COEXISTENCE AND INTEGRATION SCENARIOS OF IMS AND NEMO TO HANDLE MOBILITY IN ITS ENVIRONMENTS

Alberto Gordillo and Carolina Pinart
Telefónica I+D - c/ Emilio Vargas 6 - 28043 Madrid - +34913374952 - {albergm, cpg}@tid.es

ABSTRACT

The all-IP network convergence is becoming a reality: cellular, wireless and fixed networks support the Internet Protocol (IP), and access-network-independent IP services are being developed. All-IP should also support seamless handovers because of the increasing mobile behavior of IP nodes. In Intelligent Transport Systems and Services, seamless mobility is a must because the topology and availability of access networks change frequently due to the high mobility of vehicles. This paper explores the Network Mobility protocol and the IP Multimedia Subsystem framework as ways to provide seamless IP mobility in the ITS context, both in separate and in integrated operation.

KEYWORDS NEMO, IMS, ITS, seamless mobility

INTRODUCTION

Mobility is inherent to the environment of Intelligent Transport Systems and Services (ITS), where the mobile nodes are moving vehicles at medium to high speeds. In this environment, the network topology is constantly changing, and many ITS services involve real-time exchange of information, e.g., connected multimedia entertainment, and/or bounded delay, e.g., latency-critical safety and efficiency. Seamless mobility management should guarantee these in such a changing network environment. According to the 3rd Generation Partnership Project (3GPP), the IP Multimedia Subsystem (IMS) is a framework that supports multimedia services independently of the access technology of its users in an all-IP network. Future ITS systems will be enabled with multiple access technologies (types), e.g., IEEE 802.11p (wireless) and 3G (cellular). In this context, vehicles will communicate with other vehicles, with Road-Side Units (RSU) and with the Internet using multiple access types seamlessly, and will exchange multimedia information. Thus, IMS may become an important player here. In parallel, the Internet Engineering Task Force (IETF) is defining IP mobility protocols that may offer additional benefits to handling mobility in the ITS environment. The best candidate for that is the Network Mobility (NEMO) protocol.

This paper explores the coexistence and integration of NEMO and IMS as seamless mobility solutions, with the goal of minimizing the disruption during handovers in ITS environments. This is achieved by combining seamless mobility at the application and network layers, by using IMS and/or NEMO in these layers, respectively. Both IMS and NEMO are being considered and analyzed in ITS-related research projects and fora. The Car to Car Communications Consortium (C2C-CC) proposes NEMO to handle mobility for non-critical safety and infotainment [1]. The European COMeSafety project considers NEMO among possible network-layer mobility extensions in the ITS Station

Reference Architecture [2]. The European CVIS project includes NEMO in its mobile router [3] to maintain Internet access among handovers. Last, the German CoCar project uses IMS for seamless ITS communication and multi-network/multi-device support [4].

THE IP MULTIMEDIA SUBSYSTEM

IMS is a generic framework for delivering IP multimedia services to mobile users in the Next Generation Network. To do so, IMS provides capabilities to ease the development of multimedia services and at the same time, it makes these services independent of the access-type technology of users. As seen in Fig. 1a, IMS separates the signaling and data transmission, preventing user data from passing through the IMS architecture. For this reason, IMS can be seen as a control layer that isolates the service layer from the core network. To grant integration with the Internet, IMS uses the Session Initiation Protocol (SIP) [5] and Session Description Protocol (SDP) [6]. SIP is a signaling protocol for controlling multimedia sessions, which is text-based, allows user mobility (SIP addresses are user-location-independent) and does not transport data. Instead, it manages the establishment, termination and modification of the 'conversation'. IMS uses SIP to signal user registrations and session establishment; IMS users have an associated SIP address. Whenever a user connects to IMS, its location is registered in the core through SIP Register so that other users may contact this user by his SIP address. Analogously, when the user changes his location, he shall send another SIP Register message and a SIP Invite to every user he was communicating with.

