This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE Globecom 2011 proceedings.

Ensuring Service Resilience in the EPS: MME Failure Restoration Case

Tarik Taleb and Konstantinos Samdanis
NEC Laboratories Europe
{taleb,samdanis}@.neclab.eu

***Abstract*—In the Evolved Packet System (EPS), service resiliency can be heavily impacted by a node failure, especially in its control plane. Ensuring service resiliency via defining efficient and proactive restoration mechanisms is of vital importance. In this paper, we address the case of Mobility Management Entity (MME) failure. We propose schemes for MME failure detection and restoration considering User Equipments (UEs) in both idle and active mode. As a MME failure may impact a potential number of UEs, network signaling overload control, via randomized paging and handling signaling messages in bulk, is considered. The proposed schemes are evaluated through simulations and encouraging results are obtained.**

I. INTRODUCTION

Service resiliency is an important requirement in any communications system. Considering the case of EPS, service resiliency can be largely impacted due to link or nodal failures. Node failures may affect the control plane for example impacting MME as well as the data plane impacting nodes such as Serving gateways (S-GWs) or Packet Data Network gateways (PDN-GWs). In the control plane of EPS, MME's role is significant since it is in charge of setting up and maintaining connections to a potential number of users. Its failure significantly impacts the service provisioning, thus the importance of studying the EPS service resiliency by defining prompt, scalable and reliable mechanisms for restoration from a MME failure. Although the focus of this paper is centered on the Long Term Evolution (LTE), it can be easily extended for other radio technologies. A typical MME failure scenario following the procedures currently defined in 3GPP standards [1]-[4], is depicted in Fig.1, focusing on a terminating IMS (IP Multimedia Subsystem) call arriving after a MME failure. Such a process follows these steps:

1) The MME has failed and restarted.
2) S-GW detects the MME restart via the incremented counter in a GTP - GPRS Tunnel Protocol - echo message and removes all UE resources handled previously on this MME. It should be noted that the removal of resources is not propagated directly up to the PDN GW, i.e., the allocated IP address and a S5/S8 tunnel configuration in PDN GW remains valid.
3) IMS call establishment signaling arrives at the PDN GW.
4) Data packets stemming from IMS arrive at S-GW are discarded, due to unknown Tunnel Endpoint Identifiers (TEIDs).
5) (a) S-GW sends a reject message to PDN GW; (b) upon reception, the PDN GW removes all resources linked to the relevant IP address.
6) The loss of the SIP (Session Initiation Protocol) signaling messages leads to an error situation in IMS, impacting the service and ultimately the system resiliency.

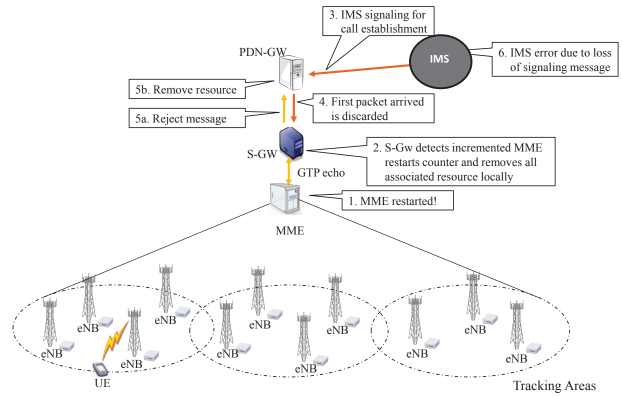

Fig. 1. MME failure scenario (as currently defined in 3GPP standards).

