^{r}(n) = R_B\left(2\beta + 2n\beta + 2\alpha h_{mh}\right),$$
$$C_{HD-NEMO}^{r}(n) = R_B\left[2(n+1)\overline{N}_p\beta + 2(\overline{N}_p - 1)\alpha h_{ac}\right],$$
$$C_{H-NEMO}^{r}(n) = 2R_B\alpha h_{mc}.$$
$$PC_{B-NEMO}(n) = \lambda_{pkt}\left(2\beta + n\beta + \alpha h_{mh} + n\alpha h_{hh} + \alpha h_{hc}\right),$$

$$PC_{HD-NEMO}(n) = \lambda_p\left(2\beta + n\beta + \alpha h_{cn}\right) + \lambda_p\left(\overline{N}_p - 1\right)\left(2\beta + n\beta + n\alpha h_{ac} + \alpha h_{cn}\right),$$

$$PC_{H-NEMO}(n) = \frac{\overline{N}_m}{\overline{N}_p}\lambda_p[2\beta + n\beta + \alpha h_{cn} + \left(\overline{N}_m - 1\right)(2\beta + n\beta + \alpha h_{ac} + \alpha h_{cn})] + \frac{\left(\overline{N}_p - \overline{N}_m\right)}{\overline{N}_p}\lambda_{pkt}\left(2\beta + n\beta + \alpha h_{mc} + \alpha h_{lc}\right),$$

$$E2E_{B-NEMO}(n) = (2 + n)\kappa + \tau h_{mh} + n\tau h_{hh} + \tau h_{hc},$$

$$E2E_{HD-NEMO}(n) = \frac{1}{\overline{N}_p}(2\kappa + n\kappa + \tau h_{cn}) + \frac{\overline{N}_p - 1}{\overline{N}_p}(2\kappa + n\kappa + n\tau h_{ac} + \tau h_{cn}),$$

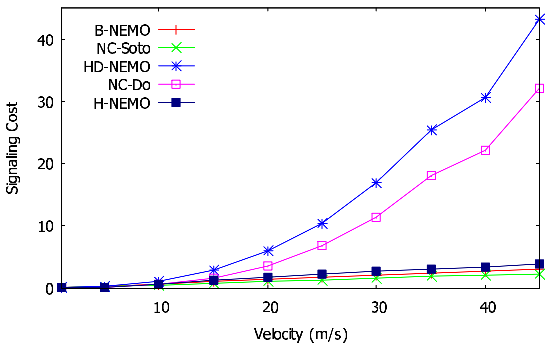


Fig. 8: Signaling cost as a function of velocity (v).

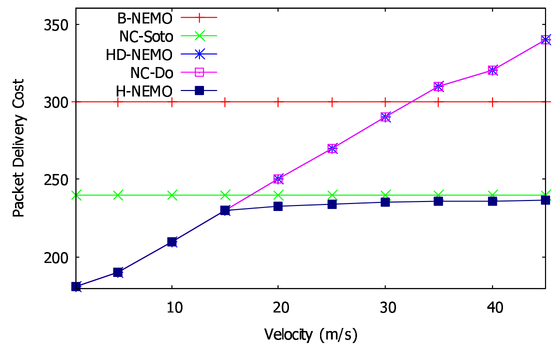


Fig. 9: Packet delivery cost as a function of velocity (v).

$$E2E_{H-NEMO}(n) = \frac{1}{\overline{N}_p}[2\kappa + n\kappa + \tau h_{cn} + (\overline{N}_m - 1)$$
$$(2\kappa + n\kappa + n\tau h_{ac} + \tau h_{cn})] + \frac{(\overline{N}_p - \overline{N}_m)}{\overline{N}_p}(2\kappa + n\kappa + \tau h_{mc} + \tau h_{lc}).$$

C. Numerical Results

In this paper, we consider the case where the MN always moves from MAR/MAG to MAR/MAG as if they were linearly deployed (the user is moving further away from the first attached MAR/MAG and never attaches back to a previously visited MAR/MAG). This assumption, while has no impact on the centralized-based approaches, represents the worst-case scenario for distributed-based ones. Hence, we have [6]: $h_{ac} = \overline{N}_p h_{mm}$, $h_{mh} = h_{mc} + h_{hl}$, and $h_{cn} = h_{lc} + h_{mc}$. The default parameter values for the analysis are introduced in Table II in which some parameters are taken from [6].

TABLE II: Parameters for the performance analysis.

Parameter	Value	Parameter	Value	Parameter	Value
α	1	β	5	m	1
τ	2	κ	15	N_0	6
h_{mm}	1 hop	h_{mc}, h_{hl}	8 hops	h_{lc}	6 hops
h_{hc}	4 hops	T_{BCE}	300s	λ_{pkt}	10
t_{L2}	50 ms	$1/\delta$	300s	R	600 m

Fig. 8 shows the signaling cost as a function of velocity (v). We can observe that the signaling cost is significantly increased in the distributed-based scheme (HD-NEMO and NC-Do) while slightly increased in our proposed and the centralized-based as well (B-NEMO and NC-Soto). The cost in H-NEMO is slightly higher than that in the centralized-based schemes. The reason is that when the velocity is increased the value of $\overline{N}_p$ is expected to be increased. As a result, an additional cost is required for updating the active prefixes in the distributed-based scheme. Our solution (H-NEMO), in this case, helps prevent the significant increase in terms of signaling cost of the distributed-based compared to the centralized-based schemes as the velocity increases.