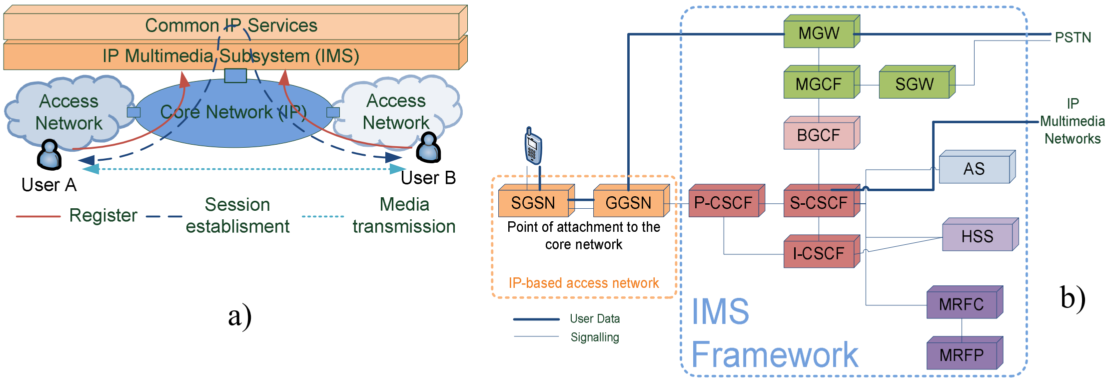

Figure 1 – a) Flows and b) architecture of IMS

The IMS architecture is based on two functionality groups (control and interworking), two servers (home subscriber, HSS module in Fig. 1b, and application, AS in Fig. 1b), and a multimedia resource function. In turn, the control functionality is split into three functions (P-CSCF, I-CSCF and S-CSCF in Fig. 1b), and the interworking functionality is divided in the BGCF, MGCF, MGW and SGW modules in Fig. 1b. These last three modules deal with circuit-switching (CS) communications. Finally, the multimedia resource function is composed of MRFC and MRFP (Fig. 1b). Table 1 briefly describes the features of the above-mentioned modules of the IMS architecture.

In the *IMS registration process*, the user equipment (UE) sends a SIP Register message to its entry point (P-CSCF in Fig. 1b), which forwards it to I-CSCF. This module sends a User Authorization Request (UAR) to the HSS, which checks that the user identification is correct and sends a User Authorization Answer (UAA) back to I-CSCF. Then, I-CSCF sends the registry to the S-CSCF module, which downloads the authentication vectors from HSS to challenge the user. S-CSCF sends the challenge to the UE, which sends back a SIP Register message with the answer. Finally, S-CSCF authenticates the user, downloads the user profile from the HSS module and notifies the UE that the user has been successfully registered. This process is illustrated in Fig. 2.

Table 1 – Main modules of the IMS architecture

	Module	**Description**
Control	Proxy Call Session Control Function (P-CSCF)	Entry point to IMS, coming from the IP-based access network of the user. P-CSCF acts as a proxy, routing the registry messages to the right I-CSCF, and applying local policies to the sessions.
	Interrogating Call Session Control Function (I-CSCF)	SIP proxy server, hiding the network topology to outer modules. It performs the initial authorization and assigns it to the user S-CSCF.
	Serving Call Session Control Function (S-CSCF)	S-CSCF is always in the user home network, and is responsible for session control, authentication and request of media resources.
Interworking	Breakout Gateway Control Function (BGCF)	Gateway selection for a CS domain communication.
	Media Gateway Control Function (MGCF)	Translates SIP messages to its equivalents in a CS domain.
	Media Gateway (MGW)	Translates multimedia IP data to a CS domain.
	Transport Signaling Gateway (SGW)	Also called T-SGW, this module interconnects different signaling networks, without interpreting application-layer messages.
Servers	Home Subscriber Server (HSS)	User database, storing user information like subscription profiles, authentication, status and location.
	Application Server (AS)	IMS nodes holding the multimedia services.
Multimedia	Multimedia Resource Function Controller (MRFC)	SIP user agent.
	Multimedia Resource Function processor (MRFP)	Provides media mixing and processing.