Intuitively, a major concern with the described solution is its reactive nature. Indeed, it awaits the restart of the failed MME, and while waiting, any communications to/from UEs handled by the affected MME will be significantly disrupted. In this paper, we propose a proactive MME restoration approach, i.e. as soon as a MME goes off, MME relocation for affected UEs should take place. The key concept behind the proposed solution is to proactively trigger MME relocation and restoration of lost state to avoid service disruption at a later stage. The proposed approach addresses UEs in both idle and active mode. For the former, it triggers all affected idle-mode UEs through "scheduled/randomized bulk paging" to re-attach to the network. As for the latter, it allows ongoing communications to proceed and triggers affected UEs to perform a Tracking Area Update (TAU) in a scheduled manner according to predefined priorities. Both mechanisms lead to the selection of a new MME for affected UEs and restoration of their context in a pro-active manner. A number of mechanisms facilitating the selection of new MMEs for affected UEs and a proactive restoration of UEs' context are also envisioned.

The remainder of this paper is organized as follows. Section II presents the state of the art. Section III presents our proposed MME restoration solutions with all details pertaining to bulk signaling including UEs in both idle and active mode. Section IV evaluates the overall approach and showcases its technical benefits. The paper concludes in Section V.

II. STATE OF THE ART

Failure recovery of mobile systems is critical for the service provision and has thus been exhaustively investigated in the recent literature. A framework that describes the initial efforts towards fault tolerance in wireless access networks is described in [5] focusing on the Global System for Mobile

978-1-4244-9268-8/11/$26.00 ©2011 IEEE

communication (GSM), where the primary means of fault provision is transport and coverage redundancy. A similar approach is also adopted in all-IP wireless networks, where the restoration and recovery policy consists in duplicating mobility information at different agents. Thus, when one fails, the backup operates instead [6].

An analogous scheme centered on UMTS is described in [7], where a per-user check pointing approach is introduced for the Home Location Register (HLR). The proposed method regulates the duplication of HLR information considering user activity to reduce the associated overhead. Given the high dynamicity of mobile users along with the increase in the number and types of mobile applications, maintaining a double mobility registration is not only an expensive process, but also an inefficient one requiring significant synchronization effort. For this reason, alternative methods that dynamically select backup mobility agents on demand are introduced. In [8] once a mobility agent fails, the system estimates the affected load and selects the backup agents initiating a system driven proactive handover considering load balancing. A distributed variant is presented in [9], while 3GPP LTE adopts the same approach for PCRF (Policy Charging and Rules Function) [4] [10]. In the PCRF case, the failure can be detected by the DIAMETER protocol [11] or at the PCRF application level. During a PCRF failure, other PCRF entities are simply utilized instead with not much impact on the data plane.

In 3GPP LTE the adoption of the S1-flex provides the fundamental means for network redundancy in the EPC creating pools of MMEs or S-GWs [12]. A pool may serve certain eNBs allowing them to be connected to multiple MMEs and S-GWs at the same time. When a certain MME or S-GW fails the affected UEs may select another entity from the same pool. To the best knowledge of the authors, no prior research work addressed failure detection and restoration of EPS's control and data plane nodes. Since the control plane is of utmost importance, this paper aims to devise fault detection and restoration methods for the control plane avoiding double mobility registration. In addition, it introduces bulk signaling, which advances the current state reducing the signaling and processing overhead within the core and access network. Our vision is to provide such failure restoration as a Self-Organized Network function [13], whereby network nodes may autonomously adapt to the MME pool accordingly, while the re-configuration process of affected MMEs is seamless.

III. MME Restoration with Bulk Signaling

The proposed solution is a combination of different mechanisms, each with a specific objective. Firstly, there are different ways for the detection of a MME failure. The first method could be via an explicit intervention/notification from the Operations & Maintenance (O&M) system. Indeed, O&M could detect MME failure i) based on feedback from supervising software daemons on the MMEs, ii) based on periodic keep-alive/echo messages, iii) having MME immediately send an alarm to O&M right before it crashes i.e. in case of partial failure, or iv) by analyzing related information (e.g., handover occurrences) from other network elements such as eNBs, S-GWs, etc. The MME failure can be detected directly by eNBs using S1-MME i.e., using keep-alive messages of the Stream Control Transmission Protocol - SCTP. Alternatively, MME failure can be detected by neighbor MMEs using S10 protocol means or by S-GWs using S11 protocol means. In the remainder of this section, we discuss how MME restoration can be separately achieved for UEs in idle mode and UEs in active mode.