Fig. 9 illustrates the packet delivery cost when the velocity is varying. It appears clearly that when the value of v is small, the distributed-based approaches (including H-NEMO) outperform the centralized-based ones. The cost for the centralized-based approaches is fixed since the packet is routed via the route MN-MAG-LMA(HA)-CN and the distance between the LMA/HA and different MAGs (between LMA/HA and CN, as well) is supposed to be constant. On the contrary, the cost in the distributed approaches increases as the value of v increases. The reason is that the packet is delivered via the tunnel between the cMAR and aMAR, which is proportional to the value of v. In all the cases, H-NEMO gives a better performance compared to the others.

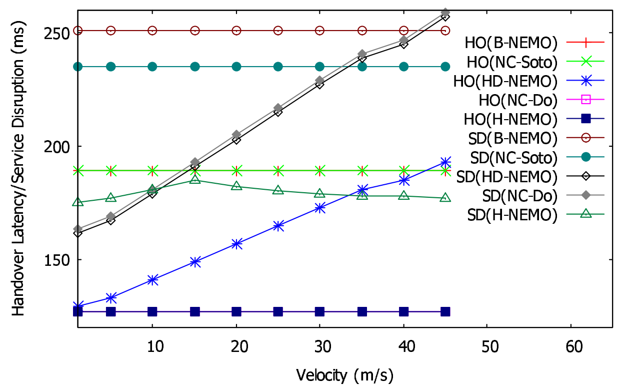
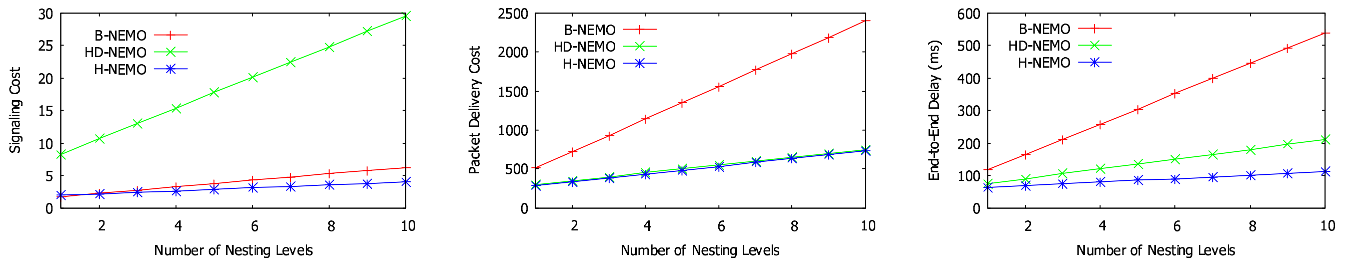


Fig. 10: Handover latency and service disruption time as a function of velocity.

Fig. 10 shows the handover latency and service disruption time as a function of v. Regarding the handover latency, it is greatly increased in case of HD-NEMO while kept constant in the others. It is because unlike the other cases the handover latency in HD-NEMO depends on the delay between cMAR and aMAR. As we can observe, the handover latency in H-NEMO (and NC-Do) is really lower than that in the others. That means, in general, DMM helps the MN quickly obtain a new address to start a new flow after handover. Concerning the service disruption, when v is small, the distributed-based schemes (including H-NEMO) are much better than the centralized-based ones. As v increases, the service disruption in the distributed-based schemes is rapidly increased, thus, becoming notably higher than that in H-NEMO. The service disruption in the centralized-schemes, in both cases, is quite high. In conclusion, H-NEMO helps mitigate the impact of the increase of the velocity to the service disruption (at a minor additional cost compared to the distributed-based approaches when v is small).

Fig. 11 depicts the end-to-end delay as a function of v. Again, when v is small, the distributed-based schemes including our solution are better than the centralized-based ones. When v increases, the end-to-end delay rapidly increases in the distributed-based approaches while keeping the same in the centralized-based ones. As a result, the delay in distributed-based becomes higher than that in centralized-based approaches. However, in our solution, the increase is prevented.



(a) Signaling cost (b) Packet delivery cost (c) End-to-End delay

Fig. 12: Performance metrics as a function of number of nesting levels.

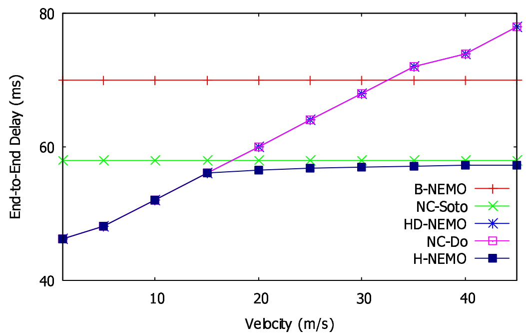


Fig. 11: End-to-End delay as a function of velocity (v).