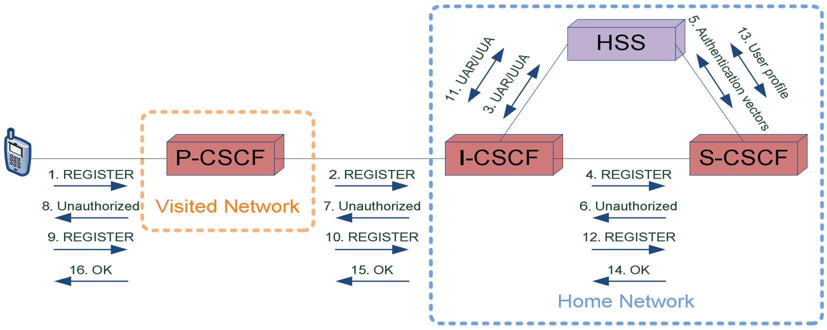

Figure 2 - IMS registration process

To *establish an IMS session*, the UE sends a SIP Invite message to the P-CSCF module (Fig. 1b) with the SDP media description for that session. P-CSCF forwards it to the I-CSCF module, which forwards the message again to the next hop (S-CSCF). Then, the user-assigned S-CSCF determines if the message has to be sent to an AS to implement the service; the S-CSCF receives the message and forwards it to the P-CSCF module of the destination user. The UE receives the SIP Invite message and replies with the SDP negotiation parameters. The communication begins when the caller and called users agree with the communication media capabilities. This is depicted in Fig. 3 for both the caller and called home networks.

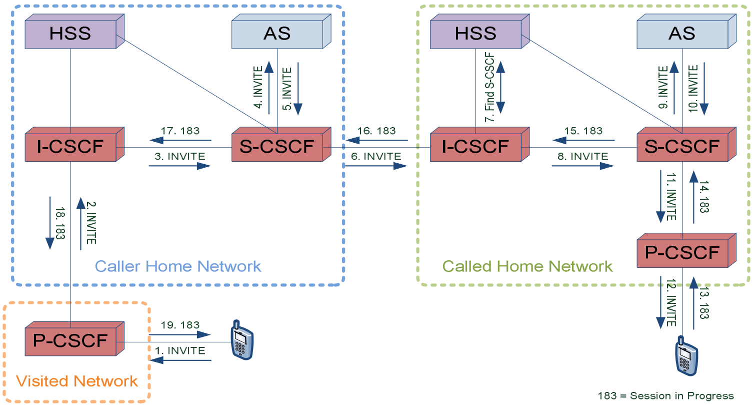

Figure 3 – IMS session establishment process

THE NETWORK MOBILITY PROTOCOL

An IP address has two functions: identifying an end point of a communication, and specifying how to reach the device that is using the address (i.e., the route to the destination). These functions limit the mobility of the IP address because they bind it to a specific route. Therefore, when the device associated to the IP address moves to another location, the IP address is no longer topologically correct and hence the connectivity is lost. The solution to this disruption involves obtaining a new address that is topologically valid in the new location of the device. However, this solution does not consider the mobility problem, because the connections that were established before the connectivity loss were associated to the ‘old’ IP address, that is, they are lost.

The IETF has proposed solutions to this mobility problem both for IP addresses and networks. Mobility Support in IPv6 (MIPv6) [7] targets the former, while the NEMO Basic Support protocol [8], which is based on MIPv6, targets the latter. Fig. 4 illustrates the operation of NEMO to handle mobility, which involves the following elements: the Mobile Router (MR), which is a router that can change its location among networks while maintaining reachability. The Mobile Network Node (MNN) is a node, fixed or mobile, of the Mobile Network (MN), which has an IP address of the block assigned to the MN. The Home Address (HoA) is a global unicast IP address topologically valid in the MR home network that is assigned to the MR. The Care-of Address (CoA) is a global unicast IP address assigned to the MR that is topologically valid in the network where it is currently attached (Visited Network in Fig. 4). The Home Agent (HA) is a router on the MR home network that represents the MR when out of its home (Fig. 4).

When the MR reaches a visited network, it obtains a CoA and notifies its new IP address to the HA. The HA is situated in the MR home network, where its address is topologically correct. As seen in Fig. 4, all data sent to the mobile network are received by the HA which forwards it to MR’s CoA by an IPv6-to-IPv6 tunnel. The MR erases the tunnel headers and forwards the data to the MNN. In the case of MN-originated data, the MR acts as a gateway: it receives data and sends it through the tunnel to the HA, which forwards the data to its destination. With NEMO, the MR hides the mobility of the network to the MNNs and, thanks to keeping its IP address, the communications of the mobile network do not need to be reestablished between location changes. These features of NEMO are considered to be important advantages for ITS; they are recommended in the standard for Communications Architecture for Land Mobile environment (CALM) [9]. However, NEMO also presents some drawbacks. Due to the absence of route optimization in NEMO Basic Support, suboptimal paths are produced in the sense that the traffic needs to cross the HA even if there is a shorter route. The HA-MR tunneling (see Fig. 4) adds a 40-byte header overhead to every packet.

