Enhanced Mobility Management Scheme in PMIP-SAE-Based Mobile Networks

Moneeb Gohar and Jin-Ghoo Choi

***Abstract*—In the existing PMIP scheme for SAE-based mobile networks such as LTE, data delivery is performed by GTP between eNB and S-GW, and also between eNBs in X2-based handover, while GRE protocol is used between S-GW and P-GW. However, the use of GTP tunnels between eNB and S-GW and between eNBs tends to induce large tunneling overhead for packets. Existing scheme also tends to incur large registration delay and route update delay after handover. To overcome the drawbacks, we propose an enhanced mobility management scheme for SAE based mobile networks. In the proposed scheme, we combine two user plane nodes, S-GW, and P-GW, as a single entity, and make it function as LMA. Further, each eNB works as a MAG of PMIP. The GRE tunnel is used between eNB/MAG and S-GW/P-GW/LMA, and also between eNB/MAGs, instead of GTP. Differently from the existing scheme, target eNB/MAG exchanges PBU and PBA messages directly with the S-GW/P-GW/LMA in the route update operation after handover, not through MME. Numerical analysis shows that the proposed scheme reduces all the tunneling overhead, registration delay, and route update delay after handover significantly.**

***Index Terms*—PMIPv6, LTE, SAE, Mobility, Handover.**

I. Introduction

THE System Architecture Evaluation (SAE) is the core network architecture for the 3rd Generation Partnership Project (3GPP) LTE system, supporting high throughput and mobility between heterogeneous access systems. For the IP mobility in SAE, Proxy Mobile IP (PMIP) has been considered in [1], [2], [3], which is denoted as PMIP-SAE in this letter. Therein, Packet Data Network (PDN) Gateway (P-GW) of SAE has been used as Local Mobility Anchor (LMA) of PMIP, and Serving Gateway (S-GW) of SAE has functioned as Mobile Access Gateway (MAG) of PMIP. It employs the GPRS Tunneling Protocol/Generic Routing Encapsulation (GTP/GRE) protocols for data delivery. In the Evolved Packet Core (EPC), each eNodeB (eNB) establishes a GTP tunnel with its S-GW/MAG and the S-GW/MAG establishes a GRE tunnel with the P-GW/LMA. While in X2-based handover, the GTP tunnel is set up between the associated eNBs. This architecture, however, incurs large tunneling overhead in the data delivery due to the GTP tunnels. Each packet shall include the GTP/UDP/IP header between eNB and S-GW/MAG and also between eNBs. Moreover, the PMIP-SAE also tends to induce large registration delay and route update delay after handover, as shown in Fig. 1a and 1b [4].

Manuscript received October 20, 2015; revised December 24, 2015 and January 9, 2016; accepted January 9, 2016. Date of publication January 25, 2016; date of current version June 8, 2016. The associate editor coordinating the review of this paper and approving it for publication was M. A. Imran.

The authors are with the Department of Information and Communication Engineering, Yeungnam University, Gyeongsan 38541, South Korea (e-mail: moneebgohar@yu.ac.kr; Jchoi@yu.ac.kr).

Digital Object Identifier 10.1109/LCOMM.2016.2521645

To overcome such limitations, we propose an enhanced mobility management scheme in PMIP-SAE, denoted by EPMIP-SAE. We first combine S-GW and P-GW as a single entity and call it the S-GW/P-GW. Each eNB works as an MAG and the S-GW/P-GW becomes an LMA. We do not apply the GTP tunnel at all to reduce the overhead. GRE tunnels are used alternatively between eNB/MAG and S-GW/P-GW/LMA and also between eNB/MAGs. In the proposed scheme, Target eNB/MAG exchanges *Proxy Binding Update* (PBU) and *Proxy Binding Acknowledgement* (PBA) messages directly with the S-GW/P-GW/LMA, not through the Mobility Management Entity (MME), since the Target eNB/MAG can obtain the information on the S-GW/P-GW/LMA from the Source eNB/MAG during the handover preparation phase.

