Signalling Effort Evaluation of Mobility Protocols within Evolved Packet Core Network

Sandra Frei[1,3], Woldemar Fuhrmann[2], Andreas Rinkel[3], Bogdan Ghita[1]

[1]Centre for Security, Communications and Network Research
University of Plymouth, United Kingdom
[2]University of Applied Sciences Darmstadt, Germany
[3]University of Applied Sciences Rapperswil, Switzerland
sandra.frei@plymouth.ac.uk
w.fuhrmann@fbi.h-da.de
arinkel@hsr.ch
bogdan.ghita@plymouth.ac.uk

Abstract: The Evolved Packet Core (EPC), IP-based, as defined by 3GPP, is the core network for the 4th generation networks. Within the EPC there are two mobility protocols defined to use on the S5/S8 interface between the Serving Gateway (SGW) and the Packet Data Network Gateway (PGW). The architecture and functionalities of the two involved gateways are different either of GPRS Tunnelling Protocol (GTP) or Proxy Mobile IPv6 (PMIPv6) is deployed as the mobility management protocol on the S5/S8 interface. This paper evaluates the signalling costs of the two mobility management protocols when dynamic QoS and policy control are applied. The two mobility protocols are functionally compared against each other in general as well as within the EPC architecture. The evaluation of the signalling effort of the two mobility protocols is done by means of handover with SGW relocation and shows that GTP performs better than the PMIPv6.

1 Introduction

The EPC is the core network of the 4th generation network and was defined by the 3GPP. It is based on IP and provides integration of several different radio access technologies, such as the Long Term Evolution (LTE), non-3GPP access technologies such as WiMAX or WLAN, as well as the integration of second and third generation networks. The mobility protocol of the second and third generation networks is the GTP, which was subsequently adjusted and improved when used in the Universal Mobile Telecommunication System (UMTS) network and once more for the use with LTE/EPC. Within the EPC, 3GPP defined an alternative mobility protocol on the S5/S8 interface between the SGW and PGW, which is the PMIPv6 [GLD+08].

Prior research provided a signalling cost evaluation of the mobility protocols GTP, PMIP and MIP for different core network architectural arrangements in LTE/Service Architecture Evolution (SAE) [WGT08]. The 3GPP standards used to identify the signalling cost of the different mobility protocols were based on the release 7 and on an early draft version of the release 8. At the time, the paper [WGT08] was written, PMIP was defined as the alternative protocol beside GTP, but neither any Quality of Service (QoS) interactions with the Policy and Charging Control Architecture (PCC) for PMIP or the Evolved Packet System (EPS) bearer with a PMIP-based S5/S8 interface were defined. The result of the paper [WGT08] was that PMIP has great advantage on reducing the signalling cost of mobility management compared with the GTP and MIP.

3GPP started to define already the release 9 within the actual standards. All the necessary QoS interactions from the gateways with the PCC are currently defined for both mobility protocols used on the S5/S8 interface. Therefore an evaluation of the signalling effort from GTP and PMIPv6 during a handover is done once more, but this paper includes the QoS interactions, which are necessary if dynamic QoS will be provided. This paper aims to evaluate the signalling effort for both GTP and PMIP mobility protocols and recommends the preferred mobility protocol on the S5/S8 interface.

The paper is organised as follows: Section 2 provides the architectural differences between GTP and PMIP as mobility protocols on the S5/S8 interface. Section 3 analyses the functional differences of the two mobility protocols in general, as well as their functionality within the EPC environment. Section 4 presents the signalling efforts of the GTP and PMIP for an S1 handover in a scenario with multiple bearers. Finally section 5 concludes the paper.