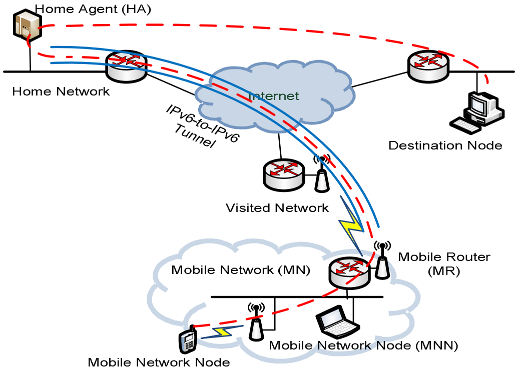

Figure 4 – Mobility management in NEMO

COEXISTENCE AND INTEGRATION SCENARIOS FOR IMS AND NEMO

We use the term ‘coexistence’ when referring to the operation of NEMO and IMS without interactions between them, and ‘integration’ when there are interactions. In the coexistence scenario, we believe that the best approach for the ITS environment is to use IMS as the mobility support mechanism for IMS-capable services, and to use NEMO otherwise. This way all the ITS services can be enabled with seamless mobility, and IMS-capable services can enjoy enhanced functionality on top of that. The integration scenario for IMS and NEMO offers a key benefit: the reduction in the amount of reestablishments, i.e, *the reduction in the delay of communications in the event of mobility management*. The rationale for this is that NEMO provides mobility at the network layer, not at the application layer, which eliminates the need for session reestablishments between handovers, i.e., reduces the handover delay of IMS sessions. Table 2 summarizes the key aspects of these coexistence and integration scenarios.

The interactions between NEMO and IMS in the integration scenario are similar, although more restricted, to those between MIPv6 and IMS, which are analyzed in [10]. That is, the mobile node uses its home address as the IP address in the SIP registration of IMS. Without this process, if a device changed the access technology, it would also change its IP address. Under IMS, this device would have to re-register and send SIP Invite messages to the nodes implied in the communication before changing the access, and then continue the communication, which would produce disruptions. Thanks to NEMO (or MIPv6), the devices of the Mobile Network (MN) maintain their IP address in the event of an access technology change, which in turn maintains the IP address/SIP address association and hence avoids the need for IMS re-register and re-invites, keeping the signaling messages down to a CoA update from the MN to the HA (Fig. 4).

However, the integration scenario presents a drawback: the presence of suboptimal paths due to the NEMO requirement of always passing through the HA, which may result in the HA (Fig. 4) managing large amounts of traffic and also in an increase in the end-to-end communication delay after the handover. This problem is solved in MIPv6 using route optimization, which prevents signaling and data flows from following suboptimal paths [7]. However, the absence of such optimization mechanisms in the NEMO Basic Support [8] results in the data flow always passing through the HA-MR tunnel. Moreover, due to IMS security requirements, the same IP address must be used in the SIP messages for the source and SIP-internal addresses. Therefore, the same IP address is required for SIP signaling and data flow; if NEMO is used together with IMS (as in the integration scenario), both SIP signaling and data must pass through the HA.

Table 2 - Summary of IMS-NEMO mobility management

	IMS (Mobility)	**NEMO**
Benefit	Direct data traffic	Maintains ongoing sessions
Drawback	Needs session reestablishment	Suboptimal routing
Layer	Application	Network
ITS applications	Network enhancement for ITS safety services (CoCar project)	Non-critical safety and infotainment applications (explored in C2C-CC, and in the CVIS and COMeSafety projects)
Coexistence scenario	IMS-capable applications	Non-IMS capable applications
Integration scenario	User mobility and QoS negotiation	Handover management

PERFORMANCE ANALYSIS

This section provides an analytical comparison of the NEMO-IMS coexistence and integration scenarios described in the previous section. The performance of these scenarios is based on the signaling cost of registration, session establishment and handover, for which the figures of merit considered are the delay (*D*) and overhead cost (*C*), where *D* is defined as the end-to-end delay of the signaling process under analysis, and *C* is the length of the messages required multiplied by the number of hops of the intended communication. $D(_{a\text{-}b}M)$ denotes the delay of the message *M* from the source *a* to the destination *b*, while and $L(_{a\text{-}b}M)$ refers to the length of the message *M*.