A. MME Restoration for Idle Mode UEs

The proposed procedure presents a paging enhancement enabling notification of all UE's that have been served by a particular MME in bulk. Indeed, the "bulk" paging is characterized by the use of MME information - which is the leading part of GUTI - Globally Unique Temporary Identity - as identifier [14]. Upon detecting a MME failure, all associated eNBs with S1-MME connection, initiate bulk paging of all UEs being served by the failed MME. During the re-attachment, eNBs relocate responding UEs to the MMEs remaining in operation, taking load balancing into account.

Due to the failure of the originally assigned MME, an eNB needs to relocate the UE to another MME by releasing the Radio Resource Control (RRC) connection. UEs will then re-establish the RRC connection and subsequently perform a TAU operation. Such a mechanism would in principle trigger many UEs to re-attach at the same time. To avoid signaling overload at the newly selected MMEs, the re-attach attempts should be spread out over time. This can be achieved by different mechanisms as will be explained later.

Paging UEs in idle mode and affected by the MME failure can be also initiated by neighboring MMEs. Generally speaking, a MME A is said to be a neighbor of MME B if both MMEs have at least one common Tracking Area [12]. Effectively, one or more neighbor MMEs may detect a MME failure and then initiate bulk paging addressing idle mode UEs. Duplicate paging shall be minimized, if not entirely avoided. This can be achieved via different methods. In case a neighbor MME detects the MME failure, it immediately starts the paging and notifies its neighboring MMEs that it has already paged the impacted UEs. This mechanism assumes that MMEs have prior knowledge on the pool of MMEs that are able to cover a failing MME. In case O&M detects the MME failure and notifies the neighbor MMEs, O&M explicitly indicates to each MME which Tracking Area it should page. Additionally, eNBs may filter out duplicated paging messages stemming from different MMEs. In case of an inevitable reception of duplicate paging messages, UEs simply consider the first message and discard the following ones.

Whilst signalling congestion, due to simultaneous attempts from many affected UEs to re-attach, can be avoided at MMEs by taking load balancing into account, this imminent congestion can be also avoided by handling signalling messages from affected UEs in bulk. Indeed, MMEs can handle only a limited number of signaling messages per second, thus eNBs could hold back signaling messages, i.e. TAUs from UEs in

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE Globecom 2011 proceedings.

order to aggregate them and send them in one single message. MMEs can follow the same operation for several location update messages, being part of the TAU procedure towards the Home Subscriber Server (HSS). For example, MMEs may wait for a predefined timeout, till a number of location update requests arrive, or both to proceed with a bulk of location update requests towards HSS.

TAU messages consist of 15 octets of mandatory fields and a set of optional fields. Since we concentrate on UEs associated with the same MME, the only UE specific parameter is the M-TMSI (MME Temporary Mobile Subscriber Identity), which identifies a device. The other fields, consisting 11 octets out of a total of 15 octets, are common to all UEs associated with the MME. Assuming N affected UEs associated with the same MME, it is possible to bulk N TAUs in a single message of (4 x N + 11) Bytes, while with individual messages (N x 15) Bytes are needed. Further aggregation towards the HSS could be made at the MME. Moreover, the effort of parsing the parameters of many messages is also reduced to a minimum, which also reduce the overall time spent for the procedure. Accordingly, signaling efficiency can be improved even if all Information Elements for the many original signaling messages would differ, just by avoiding the processing of multiple messages and by much more efficient parsing.

It is also of vital importance that eNBs are not congested with too many signalling messages from affected UEs. This can be achieved by scheduled paging. Indeed, bulk paging at MMEs or eNBs can be performed per specific groups of UEs, based on certain priority metrics, or in a randomized manner. Responses from UEs can be also carried out in a randomized manner and over a time interval following a hash function that takes UEs' unique identifiers (e.g., International Mobile Subscriber Identity - IMSI, subscription information available at UE, etc) as input values.