2 Architectural differences between the GTP-based and PMIP-based S5/S8 interface

The architecture and the functionality of the EPC entities vary, subject to the deployed mobility protocol on the S5/S8 interface. In this section, the two protocol alternatives, GTP and PMIP, are explained and the consequences which result from the different protocols are outlined. The architecture of the EPC with the interworking towards the PCC, where GTP is deployed on the S5/S8 interface, is shown in Figure 1 A. At the Packet Data Network Gateway (PGW), several Service Data Flows (SDF) are pending on the downlink direction. Through downlink Traffic Flow Templates (TFT), the SDFs are mapped on the appropriate EPS bearer, which supports a certain QoS and is represented by one GTP tunnel on the interfaces S5/S8 and S1-U as well as the LTE-Uu interface. The filters used to do the mapping are [source_IP, destination_IP, source_port, destination_port, transport_protocol_ID] IP-5-tuples. The TFTs are provided by the Policy Control Charging Rules Function (PCRF) towards the Policy Charging Enforcement Function (PCEF) within the QoS policy information and PCC rules. The PCEF is located in the PGW[1].

Within the other EPS components, the SGW and the evolved Node B (eNB), there is a 1 1 mapping between the different bearers in both up and downlink direction. The UE receives the same QoS policy information and PCC rules like the PCEF and this information includes also the uplink TFTs to map the traffic to be sent towards the appropriate radio bearer.

[1] The PGW also manages traffic charging, which is outside the scope of this paper

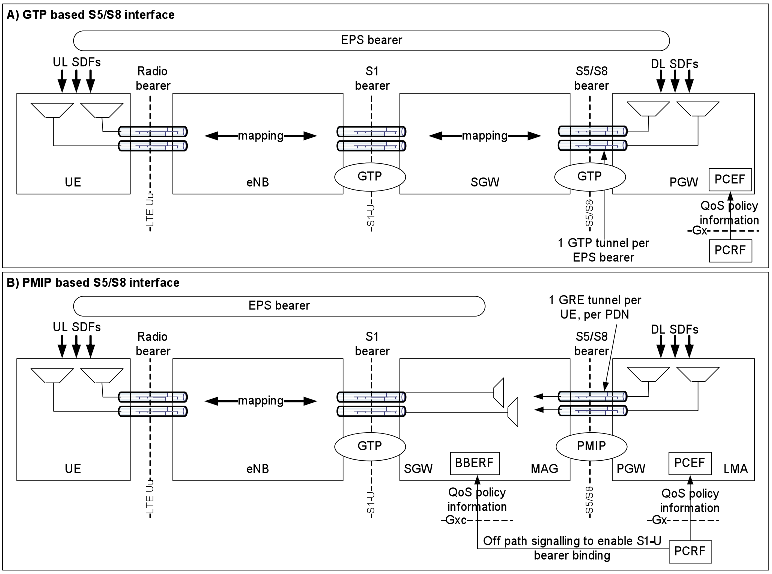

Figure 1: Architectures for GTP and PMIP-based S5/S8 interface

When PMIP is deployed on the S5/S8 interface, the architecture changes significantly compared with the GTP-based S5/S8 interface. A major difference is the shortened EPS bearer. It lasts only from the UE over the eNB to the SGW. The PGW is not anymore part of the EPS bearer. In Figure 1 B) there are filters on both EPC entities, the PGW and the SGW. The SGW takes over the bearer binding functionality, which is done in the PGW if GTP is deployed on the S5/S8 interface. As a result of this new functionality the SGW has to perform, the PCRF must provide the necessary QoS policy information towards the Bearer Binding Event Reporting Function (BBERF) entity, which is similar to the PCEF but located within the SGW. Only with the QoS policy information it is possible for the SGW to map the incoming data flows towards the appropriate EPS bearer. The reasons to still apply filters in the PGW are on one hand because of the charging in the PGW and on the other hand because the traffic shaping is performed by the PGW. But the filters on the PGW have no relevance on the bearer binding. Between the PGW and the SGW Generic Routing Encapsulation (GRE) tunnels are established to separate the traffic for each UE and PDN. This separation is to enable the UE to do a handover per PDN. Therefore, if a UE has connections to multiple PDNs, it is possible to independently do a handover for a single PDN connection. The GRE tunnels do not support any QoS. The QoS is enforced by the traffic shaping on the PGW and on the eNB as well as with the EPS bearer mapping at the SGW.

3 Functional differences between GTP and PMIP

GTP and PMIP are both mobility protocols, but GTP includes additional functionality besides mobility. Table 1 compares the support of several functions between the two mobility protocols GTP and PMIPv6.