Coexistence scenario

NEMO registration costs are $D_{NEMO-HO} = D(_{MR-AR}Rsol) + D(_{AR-MR}Rad) + D(_{MR-HA}BU) + D(_{HA-MR}BU)$ and $C_{NEMO-HO} = L(_{MR-AR}Rsol) + L(_{AR-MR}Rad) + N_{MR-HA}(L(_{MR-HA}BU) + L(_{HA-MR}BU))$, considering the signaling needed to obtain an IP address from the access router (CoA) and to match the CoA to the home addredd (HoA), as illustrated in Fig. 5. The signaling for the NEMO registration process is the same as for a *NEMO handover* process.

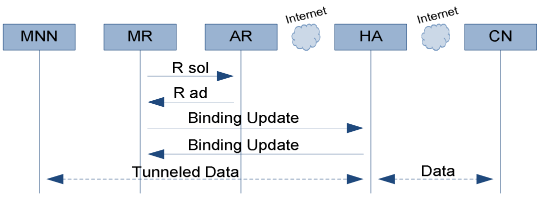

Figure 5 - NEMO flow for registration or handover

IMS registration costs are calculated as $C_{IMS-\mathrm{Reg}} = N_{MN-SCSCF}*(L_{REG1} + L_{401} + L_{REG2} + L_{200})$ and $D_{IMS-\mathrm{Reg}} = D(_{MN-PCSCF}REG1) + D(_{PCSCF-SCSCF}REG1) + D(_{SCSCF-PCSCF}401) + D(_{PCSCF-MN}401) + D(_{MN-PCSCF}REG2) + D(_{PCSCF-SCSCF}REG2) + D(_{SCSCF-PCSCF}200) + D(_{PCSCF-MN}200)$ (see Fig. 6 left), where IMS uses SIP for signaling. Since SIP is a text-based protocol, SIP messages are easily readable by humans but they are more inefficient than binary encoded messages, which results in the cost of SIP messages being higher than the cost of NEMO's. The costs for an *IMS session establishment*, including coding negotiation and QoS reservation (Fig. 6 right), are $C_{IMS-Ses} = N_{MN-CN}*(L_{INV} + L_{183} + L_{PRA} + L_{UPD} + L_{200}*2 + L_{RIN} + L_{ACK})$ and

$D_{IMS-Ses} = D\left({}_{MN-PCSCF}INV\right) + D\left({}_{PCSCF-CN}INV\right) + D\left({}_{CN-PCSCF}183\right) + D\left({}_{PCSCF-MN}183\right) + D\left({}_{MN-PCSCF}PRA\right) + D\left({}_{PCSCF-CN}PRACK\right)$ $+2*\left(D\left({}_{CN-PCSCF}200\right) + D\left({}_{PCSCF-MN}200\right)\right) + D\left({}_{MN-PCSCF}UPD\right) + D\left({}_{PCSCF-CN}UPD\right) + D\left({}_{CN-PCSCF}RIN\right) + D\left({}_{PCSCF-MN}RIN\right)$ $+D\left({}_{MN-PCSCF}ACK\right) + D\left({}_{PCSCF-CN}ACK\right) = DA\left({}_{MN\leftrightarrow PCSCF}X\right) + DA\left({}_{PCSCF\leftrightarrow CN}X\right)$. Note that the signaling delay contains $DA({}_{a\leftrightarrow b}x)$, which is the sum of all IMS signaling messages between *a* and *b*. The session establishment delay can be split in two parts; up to and from the P-CSCF module (Fig. 1b): $DA({}_{MN\leftrightarrow P\text{-}CSCF}x)$ and $DA({}_{P\text{-}CSCF\leftrightarrow CN}x)$ in the $D_{IMS\text{-}Ses}$ expression above.

The *handover in IMS* implies the process of re-register with its S-CSCF module and a re-invite for each active session before the handover. In the case of a mobile network, it implies the re-register and the re-invite of all the IMS clients and sessions in the network, and is represented as N_c (number of mobile IMS clients) and N_s (number of active IMS sessions). The handover delay of IMS corresponds to the maximum delay of all re-registers and re-invites. $\mathrm{D}_{\mathrm{IMS\text{-}Reg}}CliX$ is the delay of IMS registration of client *X* and $\mathrm{D}_{\mathrm{IMS\text{-}Ses}}SesXY$ to the delay of IMS session establishment between clients *X* and *Y*. The *handover cost* is therefore the sum of the cost of all re-registers and re-invites: $D_{IMS-HO} = \underset{D}{Max}\left(D_{IMS-\mathrm{Reg}}CliX + D_{IMS-Ses}SesXY\right)$ and $C_{IMS-HO} = N_c * C_{IMS-\mathrm{Reg}} + N_S * C_{IMS-Ses}$.