B. MME Restoration for Affected UEs in Active Mode

Regarding UEs in active mode, which are affected by a MME failure, the objective is to restoret their contextual information, without impacting the ongoing sessions. In contrast to existing high cost solutions that duplicate UE context information, our approach is based on more intelligent, cooperative behavior among network elements, which allows a considerably simpler and thus cheaper MME implementation.

In particular, a newly selected MME recovers the state information for the eNB (step 1 in Fig. 2) and S-GW (step 3 in Fig. 2). The state information recovered from the S-GW (in step 3) include per UE bearer information such as IMSI, Mobile Equipment Identity, S-GW TEID (Tunnel Endpoint Identifier) for S11/S4 interfaces, PDN-GW IP address and TEID for S5/S8 interfaces, eNB TEID for S1-u interface, PDN charging characteristics, EPS bearer QoS, etc. The state information recoverable by the new MME from eNBs (in step 1 in Fig. 2) include per UE, EPS bearers information (TEID and eNB IP address) and Aggregate Maximum Bit Rate (AMBR). Since EPS bearer information must be exchanged for many UEs, the information exchange between eNB/S-GW

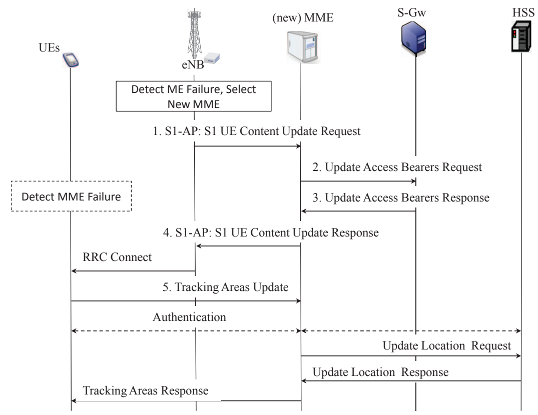

Fig. 2. eNB-initiated MME restoration for affected UEs in connected mode.

and MME (steps 1-4) can also be achieved by means of bulk signaling. The flowchart of the proposed solution is shown in Fig. 2. The mechanism is applied by each eNB being in a tracking area that was serviced by the failed MME. It concerns only UEs in connected mode registed with the failed MME. The steps of this solution are as follows:

1) eNB detects MME failure and selects a new one taking into account load balancing. MME selection can be performed for an individual active UE or for a UE set with common factors i.e. same S-GW and defined by a unique identifier (e.g., Connection Set ID allocated locally [3]). Prioritization among the UEs or UE sets can also be envisioned, i.e. intuitively UEs with imminent handoffs should be prioritized over other UEs.
2) eNB sends UE's S1 bearer information to the selected MME requesting a UE context update. Some of the provided context could be UE's IMSI, corresponding S-GW, Reason for Update, etc. A bulk of update requests can also be performed for each formed set of UEs.
3) MME then sends an Update Access Bearer request to the corresponding S-GW querying UE's S1 bearer information. MME, in turn, can also group UEs into different sets, uniquely and locally identified, and send a bulk of update bearer requests for each formed set.
4) In response, S-GW sends an Update Access Bearer Response. Here, the information on the corresponding PDN-GW can also be included.
5) As a confirmation, the newly selected MME responds with a S1 UE context update response to the eNB.
6) When UE detects MME failure (e.g., based on error message following an attempt to initiate a new PDN connection using old GUTI) or is triggered to perform TAU (e.g., by eNB via a RRC connect signaling message), it sends a TAU. MME relocation will then take place without impacting the user plane.

It should be noted that while in the above described flow, the TAU request is handled for each individual UE in active mode, the same bulk signaling handling described for UEs in idle mode could be applied [15].

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE Globecom 2011 proceedings.