II. Existing PMIP-SAE Mobility Scheme

A. Initial Registration

We briefly review the previous IP mobility scheme for SAE networks, PMIP-SAE, studied in [1], [2], [3], [5]. The network model is illustrated in Fig. 2a, where P-GW of SAE is used as an LMA of PMIP, and S-GW of SAE is employed as an MAG of PMIP.

We describe the initial procedure of PMIP-SAE in Fig. 2b [2], [3]. When a User Equipment (UE) establishes a radio link with the eNB, it sends an *Attach Request* message to MME. Then, the security-related procedures are performed between UE and MME. MME updates the associated Home Subscriber Server (HSS) and to establish a transmission path, sends a *Create Session Request* message to the S-GW/MAG. Receiving the request from MME, S-GW/MAG sends a *PBU* message to the P-GW/LMA, and P-GW/LMA responds with a *PBA* message. Subsequently, S-GW/MAG sends a *Create Session Response* message to the MME, and the MME notifies the information received from S-GW/MAG to eNB by the Initial *Context Setup Request* message. This signaling message also contains the Attach Accept notification, which is the response for the initial *Attach Request* message. Now, eNB returns the Initial *Context Setup Response* message to MME, and UE sends an *Attach Complete* message to the MME. To set up the data bearer, MME sends a *Modify Bearer Request* message to S-GW/MAG, and the S-GW/MAG responds with a *Modify Bearer Response*. UE now sends a data packet to P-GW/LMA, and it forwards the packet to users in the Internet.

B. Protocol Stack

Fig. 3 shows the protocol stack for the GTP/GRE-based data delivery in PMIP-SAE networks. The radio access uses the

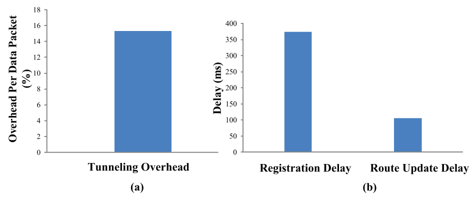


Fig. 1. PMIP-SAE overhead and delay.

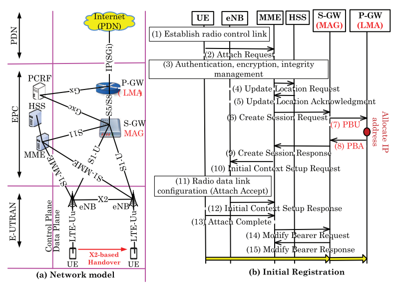


Fig. 2. Existing PMIP-SAE mobility scheme.

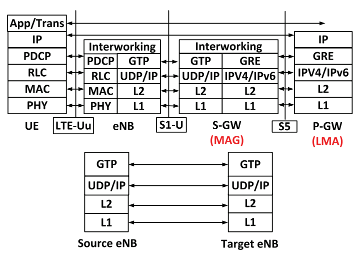


Fig. 3. Protocol stack of PMIP-SAE.

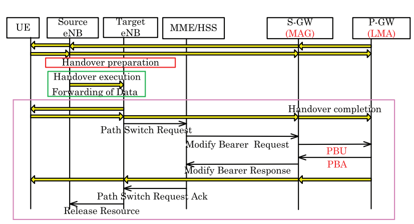


Fig. 4. Route update operation after handover.

Medium Access Control (MAC), Radio Link Control (RLC), and Packet Data Convergence Protocol (PDCP) protocols. GTP is used between eNB and S-GW/MAG and GRE is used between S-GW/MAG and P-GW/LMA. GTP encapsulates the original IP packet into an outer IP packet in between eNB and S-GW/MAG, and GRE encapsulates the original packet into an outer IP packet in between S-GW/MAG and P-GW/LMA.

C. Route Update Operations After Handover

Fig. 4 shows the route update operation after handover of UE from Source eNB to the Target eNB. First, Target eNB sends a *Path Switch Request* message to MME. Then the MME sends a *Modify Bearer Request* message to S-GW/MAG. After receiving the message, S-GW/MAG sends a *PBU* message to P-GW/LMA. P-GW/LMA responds with a *PBA* message to the S-GW/MAG, and S-GW/MAG sends a *Modify Bearer Response* message to the MME. Then, MME sends *Path Switch Request Ack* to Target eNB, and it sends *Release Resource* to Source eNB to complete the handover procedure.