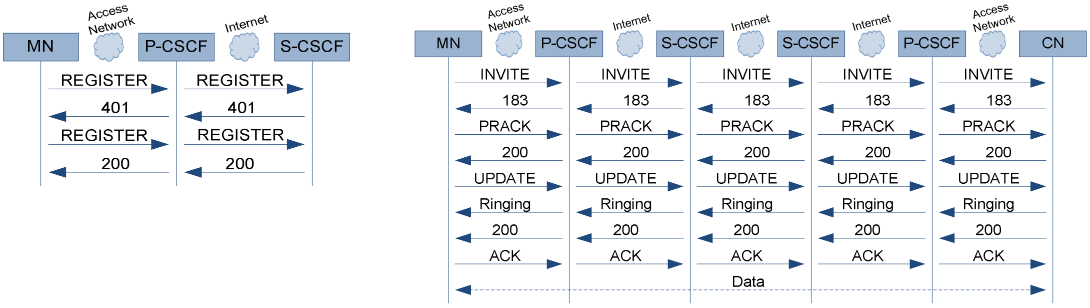

Figure 6 – Flow for IMS registration (left) and session establishment (right)

Integration scenario

In the *NEMO-IMS* integrated solution, the signaling cost involves the registration of the NEMO HoA as the IP of the IMS user identity, as described in the first section. Due to the NEMO MR-HA tunnel, the SIP signaling follows a suboptimal path (Fig. 7). Due to this, the *cost of registration, session establishment and handover* is the sum of the NEMO registration delay and the IMS registration delay, the latter is the maximum delay of the IMS clients in the mobile network, of the previous formulas but considering that the IMS messages follow the suboptimal path, i.e., MR – HA – P-CSCF. For that reason, the *delay of IMS registration* is modified by subtracting the messages between MN and PCSCF, i.e., the optimal route, and adding the messages of the suboptimal route, i.e., between MR and HA and between HA and P-CSCF. Hence, costs are: $D_{NEMOIMS-\mathrm{Reg}} = \underset{D}{Max}\left(D_{IMS-\mathrm{Reg}}CliY - DA\left({}_{MN\leftrightarrow PCSCF}X\right)CliY + DA\left({}_{MN\leftrightarrow HA}X\right)CliY + DA\left({}_{HA\leftrightarrow PCSCF}X\right)CliY\right) + D_{NEMO-HO}$ and $C_{NEMOIMS-\mathrm{Reg}} = C_{NEMO-HO} + N_c * \left(N_{MN-HA} + N_{HA-PCSCF} + N_{PCSCF-SCSCF}\right) * \left(2 * L_{REG} + L_{401} + L_{200}\right)$. Therefore, the cost is the sum of the NEMO Registration and the IMS Registration, with a modification in the number of hops to take into account the suboptimal path, that is: $D_{NEMOIMS-Ses} = D_{IMS-Ses} - DA\left({}_{MN-PCSCF}X\right) - DA\left({}_{PCSCF-MN}X\right) + DA\left({}_{MN-HA}X\right) + DA\left({}_{HA-PCSCF}X\right) + DA\left({}_{PCSCF-HA}X\right) + DA\left({}_{HA-MN}X\right)$ and $C_{NEMOIMS-Ses} = \left(N_{MN-HA} + N_{HA-PCSCF} + N_{PCSCF-SCSCF} + N_{SCSCF-CN}\right) * \left(L_{INV} + L_{183} + L_{PRA} + L_{UPD} + L_{200} * 2 + L_{RIN} + L_{ACK}\right)$. Similarly, the delay and cost of the session establishment are the delay and cost of IMS following the suboptimal path: $D_{NEMOIMS-HO} = D_{NEMO-HO}$ and $C_{NEMOIMS-HO} = C_{NEMO-HO}$. The handover cost is that of NEMO; thanks to the network-layer mobility, IMS agents keep their IP addresses from the NEMO network prefix, while maintaining their sessions.