IV. Performance Evaluation

Having described our proposed MME restoration scheme, we now direct our focus on its performance evaluation, using Matlab. We compare the proposed bulk signaling against the conventional scheme whereby all UEs are notified at the same time or are notified in a random manner. The radio access network (RAN) is represented as a graph G(V,E) with V nodes denoting eNBs and E edges indicating adjacency between neighbor eNBs, i.e. a UE may perform a handover among eNBs incident to a particular edge. The adopted RAN conforms a scale-free network paradigm, whose degree distribution follows a power law, representing a typical wireless network. RAN consists of 50 eNBs with every eNB having an upper bound capacity of 50 UEs. eNBs are associated with a pool of 5 MMEs that may handle up to 200 UEs each.

MMEs indicate their available capacity towards the associated eNBs and perform an update once a significant load change occurs by advertising a quantized weight. Thus, eNBs may perform load balancing to equally spread UEs among the pool of MMEs. In the simulation study, MMEs support four equally spaced update classes employing also a hysteresis mechanism, i.e. the update boundaries are [50,100,150] when the load is increasing and [25,75,125] when the load decreases. UEs enter the RAN following a Poisson distribution with an inter-arrival rate $\lambda = \{10,15,20,25,30\}$. Each UE camps in this specific MME pool area according to a random exponential distribution with a mean μ = 30 min and may reside at a certain eNB based also on an exponential distribution with a mean n = 10 min. The simulation initiates a random instant of the RAN and then starts considering the incoming traffic development associating arriving UEs to specific eNBs with a uniform probability. Each eNB selects a MME for each newly attached UE based on the advertised MME weights. The simulation develops for the first 50 minutes and then a randomly selected MME fails.

Once a failure occurs, the affected UEs perform a TAU update based on the following three schemes, namely, $i)$ with no intelligence TAU; whereby all UEs perform an update at the same time, $ii)$ with random notification; whereby UEs are notified when to perform an update based on a random uniform distribution with mean of 100ms, and $iii)$ using bulk signaling; whereby UEs are randomly notified as in $ii)$ and additionally eNBs perform bulk signaling. It should be noted that each eNB is assumed to be able to handle up to 20 TAU requests at the same time; additional TAU requests are dropped. The simulation process is repeated 100 times and the average signaling overhead associated with each different case is plotted in Fig. 3. Clearly with bulk signaling the communication overhead is reduced, especially as the inter-arrival rate increases, outperforming the other conventional schemes. Since bulk signaling avoids redundant information, new MMEs only need to process absolutely essential data reducing also the processing overhead. Accurate assessments of such processing overhead reduction depend on the specific MME characteristics, which may vary among different vendors. However, the reduction is envisioned to be proportional to the TAU overhead depicted in Fig.3.

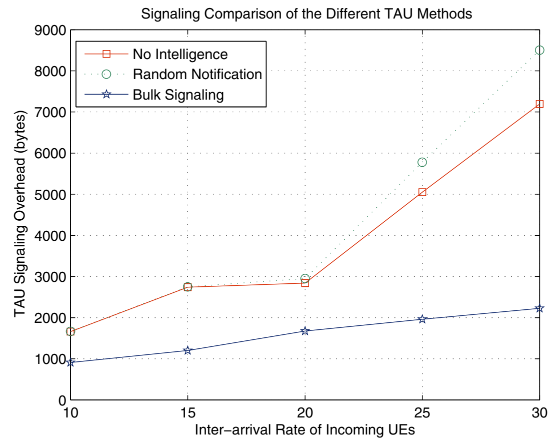

Fig. 3. TAU signaling methods with respect to communication overhead.