III. Proposed EPMIP-SAE Scheme

A. Network Model

Fig. 5a shows an overview of the proposed enhanced mobility management scheme in PMIP-SAE, denoted by enhanced PMIP-SAE (EPMIP-SAE). In our scheme, each eNB functions as MAG and S-GW/P-GW works as LMA. In the proposed scheme, the S-GW/P-GW can be configured on the same device as a single entity. It is possible for MME to select same S-GW/P-GW, when UE connects from a home network. This combined S-GW/P-GW can implement a single S-GW/P-GW session for the UE. This session acts as an S-GW on the ingress side and P-GW on the egress side. When UE moves to another eNB region in different S-GW area in the same network or a different network, the MME will select another S-GW for the session and thus the S-GW/P-GW session for UE transitions to P-GW session only and the session exposes S5/S8 interfaces to the S-GW. In the Proposed scheme, GTP control plane is used for control packets. Based on UE, each eNB/MAG will decide about the data plane. If UE is PMIP user, then PBU/PBA option is used in initial registration as describe in Section III-C. Based on PBU/PBA, the GRE tunnel is established between eNB/MAGs and S-GW/P-GW/LMA for data plane. If UE is not PMIP user, then PBU/PBA option is not used in initial registration and thus the GTP tunnel is established between eNB/MAG and S-GW/P-GW/LMA for data plane. A UE initially communicates with the Internet user in the Source eNB/MAG domain, and moves to the Target eNB/MAG domain later under the same S-GW.

B. Protocol Stack

Fig. 5b shows the protocol stack for the proposed scheme. The radio access uses the MAC, RLC, and PDCP protocols. GRE tunnel is used between eNB/MAG and S-GW/P-GW/LMA. GRE encapsulates the original IP packet into an outer IP packet in between eNB/MAG and S-GW/P-GW/LMA as well as between Source and Target eNB/MAGs.

C. Initial Registration

Fig. 5c illustrates the initial registration procedure of the proposed scheme. When UE establishes a radio link with eNB, it sends an *Attach Request* message to eNB/MAG. Then, eNB/MAG sends *Attach Request* message incorporating PBU to MME. Then, the security-related procedures are performed between UE and MME. MME updates the associated HSS and to establish the transmission path, sends a *Create Session Request* to S-GW/P-GW/LMA, which incorporates the *PBU* message. When S-GW/P-GW/LMA receives the request from MME, it responds with a *Create Session Response* message containing the PBA. Now, MME sends the information from S-GW/P-GW/LMA to the eNB/MAG with *Initial Context Setup Request* containing the *PBA* message. This signaling message contains the *Attach Accept* notification also, which is the response for the *Attach Request*. Then, eNB/MAG responds with an *Initial Context Setup Response* to MME, and UE sends an *Attach Complete* message to the MME. For data delivery,

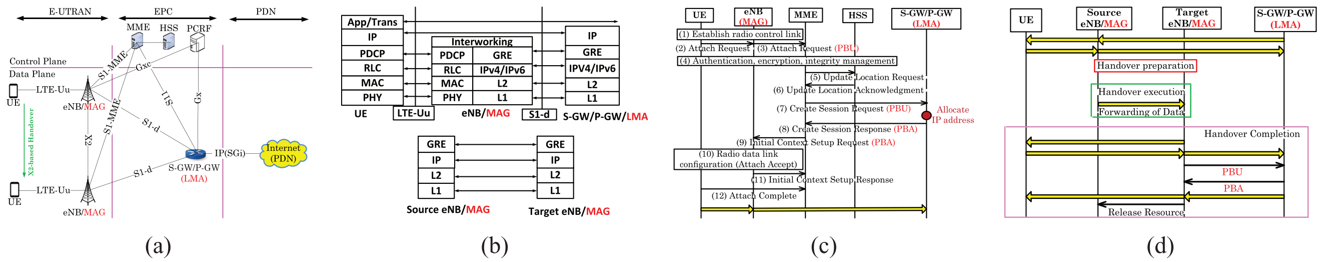


Fig. 5. (a) Network model for EPMIP-SAE, (b) Protocol Stack, (c) Initial Registration, and (d) Route update operations after handover for Proposed Scheme.