	GTP	PMIPv6
QoS signalling	+	-
Bearer signalling	+	-
Packet forwarding during handover	+	-
Mobility tracking	+	+, always active
Paging	+	+, because always active
Network based mobility management scheme	+	+

Table 1: Functions supported by GTP / PMIPv6

As shown, GTP supports QoS signalling and is able to establish and modify bearers. In case of a handover, data can be forwarded since GTP can be used to establish the necessary tunnels for forwarding data on the user plane. From the PMIPv6 point of view, it does not provide functions to decrease the power consumption of the end device, since the end device is always active. The ability to change the state of the end device from active to idle mode is left to the access technology.

Within the EPC, both mobility protocols can be used. As shown in Figure 2, communication between the eNB, MME and SGW uses GTP (area 1), and either of GTP or PMIP can be deployed on the S5/S8 interface (area 2). The fact that, regardless of the protocol used on the S5/S8 interface (see Figure 2 area 2), GTP is used towards the eNB (see Figure 2 area 1) will eliminate some of the functional drawbacks of the PMIPv6.

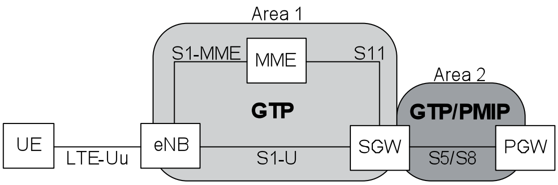

Figure 2: EPC with deployed mobility protocols

Table 2 lists which functional drawbacks of the PMIPv6 can be compensated by the GTP of area 1.

	Area 1/GTP	Area 2/GTP	Area 2/PMIP
QoS signalling	+	+	-
Bearer signalling	+	+	-
Packet forwarding during handover	+	Is performed in area 1	Is performed in area 1
Mobility tracking	+	Is performed in area 1	Is performed in area 1
Paging	+	Is performed in area 1	Is performed in area 1
Network based mobility management scheme	+	+	+

Table 2: Functions supported by the areas with deployment variations of the mobility protocol

Since area 1 provides GTP functionality, the performance of the handover process can be improved as well as the power consumption behaviour of the UE, through mobility tracking and paging functions, provided by area 1. But the QoS and bearer signalling functionalities are still not available on the PMIP-based S5/S8 interface.

Within area 1, the signalling effort is the same, independent of the mobility protocol used in area 2. The signalling effort of the area 2 is different for GTP and PMIP. As described previously, the architecture as well as the functionality of the entities depends on the deployed mobility protocol in area 2. If the S5/S8 interface is PMIP-based, additional signalling has to be performed because the PGW can only map the incoming data traffic per UE and per PDN towards a GRE tunnel. The bearer binding is done in the SGW and therefore additional QoS information has to be provided by the PCRF to the BBERF, located within the SGW. A dynamic QoS can only be performed within the EPS bearer, which is from eNB to SGW. In the case where GTP is deployed on the S5/S8 interface, the bearer binding can be performed directly by the PGW with one GTP tunnel per EPS bearer. As a result, dynamic QoS can be provided from the UE to the PGW by using EPS bearers, with each of it has its own QoS capabilities. To perform QoS on the S5/S8 interface when PMIP is used, Differentiated Services (DiffServ) with the appropriate Differentiated Services Code Points (DSCPs) could be used, but the limitation is that the provided QoS with DSCP is static.

4 Handover signalling effort for GTP and PMIP

As previously discussed, the signalling effort for both mobility protocols is the same for area 1. There are two different scenarios, depending on the deployed mobility protocol on the S5/S8 interface. Scenario 1 is applied if GTP is deployed on the S5/S8 interface and the scenario 2 is with a PMIPv6-based S5/S8 interface. This section analyses the signalling effort for an S1 handover with SGW relocation in both scenarios.

The signalling effort for an S1 handover with SGW relocation is also applicable for an X2 handover with SGW relocation. The signalling efforts consist of all the signalling effort of all protocols above the layer 2, where IPv6 is used as the layer 3 protocol. The following equations use the abbreviation *tse* for the total signalling effort above layer 2. The unit of *tse* is Byte.