The “random notification” and the “no intelligence” approaches perform similarly for relatively low inter-arrival rates. As the inter-arrival rate increases, the “no intelligence” scheme produces less overhead because it results in drops of TAU at eNBs. Such drops are associated with the fact that UEs are notified instantly and simultaneously respond with TAU, overloading eNBs. It should be recalled that bulk signaling also adopts the random notification, avoiding TAU drop at eNBs. Considering, the transient period, i.e. the time duration immediately following a MME failure until the instant where all affected UEs are registered with a new operating MME; the three approaches exhibit a different behavior. The “no intelligence” one ensures the shortest transit period because of the instant notification, while in the other schemes that employ a random notification it primarily depends on the notification window, i.e. the time period during which an eNB notifies the affected UEs. Short notification windows may produce congestion, while longer ones may increase the transit period delay and ultimately impact the restoration latency.

Comparing the transient period of the “random notification” and bulk signaling approach we observe that the bulking signaling always results in higher transit times as illustrated in Fig. 4. Effectively, the “random notification” allocates a TAU time to each affected UE following a uniform distribution within a 100ms interval, thus the maximum transient period is on average below 100ms. The bulk signaling propagates each bulk either once 20 TAU messages are received at an eNB or when 20ms are elapsed since the last bulk message was sent. It should be noted that short bulk wait intervals result in shorter restoration latency but may cause higher overhead, especially when the duration of the failure transit window increases and vice versa. In addition, short failure transient times reduce the overall TAU delay at the cost of potential congestion at each involved eNB.

Once a MME failure occurs, TAU messages from affected UEs towards operating MMEs may overload some of them. Usually an overloaded MME forwards incoming UEs towards another less congested. For the sake of simplicity, in this

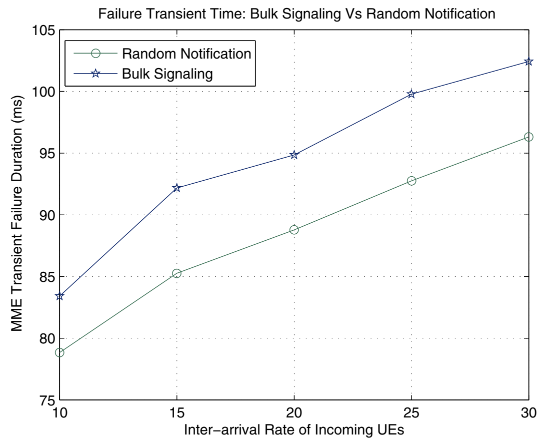

Fig. 4. MME failure transient duration for different UE inter-arrival rates.

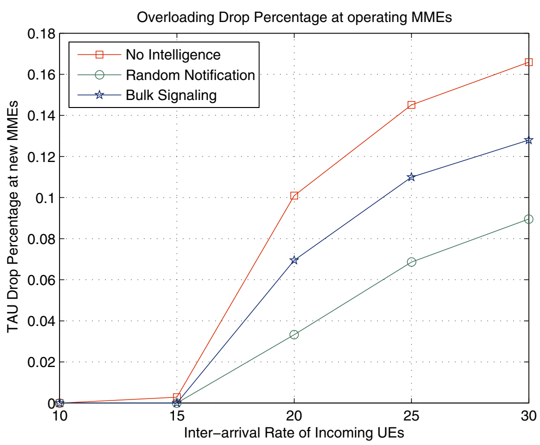

Fig. 5. TAU drop ratio at new MMEs.

study we dropped such UEs in order to compare the proposed different failure recovery schemes. Fig. 5 depicts the MME drop ratio associated with the three schemes. In general, the weighted MME approach followed in this study provides a degree of load balancing among the pool of candidate MMEs. However, in case of "no intelligence" approach where all UEs instantly perform a TAU at the same time the choice of the optimal new MME for each UE may overload certain MMEs that advertised a high weight before the failure incidence. Thus with the "no intelligence" approach MME do not have adequate time to update their weight accordingly. In contrast, "random notification" provides enough time for MMEs to update their weight. Thus, eNBs have more accurate information for selecting the new MME. The bulk signaling approach share some similarity with the "random notification" but the TAU is performed in bulks. Thus, it provides enough time for MMEs to update their weight after each round of bulk processing. However, within each bulk update round MMEs that advertise lower weights may receive a higher amount of TAU bulks at the same time from different eNBs. For this reason the MME load balancing performance using bulk signaling is not as fine grained as in the case of "random notification". Intuitively, the drop ratio in case of the bulk signaling scheme depends on the size of the bulk; smaller bulk sizes may decrease such drop but may increase the overhead.