UE sends a data packet to eNB/MAG and then to S-GW/P-GW/LMA. The S-GW/P-GW/LMA forwards data packets to the Internet users.

D. Route Update Operations After Handover

We illustrate the route update operation after handover in Fig. 5d. After a UE moves from Source eNB to Target eNB, the Target eNB/MAG exchanges directly the *PBU* and *PBA* messages with S-GW/P-GW/LMA for route optimization. Since, the Target eNB/MAG has obtained the information's on the S-GW/P-GW/LMA and also about tracking area identity (TAI) list for the cluster of cells under the same S-GW area that UE can move around without MME update, from the Source eNB in handover preparation phase. It is noted that UE obtains TAI list when it attaches to an LTE network and the UE sends TAI list to the Source eNB when detects measurement event. When UE moves to another eNB region in different S-GW area in the same network or a different network and the new tracking area is not present in the TAI list, the Target eNB will switch the path through MME and also update location. Then, Target eNB/MAG sends *Release Resource* to Source eNB/MAG to complete the handover procedure.

IV. Performance Analysis

In this section, we analytically compare two mobility schemes, PMIP-SAE and EPMIP-SAE, in terms of the tunneling overhead, registration delay, and route update delay after handover.

A. Analysis Model

We summarize the notations for our analysis in Table 1. We denote by $T_{x-y}(S)$ the transmission delay of a message of size S from node x to node y via a wireless link. Note that $T_{x-y}(S) = [(S/B_{wl}) + L_{wl}]$. Meanwhile, $T_{x-y}(S, H_{x-y})$ denotes the transmission delay of a message of size S from node x to node y via wired links, where H_{x-y} represents the hop count between x and y. It is easy to obtain $T_{x-y}(S, H_{x-y}) = H_{x-y} \times [(S/B_w) + L_w + T_q]$.

B. Tunneling Overhead (TO)

In the existing PMIP-SAE architecture using a GTP tunneling between eNB and S-GW/MAG and between eNBs, a data packet is encapsulated by the 20-bytes IP header, 8-bytes UDP header, and 8-bytes GTP header, which is 36 bytes in total. In the proposed EPMIP-SAE architecture using GRE tunneling

TABLE I
Parameters Used for Analysis

S_c and S_d	Size of control and Data packets (bytes), respectively.
B_w and B_{wl}	Wired and wireless link bandwidth (Mbps), respectively.
L_w and L_{wl}	Wired and wireless link delay (ms), respectively.
H_{a-b}	Hop count between node a and b in the network
T_q	Average queuing delay at each node

between eNB/MAG and S-GW/P-GW/LMA combined node and between eNB/MAGs, each data packet is encapsulated by the 20-bytes IP header and 4 bytes GRE header, which is total 24 bytes. We define the tunneling overhead of PMIP-SAE and EPMIP-SAE as follows.

$$TO_{PMIP-SAE} = \frac{GTP/UDP/IP}{\text{Data Packet Size}(S_d)+GTP/UDP/IP} \times 100$$

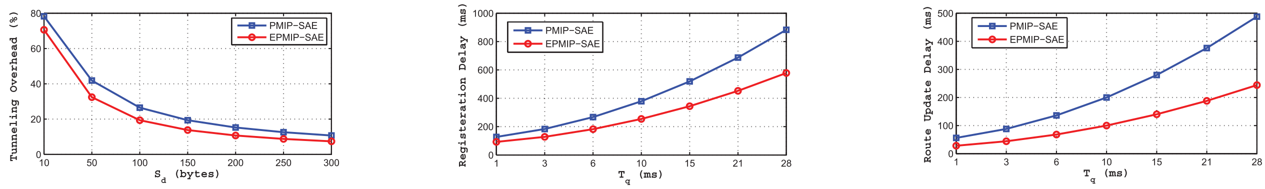
$$TO_{EPMIP-SAE} = \frac{GRE/IP}{\text{Data Packet Size}(S_d)+GRE/IP} \times 100$$