4.1 Signalling effort for S1 handover within Scenario 1 (GTP)

The signalling effort on the S5/S8 interface using GTP as the mobility protocol contains the GTP request Modify Bearer Request (MBReq) which is sent from the SGW to the PGW and the GTP response Modify Bearer Response (MBResp) sent from the PGW to the SGW. For calculating the signalling effort, the sequence diagrams of the 3GPP standard [TS09f] were used to identify the relevant messages and its Information Elements (IEs). Furthermore the 3GPP standard [TS09a] was used to calculate the message sizes on the basis on the identified IEs. Since the signalling effort is here defined as the whole signalling effort above layer 2 the UDP [Po80] and IPv6 [DH98] headers (UDP_hdr, IP_hdr) have to be added to the signalling effort of the GTP message, resulting in:

$$tseMBReq=MBReq+IP\_hdr+UDP\_hdr=62+40+8=110$$

The Modify Bearer Response message has the following signalling effort:

$$tseMBResp=MBResp+IP\_hdr+UDP\_hdr=44+40+8=92$$

The whole signalling effort for both messages is 202 Bytes.

4.2 Signalling effort for S1 handover within Scenario 2 (PMIP)

The necessary actions, which have an impact on the signalling effort and are performed when using PMIP on the S5/S8 interface, are shown in Figure 3. First the SGW has to establish a control session towards the PCRF over the Gxc interface to get the necessary QoS policy information to perform the bearer binding at the SGW.

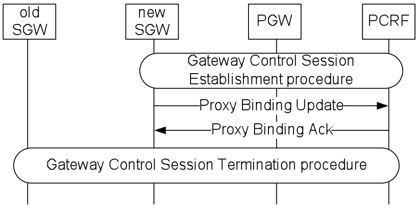

Figure 3: S1 handover procedure with PMIP-based S5/S8 interface

Then the proxy binding update with the appropriate response can be processed and finally the old SGW has to terminate its session towards the PCRF.

To get the QoS information, the SGW performs a Gateway Control Session Establishment (GCSE) procedure towards the PCRF. This is done by sending a Credit Control Request (CCR) Diameter message (CCR_msg) from the SGW to the PGW and receiving a Credit Control Answer (CCA) Diameter message (CCA_msg) from the PCRF, which contains all the necessary QoS information to enable the SGW doing the bearer binding. The layer 4 protocol is the Stream Control Transmission Protocol (SCTP) and the header and chunk (SCTP_hdr_chunk) have to be added to get the total signalling effort for the QoS Request. Based on the sequence diagrams and the message descriptions of the 3GPP standards [TS09c, TS09d, TS09e, TS09g], the RFC definitions of the diameter protocol [CLG+03, HMK+05], the PMIPv6 definitions [GLD+08, JPA04, KND08, MKGL09, TS09b], the IPv6 definition [DH98], the UDP definition [Po80] and the SCTP definition [St07], the calculations of the signalling efforts have been made. The following signalling efforts are identified to request the QoS information from the PCRF

tseQoSReq=CCR_msg+SCTP_hdr_chunk+IP_hdr=336+28+40=404

and to provide QoS information from the PCRF to the SGW:

tseQoSResp=CCA_msg+SCTP_hdr_chunk+IP_hdr=972+28+40=1040

Following the GSCE, a Proxy Binding Update (PBU) message will be sent from SGW to the PGW, and then the Proxy Binding Acknowledge (PBA) message will be sent from the PGW to the SGW.

tsePBU=PBU+IP_hdr=104+40=144

tsePBA=PBA+IP_hdr=104+ 40=144

Finally the session between the old SGW and the PCRF has to be terminated. The termination of the session is initiated by the SGW sending a CCR Diameter message to the PCRF and the PCRF responds with a CCA Diameter message. The identified signalling efforts for the session termination request and the response result in:

tseSessionTerminationReq=CCR_msg+SCTP_hdr_chunk+IP_hdr=212+28+40=280

tseSessionTerminationResp=CCA_msg+SCTP_hdr_chunk+IP_hdr=188+28+40=256

The percentage of the signalling effort per procedure is shown in Figure 4. The PBU and the PBA have the lowest impact, followed by the Gateway Control Session Termination (GCST) procedure. The highest impact is caused by the GCSE procedure, which contains the QoS information.