V. Conclusions

In this paper, we proposed a set of solutions to ensure service resiliency in EPS in case of a MME failure through proactive resilience mechanisms across nodes in the radio access and core network. In the proposed solution, upon a MME failure detection, concerned eNBs perform bulk paging of idle mode UEs affected by the MME failure, and characterized by the use of the failed MME information as identifier in the paging message and information relevant for overload avoidance. Overload of particular MMEs/eNBs is avoided via scheduled/prioritized paging and/or scheduled responses from UEs. To minimize the signalling load on the network and the processing load at the receiving end-points, concerned nodes aggregate signalling messages in a bulk fashion. The proposed MME restoration approach addresses UEs in both idle and active mode states. Service disruption is minimized and the service resiliency is accordingly ensured. Simulations are conducted using Matlab and the good performance of the proposed schemes is confirmed.

Acknowledgment

The work presented in this paper is supported by the European Commission, under Grant No. FP7-224619 (the ResumeNet project).

References

[1] 3rd Generation Partnership Project, "Non-Access-Stratum (NAS) protocol for Evolved Packet System (EPS)", TS 24.301, Work in Progress, V10.3.0, Jun. 2011.

[2] 3rd Generation Partnership Project, "Architecture enhancements for non-3GPP accesses", TS 23.402, Work in Progress, V10.4.01, Jun. 2011.

[3] 3rd Generation Partnership Project, "Restoration Procedures", TS 23.007, Work in Progress, V10.4.0, Jun. 2011.

[4] 3rd Generation Partnership Project, "Study of EPC Nodes Restoration", TR 23.857, Work in Progress, V1.3.0, Jun. 2011.

[5] D. Tipper, et. al., "Providing Fault Tolerance in Wireless Access Networks", IEEE Commun. Magazine, Vol.40, No.1, Jan 2002.

[6] S. Pack, T. You, and T. Choi, "Performance Analysis of Robust Hierarchical Mobile IPv6 for Fault Tolerant Mobile Services", IEICE Trans. on Communications, Vol.E87-B, No.5, May 2004.

[7] Y-B. Lin, "Per-user Checkpointing for Mobility Database Failure Restoration", IEEE Trans. on Mobile Computing, Vol.4 , No. 2, Mar/Apr 2005.

[8] J-W Lin and J. Arul, "An Efficient Fault Tolerant Approach for Mobile IP in Wireless Systems", IEEE Trans. on Mobile Computing, Vol.2, No.3, Jul/Sep 2003.

[9] J-W Lin and M-F Yang, "Fault-Tolerant Design for Wide-area Mobile IPv6 Networks", J.l of systems and Software, Vol.82, No.9, Sep 2009.

[10] 3rd Generation Partnership Project, "Study on PCRF failure and restoration," TR 29.816, V10.0.0, Oct. 2010.

[11] P. Calhoun, "Diameter Base Protocol", Network Working Group, RFC 3588, Sep. 2003.

[12] A. Kunz, T. Taleb, and S. Schmid, "On Minimizing SGW/MME Relocations in LTE", in Proc. ACM IWCMC'10, Caen, Jun. 2010.

[13] 3rd Generation Partnership Project, "E-UTRAN; Self-configuration and self-optimizing network use cases and solutions", TR 36.902, V9.3.0, Dec. 2010.

[14] 3rd Generation Partnership Project, "General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access", TS 23.401, Work in Progress, V10.4.0, Jun. 2011.

[15] T. Taleb and A. Kunz, "Machine Type Communications in 3GPP Networks: Potential, Challenges, and Solutions", to appear in IEEE Commun. Magazine.