We can see that NEMO forces MN-PCSCF signaling to pass through the HA, which increases the delay and the signaling cost of the integrated solution. A workaround proposed in [11] is that MIPv6 or NEMO is used for long-live TCP connections and SIP for real-time communications over UDP, although this would require that the mobile hosts identify when to use its HoA or its CoA. Nevertheless, it improves the handover latency and cost by reducing the signaling to the same as the NEMO handover, and makes it independent of the number of ongoing communications. Another side problem is that NEMO signaling does not negotiate new QoS parameters, unlike SIP signaling.

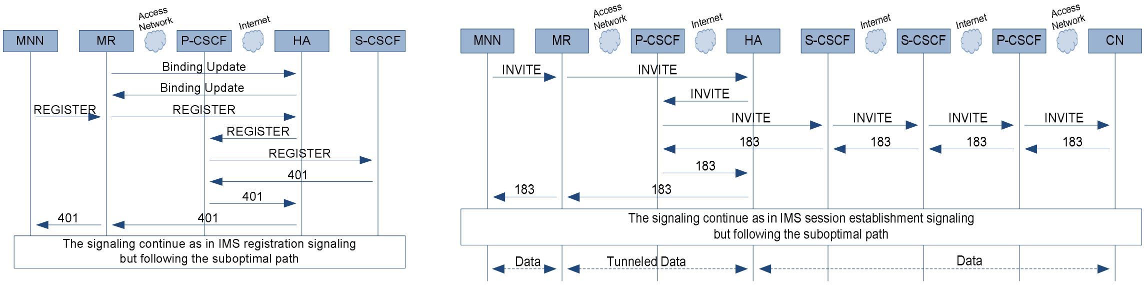

Figure 7 - Registration (left) and session establishment (right) in NEMO-IMS

CONCLUSIONS

Integrating NEMO and IMS to handle mobility in ITS environments improves the handover latency and associated cost (*D* and *C*) by reducing the IMS signaling to the NEMO handover signaling, making it in turn independent of the IMS clients that are in the ITS network and also of the number of ongoing communications, and without introducing overhead in the data plane. The price to pay for these advantages is twofold: unlike IMS-SIP, NEMO does not negotiate new QoS parameters. On the other hand, traffic passes through the HA, which may increase the end-to-end delay and decrease the scalability of NEMO-IMS. Although small networks and delay-tolerant ITS services would not notice the effects of the first drawback, a solution would encompass the introduction of QoS-awareness in NEMO. Solving the second drawback involves route optimization, as in MIPv6. Coexistence scenarios of IMS and NEMO are also possible: our recommendation is that IMS-capable ITS services use IMS mobility management mechanisms, and that the remaining ITS services use NEMO for seamless mobility.

REFERENCES

[1] R. Baldessari, A. Festag, M. Lenardi, "C2C-C Consortium Requirements for Usage of NEMO in VANETs", IETF Internet Draft Version 00, 2007.
[2] COMeSafety, “European ITS Communication Architecture. Overall Framework. Proof of Concept Implementation”, October 2008.
[3] K. Evensen, "CVIS Communications", CVIS Event, December 2008.
[4] Y. Zang, *et al.*, "Towards a European Solution for Networked Cars-Integration of Car-to-Car Technology into Cellular Systems for Vehicular Communication in Europe", IV Fully Networked Car, 2009 Workshop, March 2009.
[5] J. Rosenberg, *et al.*, "SIP: Session Initiation Protocol", IETF RFC 3261, 2002.
[6] M. Handley, *et al.*, "SDP: Session description protocol", IETF RFC 2327, 1998.
[7] D. Johnson, *et al.*, "Mobility Support in IPv6", IETF RFC 3775, 2004.
[8] V. Devarapalli, *et al.*, "Network Mobility Basic Support Protocol", IETF RFC 3963, 2005.
[9] CALM ISO TC 204 Working Group 16, http://www.isotc204wg16.org/.
[10] S. Faccin, P. Lalwaney, B. Patil, "IP multimedia services: analysis of mobile IP and SIP interactions in 3G networks", in IEEE Communications Magazine, vol. 42, pp. 113-120, 2004.
[11] E. Wedlund H. Schulzrinne, "Mobility support using SIP", in Proc. 2nd ACM International Workshop on Wireless Mobile Multimedia. Seattle (WA-US), 1999.