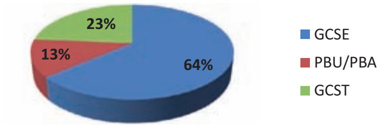

Figure 4: Percentage of the signalling effort per procedure

The overall signalling effort for the necessarily performed procedures, with PMIP as the mobility protocol, is 2268 Bytes. This is factor 11.2 more signalling effort for the PMIP than the GTP requires. This result, that GTP performs better than PMIPv6, is completely the opposite of the results of the paper [WGT08]. This is caused by the fact that in the paper [WGT08] no signalling effort is considered concerning the QoS information. For PMIPv6 these QoS information have to be sent from the PCRF to the SGW. This is not necessary if GTP is used as the mobility protocol. And as showed in this section, the signalling effort of the QoS information is significant and has the highest impact on the overall signalling effort. This impact increases even more, if multiple bearers are used, showed in the next section.

4.3 Signalling effort for S1 handover with multiple bearers within Scenario 1 and 2

A UE can hold a maximum of 11 bearers. This maximum value has effect on both scenarios. But the GTP and PMIP-based S5/S8 interface causes different additional signalling effort (*se*) for an additional bearer. In scenario 1 (*s1*) the additional information for one bearer consists of the bearer context, which is an Information Element (IE) within the GTPv2-C.

$$\Delta s1\_se\_per\_bearer = \text{bearer context} = 43 \text{ Bytes}$$

Scenario 2 (*s2*) requires an additional QoS-Rule-Install Attribute Value Pair (AVP) within the CCA diameter message for one additional bearer.

$$\Delta s2\_se\_per\_bearer = \text{QoS-Rule-Install AVP} = 792 \text{ Bytes}$$

If PMIP is the used mobility protocol the additional signalling effort for an extra bearer is 18.4 times bigger than with GTP. A number of 4 to 6 bearers is common. If we assume a UE has 2 connections to different PDNs. One connection is towards the IMS, the other connection towards the internet. One default bearer is needed for each PDN. Furthermore assume that the UE has 2 dedicated bearers for the IMS, one for voice service and the other one for streaming service, as well as one dedicated bearer for the internet PDN. An overview of the establishment procedures of default and dedicated bearer including a signalling effort evaluation is given in [FFRG09]. This makes a total of 5 bearers for the UE.

$$total\_bearer\_number = PDN\_number + dedicated\_bearer\_number$$

Figure 5 shows the percentage of the signalling effort of the GCSE procedure within the scenario 2 in relation to the overall signalling effort of the scenario 2. The impact of the

GCSE procedure, where the session from the SGW towards the PCRF is established and the needed QoS information are sent to the SGW increases from 64 % to 92 %.

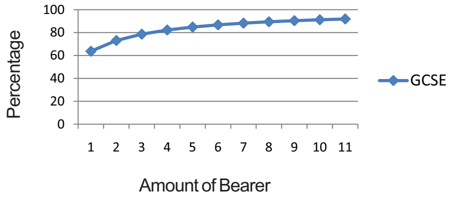

Figure 5: Percentage signalling effort of GCSE to overall signalling effort

The impact of the QoS information on the signalling effort is tremendous. And this impact appears not only during handover, but also during bearer modification, QoS modification and if a dedicated bearer is activated.