C. Registration Delay (RD)

In PMIP-SAE, the registration operation is performed as follows. When a UE enters the eNB region, it establishes a radio link and sends an *Attach Request* message to MME. This operation takes $T_{UE-MME}(S_c) = T_{UE-eNB}(S_c) + T_{eNB-MME}(S_c)$. Then, MME exchanges the *Update Location Request* and *Response* messages with HSS. It takes $2 \times T_{MME-HSS}(S_c)$. MME also sends a *Create Session Request* message to S-GW/MAG while taking $T_{MME-SGW}(S_c)$. After that, S-GW/MAG performs the binding operation with P-GW/LMA by exchanging *PBU* and *PBA* messages, taking $2 \times T_{SGW-PGW}(S_c)$. Then, S-GW/MAG responds with a *Create Session Response* message to MME, which takes TMME-SGW(Sc). MME exchanges the *Initial Context Setup Request* and *Response* messages with eNB. This operation takes $2 \times T_{eNB-MME}(S_c)$. Then, UE sends the *Attach Complete* message to MME, which takes $T_{UE-MME}(S_c) = T_{UE-eNB}(S_c) + T_{eNB-MME}(S_c)$. MME now exchanges the *Modify Bearer Request* and *Response* messages with S-GW/MAG. It takes $2 \times T_{MME-SGW}(S_c)$. Accordingly, the registration delay (RD) of PMIP-SAE can be represented as follows.

$$RD_{PMIP-SAE} = 2T_{UE-eNB}(S_c) + 4T_{eNB-MME}(S_c) + 2T_{MME-HSS}(S_c) + 4T_{MME-SGW}(S_c) + 2T_{SGW-PGW}(S_c)$$

In EPMIP-SAE, the registration operation is performed as follows. When a UE enters the eNB/MAG region, it establishes a radio link and sends an *Attach Request* message to eNB/MAG. This operation takes $T_{UE-eNB}(S_c)$. Then, eNB/MAG sends an *Attach Request* incorporating the *PBU* message to MME, which takes $T_{eNB-MME}(S_c)$. The MME



(a) Impact of S_d on tunneling overhead.

(b) Impact of T_q on registration delay.

(c) Impact of T_q on route update delay.

Fig. 6. Impact of different parameters on tunneling overhead, registration delay and route update delay.

exchanges the *Update Location Request* and *Response* messages with HSS, while taking $2 \times T_{MME-HSS}(S_c)$. After that, MME also sends a *Create Session Request* with a *PBU* message to S-GW/P-GW/LMA, which takes $T_{MME-SGW}(S_c)$. Then, S-GW/P-GW/LMA responds with a *Create Session Response* with a *PBA* message to MME. It takes $T_{MME-SGW}(S_c)$. MME exchanges the *Initial Context Setup Request* and *Response* messages with the eNB/MAG. This operation takes $2 \times T_{eNB-MME}(S_c)$. After that, UE sends the *Attach Complete* message to MME, which takes $T_{UE-MME}(S_c) = T_{UE-eNB}(S_c) + T_{eNB-MME}(S_c)$. Accordingly, the registration delay of EPMIP-SAE can be represented as follows.

$$RD_{EPMIP-SAE} = 2T_{UE-eNB}(S_c) + 4T_{eNB-MME}(S_c) + 2T_{MME-HSS}(S_c) + 2T_{MME-SGW}(S_c)$$

D. Route Update Delay (RUD)

In PMIP-SAE, when a UE moves to another eNB region by handover and attaches to Target eNB, the Target eNB sends *Path Switch Request* to MME. Then, MME sends *Modify Bearer Request* to S-GW/MAG. After that, S-GW/MAG exchanges *PBU* and *PBA* messages with P-GW/LMA. S-GW/MAG now responds with *Modify Bearer Response* to MME, and the MME responds with *Path Switch Request Ack* to Target eNB. So, we obtain the route update delay after handover (RUD) of the PMIP-SAE as follows:

$$RUD_{PMIP-SAE} = 2T_{eNB-MME}(S_c) + 2T_{MME-SGW}(S_c) + 2T_{SGW-PGW}(S_c)$$

In EPMIP-SAE, when a UE moves to another eNB/MAG region by handover and attaches to Target eNB/MAG, the Target eNB/MAG exchanges PBU and PBA messages with S-GW/P-GW/LMA. So, we obtain the route update delay after handover (RUD) of the EPMIP-SAE as follows. $RUD_{EPMIP-SAE} = 2T_{eNB-SGW}(S_c)$