It is worthy to notice that the similar results can be obtained by fixing the value of the velocity while varying the value of $1/\delta$. This scenario illustrates the effect of the flow duration to the signaling cost, packet delivery cost, handover latency, and end-to-end delay.

Finally, the nested NEMO is investigated. Fig. 12 shows the signaling cost, packet delivery cost, and end-to-end delay as a function of nesting levels (n) while the velocity is set to a value of 20 m/s. As can be seen in this figure, when the number of nesting levels increases, the signaling cost in HD-NEMO is quickly increased since an additional cost is required to update the active prefixes, especially via the wireless link. On the other hand, the packet delivery cost in B-NEMO is notably higher than that in HD-NEMO and H-NEMO. The difference is greatly increased as the value of n increases because B-NEMO suffers from the pinball routing problem. In conclusion, H-NEMO always gives a better performance in comparison with B-NEMO and HD-NEMO (In addition, the metrics in H-NEMO are slightly increased as n increases).

V. Conclusion

Driven by the fact that DMM may not be a suitable scheme for vehicles, this paper proposes a hybrid centralized-distributed architecture for NEMO. This solution inherits the advantages of DMM while mitigating its drawbacks in the situation when a node is running a long-lived flow and/or a node with high mobility. The numerical results showed that H-NEMO generally offers a better performance in terms of signaling cost, packet delivery cost, handover latency (and service disruption time), and end-to-end delay compared with the current proposals for NEMO (including both centralized-based and distributed-based schemes). In the next step, more performance metrics will be considered regarding the impact of the number of nodes and flows. Also, to achieve the realistic results, experiments will be conducted based on a near-to-real testbed as described in [6].

References

[1] Gartner, Inc: http://www.gartner.com/newsroom/id/2791017, Jul 2014.

[2] Ericsson, "Ericsson Mobility Report: On the Pulse of the Networked Society", Nov 2014.

[3] Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2014-2019", Feb 2015.

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

[5] H.A. Chan, D. Liu, P. Seite, H. Yokota, and J. Korhonen, "Requirements for Distributed Mobility Management", RFC 7333, Aug 2014.

[6] T.-T. Nguyen, "Optimization of Mobility Mechanisms for IP-based Multicast Flows", Ph.D. thesis, EURECOM, Biot, France, May 2014.

[7] F. Giust, C.J. Bernardos, A. De La Oliva, "Analytic Evaluation and Experimental Validation of a Network-based IPv6 DMM Solution", IEEE Transaction on Mobile Computing, no. 99, Feb 2014.

[8] J.-H. Lee, J.-M. Bonnin, P. Seite, and H.A. Chan, "Distributed IP Mobility Management from the Perspective of the IETF: Motivations, Requirements, Approaches, Comparison, and Challenges", IEEE Wireless Communications, vol. 20, no. 5, pp. 159-168, Oct 2013.

[9] T.-T. Nguyen and C. Bonnet, "On the Efficiency of Dynamic Multicast Mobility Anchor Selection in DMM: Use Cases and Analysis", in Proc. IEEE ICC, Jun 2014.

[10] I. Soto, C.J. Bernardos, M. Calderon, A. Banchs, and A. Azcorra, "Nemo-enabled Localized Mobility Support for Internet Access in Automotive Scenarios", IEEE Communications Magazine, pp. 152-159, 2009.

[11] S. Jeon and Y. Kim, "Cost-efficient Network Mobility Scheme over Proxy Mobile IPv6 Network", IET Communications, vol. 5, no. 18, pp. 2656-2661, 2012.

[12] P. Sornlertlamvanich, S. Kamolphiwong, R. Elz, and P. Pongpaibool, "NEMO-based Distributed Mobility Management",in Proc.WAINA, 2012

[13] P. Ernest, H.A. Chan, O. Falowo, L. Magagula, and S. Cespedes, "Network-based Distributed Mobility Management for Network Mobility", in Proc. CCNC, Jan 2014.

[14] T.-X. Do and Y. Kim, "Network Mobility Support in the Distributed Mobility Management", IETF Draft (work-in-progress), Oct 2014.

[15] S. Gundavelli, K. Leung, V. Devarapalli, K. Chowdhury, and B. Patil, "Proxy Mobile IPv6", RFC 5213, Aug 2008.

[16] T.-T. Nguyen and C. Bonnet, "A Hybrid Centralized-Distributed Mobility Management for Supporting Highly Mobile Users", to appear in IEEE ICC, Jun 2015.

[17] C. J. Bernardos, A. de la Oliva, and F. Giust, "A PMIPv6-based Solution for DMM", IETF Draft (work-in-progress), Mar 2015.

[18] S. Figueiredo, C. Guimaraes, R.L. Aguiar, T.-T. Nguyen, L. Yadin, N. Carapeto, and C. Parada, "Broadcasting User Content over Novel Mobile Networks", in Proc. IEEE ICC, Jun 2013.

[19] C. Makaya and P. Samuel, "An Analytical Framework for Performance Evaluation of IPv6-based Mobility Management Protocols", IEEE Trans. Wireless Communications, vol. 7, no. 3, pp. 972-983, Mar 2008.