5 Conclusion

The functional differences of the two mobility protocols, the GTP and the PMIP, have been discussed in this paper. After reviewing the two mobility protocols in general they were analysed as well in the EPS environment, and were compared with the general functionalities of the mobility protocols. The analysis indicates that GTP in area 1 provides additional and important functionalities which the PMIP cannot deliver. Furthermore the signalling effort for a S1 handover with relocation of the SGW was analysed. The result was an 11.2 times bigger signalling effort for PMIP than with GTP deployed on the S5/S8 interface. This additional signalling effort of PMIP is mainly caused by the QoS information sent from the PCRF towards the SGW to enable the SGW to perform the bearer binding. Finally, the signalling effort of one to a maximum of 11 established bearers was analysed in both scenarios. Through this analysis it was shown that the signalling effort for the QoS information is huge and accounts for 64% to 92 % of the overall signalling effort if PMIPv6 is deployed on the S5/S8 interface, depending on the numbers of bearers. The QoS information are also required to be sent from PCRF to the SGW during bearer modification, QoS modification and if a dedicated bearer is activated. As a result of these signalling effort evaluations one can say that the GTP performs clearly better during the S1 handover and in every situation when the QoS information has to be sent from the PCRF to the SGW, as the alternative mobility protocol, the PMIPv6.

References

[CLG+03] Calhoun, P.; Loughney, J.; Guttman, E.; Zorn, G.; Arkko, J.: *Diameter Base Protocol: RFC 3588*. IETF, 2003

[DH98] Deering, S.; Hinden, R.: *Internet Protocol, Version 6 (IPv6) Specification: RFC 2460*, IETF, 1998

[FFRG09] Frei, S.; Fuhrmann, W.; Rinkel, A.; Ghita, B.: Reference architecture for end-to-end QoS in heterogeneous wireless network environments: *Proceedings of the Third International Conference on Internet Technologies and Applications (ITA 09)*, Glyndwr University, Wrexham, UK, 2009; pp. 630–639

[GLD+08] Gundavelli, S.; Leung, K.; Devarapalli, V.; Chowdhury, K.; Patil, B.: *Proxy Mobile IPv6: RFC 5213*, IETF, 2008

[HMK+05] Hakala, H.; Mattila, L.; Koskinen, J-P.; Stura, M.; Loughney, J.: *Diameter Credit-Control Application: RFC 4006*. IETF, 2005

[JPA04] Johnson, D.; Perkins, C.; Arkko, J.: *Mobility Support in IPv6: RFC 3775*, IETF, 2004

[KND08] Korhonen, J.; Nilsson, U.; Devarapalli, V.: *Service Selection for Mobile IPv6: RFC 5149*, IETF, 2008

[MKGL09] Muhanna, A.; Khalil, M.; Gundavelli, S.; Leung, K: *GRE Key Option for Proxy Mobile IPv6: draft-ietf-netlmm-grekey-option-09*, IETF, 2009

[Po80] Postel, J.: *User Datagram Protocol: RFC 768*, IETF, 1980

[St07] Stewart, R.: *Stream Control Transmission Protocol: RFC 4960*, IETF, 2007

[TS09a] Technical Specification Group Core Network and Terminals: 3GPP Evolved Packet System (EPS); *Evolved General Packet Radio Service (GPRS) Tunnelling Protocol for Control plane (GTPv2-C)*. 3GPP TS 29.274 V9.1.0, Stage 3, Release 9, 2009

[TS09b] Technical Specification Group Core Network and Terminals: *Numbering, addressing and identification*. 3GPP TS23.003 V9.1.0, Release 9, 2009

[TS09c] Technical Specification Group Core Network and Terminals: *Policy and Charging Control signalling flows and QoS parameter mapping*. 3GPP TS 29.213 V9.1.0, Release 9, 2009

[TS09d] Technical Specification Group Core Network and Terminals: *Policy and Charging Control over Gx reference point*. 3GPP TS 29.212 V9.1.0, Release 9, 2009

[TS09e] Technical Specification Group Services and System Aspects: *Architecture enhancements for non-3GPP accesses*. 3GPP TS 23.402 V9.3.0, Release 9, 2009

[TS09f] Technical Specification Group Services and System Aspects: *General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access*. 3GPP TS 23.401 V9.3.0, Release 9, 2009

[TS09g] Technical Specification Group Services and System Aspects: *Policy and Charging Control Architecture*. 3GPP TS 23.203 V9.3.9, Release 9, 2009

[WGT08] Wang, M.; Georgiades, M.; Tafazolli, R.: Signalling Cost Evaluation of Mobility Management Schemes for different Core Network Architectural Arrangements in 3GPP LTE/SAE. *Vehicular Technology Conference*, IEEE VTC, 2008; pp. 2253–2258