E. Numerical Results and Discussion

Based on the analytical equations, we compare the performance of PMIP-SAE and EPMIP-SAE. For the numerical analysis, we configure the default parameters as follows by referring to [4]: $H_{eNB-SGW} = 2$, $H_{SGW-PGW} = 3$, $H_{eNB-MME} = 2$, $H_{MME-HSS} = 3$, $H_{MME-SGW} = 2$, $L_{wl} = 10ms$, $L_w = 2ms$, $T_q = 5ms$, $S_c = 50$ bytes, $S_d = 200$ bytes, $B_{wl} = 11$ Mbps, and $B_w = 100$ Mbps, respectively. Among the parameters, we note that T_q and S_d depend on the network conditions. Thus, we observe the performance of candidate schemes by varying those parameters.

We plot the tunneling overhead for varying payload sizes in Fig. 6a. We observe that the payload size impacts the tunneling overhead significantly for both PMIP-SAE and EPMIP-SAE schemes. This is because the GTP/IP/UDP headers are added with the payload in the PMIP-SAE scheme between eNB and S-GW/MAG and also between eNBs in X2-based handover, and in the EPMIP-SAE scheme, the GRE/IP headers are added with the payload between eNB/MAG and S-GW/P-GW/MAG and also between eNB/MAGs. The proposed scheme shows smaller overhead than the existing scheme because the GTP/IP/UDP headers are 36 bytes while the GRE/IP headers are 24 bytes only. We compare the registration delay of both mobility schemes for various average queuing delays (T_q) at each node in Fig. 6b. The registration delay increases linearly with T_q, for both mobility schemes, but our proposed scheme gives much better performance especially for the large queuing delay. Fig. 6c shows the impact of average queuing delay (T_q) on the route update delay. The route update delay increases linearly with T_q, for both mobility schemes. Note that the proposed scheme provides better performance than the existing scheme because in our scheme, Target eNB/MAG exchanges PBU and PBA messages directly with S-GW/P-GW/LMA, not through MME.

V. Conclusions

In this letter, we proposed an enhanced mobility management scheme for SAE based mobile networks. We combined S-GW and P-GW of SAE and let the combined entity work as an LMA of PMIP, and eNB functions as an MAG of PMIP. GRE tunnels are used between eNB/MAG and S-GW/P-GW/LMA and also between eNB/MAGs, instead of GTP to reduce the header overhead. In handover, Target eNB/MAG exchanges PBU and PBA messages directly with S-GW/P-GW/LMA, not through the MME. Numerical analysis demonstrates that our scheme shows better performance than the previous mobility scheme for SAE based networks.

References

[1] S. Gundavelli *et al.*, "Proxy mobile IPv6," Internet Engineering Task Force, Internet Society, USA, IETF RFC 5213, Aug. 2008.
[2] J. Laganier *et al.*, "Mobility management for All-IP network," *NTT DOCOMO Tech. J.*, vol. 11, no. 3, pp. 34–39, Sep. 2009.
[3] 3GPP TR23.402, "Tech. Spec. Group services and system aspects: Architecture enhancements for non-3GPP accesses," v10.7.0, Mar. 2012.
[4] M. Gohar *et al.*, "Trill-based mobile packet core network for 5G mobile communication systems," *Wireless Pers. Commun.*, vol. 87, no. 1, pp. 125–144, Jan. 2016.
[5] 3GPP TR23.401, "Tech. Spec. Group services and system aspects; General packet radio service (GPRS) enhancements for evolved universal terrestrial radio access network (E-UTRAN) access," v10.7.0, Mar. 2012.