# On Efficient Data Anchor Point Selection in Distributed Mobile Networks

Tarik Taleb*, Adlen Ksentini‡
*NEC Europe, Heidelberg, Germany
Email: talebtarik@ieee.org
‡IRISA/University of Rennes 1
Email: adlen.ksentini@irisa.fr

***Abstract*—Existing gateway selection mechanisms base their selection on gateway load and/or geographical proximity of users to the gateways. In this paper, we mainly argue the need for other metrics to improve the gateway selection mechanisms in distributed mobile networks. We therefore propose considering the end-to-end connection and the service/application type as two important additional metrics in the selection of data anchor gateways in the context of the Evolved Packet System (EPS). To enable this, two solution variants are proposed. Simulations were also conducted to evaluate the performance of the proposed solutions and encouraging results are obtained.**

## I. Introduction

Along with the ever-growing community of mobile users and the tremendous increase in the traffic associated with a wide plethora of emerging bandwidth-intensive mobile applications, mobile operators are facing a challenging task to accommodate such huge traffic volumes, far beyond the original network capacity. Given the relatively low increase in the Average Revenues per Users (ARPU), mobile operators are seeking economically feasible and cost-effective solutions for accommodating the increasing mobile network traffic with minimal investment into the existing infrastructure. An important solution that they consider is selectively offloading traffic as close to the Radio Access Network (RAN) assuming a decentralized architecture of the user plane of the mobile network deployment [1][2][3][4]. In such decentralized network, data anchor gateways are selected by the Mobility Management Entities (MMEs) based on only geographical proximity of the gateway to the user and/or their loads.

Fig. 1 depicts a shortcoming of current gateway selection function, which we aim solving in this paper. In this figure, the numbers nearby the links represent a routing cost associated with the link (e.g., hop count and bandwidth availability). At a certain point in time, a UE is connected to a server via PDN-GW1 (Packet Data Network Gateway, playing the role of a data anchor gateway). This PDN-GW1 could have been selected by MME based on geographical/topological distance to User Equipment (UE). Later on, the same UE wants to connect to a different server, namely target server. The UE has two options: *i*) either connect to target server via the current gateway, PDN-GW1, or *ii*) establish a new PDN connection via another PDN-GW, PDN-GW2 in the figure. From the figure, the route (UE - PDN-GW1 - target server) incurs a cost of 6, whereas the alternative route (UE PDN-GW2 - target server) incurs a cost of 5. Given the above, the paper addresses the following questions: shall the UE still use PDN-GW1 or shall it request a new PDN connection via PDN-GW2. In case of the latter, which shall intuitively save routing cost, how can the network and/or UE trigger the establishment of a new PDN connection?

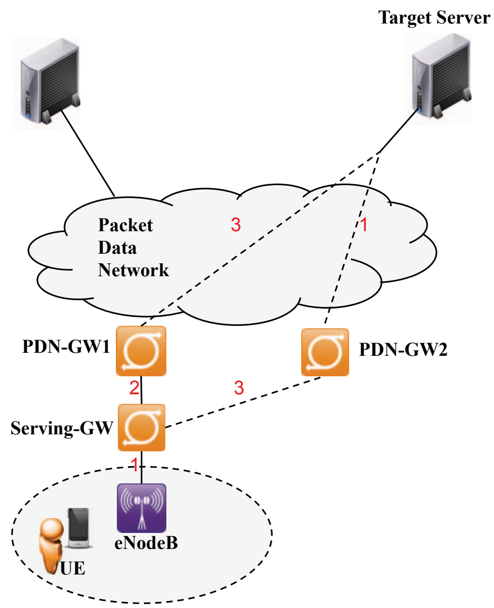


Fig. 1. Problem formulation.

As a solution to the above-described problem, we propose that in addition to load balancing and geographical/topological proximity, network should select a data anchor gateway based also on end-to-end (E2E) connection. In addition to that, despite it is not reflected in the scenario of Fig. 1, application type is also an important parameter that shall be considered in the selection of data anchor gateways. Indeed, it is likely that certain functionalities (e.g., content caches) or services (e.g. Machine Type Communications or M2M services) may be available only via specific data anchor gateways (e.g., PDN-GWs). It could be also that a certain service is best experienced at the end-user only when it is anchored at a gateway with specific processing specifications and/or performing at a load below a certain threshold, which could be in general not

978-1-4673-3122-7/13/$31.00 ©2013 IEEE

critical for the overall performance of the gateway. So in case the network has to select between two gateways, one with fair load and another with low load, the one with fair load could be selected for applications that are less sensitive to gateway performance under loads, and applications that are more sensitive to loaded gateways will be assigned to the gateway with low load. Hence, the application type could also enhance the anchor point selection.

The remainder of this paper is organized as follows. Section II presents some state of the art work. Section III describes the different envisioned solution variants. The performance evaluation of the proposed solutions is conducted in Section IV. The paper concludes in Section V.

## II. State of The Art

The Evolved Packet System (EPS) is designed to encompass different 3GPP accesses (i.e., 2G, 3G and LTE) as well as non-3GPP access (e.g., WiMAX, CDMA2000, 1xRTT, etc.). The richness of EPC accesses gave birth to a new 3GPP entity called ANDSF (Access Network Discovery Selection Function) that assists UEs to find the best or most suitable access out of the many available ones [5]. Initially, the design of ANDSF in Release 8 was for helping UEs to select from non-3GPP accesses. Recent 3GPP standards activities, within Release 12, head towards adding functionalities to ANDSF and UEs to assist UEs in selecting among different available PDN connections [6].

In general, the variety of 3GPP and non-3GPP accesses in EPS has led to different interesting 3GPP study items whereby UEs are allowed to have simultaneous accesses to different networks using different access technologies. In [8], the 3GPP System Architecture working group investigated different possibilities for dynamic IP flow mobility between 3GPP and non-3GPP accesses. The study proposed allowing a UE, equipped with multiple network interfaces, to establish multiple PDN connections to different Access Point Names (APNs) via different access systems and to selectively transfer PDN connections between the accesses with the restriction that multiple PDN connections to the same APN shall be kept in one access. In [4], a solution is proposed enabling a UE to know how and when to establish a new optimized PDN connection for launching new IP sessions to a particular APN, without compromising the ongoing (old) PDN connections to the same APN. In [3], a solution is proposed to support Selective IP Traffic Offload (SIPTO) [1]. In this solution, the user plane of a mobile network is assumed to be decentralized and the objective is to enable a per-flow offload of certain IP traffic as near to the edge of the operator network as possible. This is achieved with the involvement of the mobile network's Domain Name Server (DNS) that informs a UE of the gateway to connect to for establishing a particular flow and that is upon making a DNS resolution request. The gateway selection is solely based on the geographical proximity of the UE to the gateway and the gateway load. In this paper, we argue that selection of gateways based on only geographical proximity and/or load is not sufficient. Indeed, considering the E2E connection and/or the application type is also vital.

As will be described later, our proposed solution also uses DNS for the selection of gateways taking into account E2E connection and/or application type. However, it shall be stated that unlike many "location-based" DNS resolution mechanisms that usually resolve a target host (given request from a source host), our proposed solution uses DNS to determine a suitable intermediate PDN gateway that is best positioned between UE as source host and a target host.

## III. Envisioned Solution

In this paper, we make two main assumptions. First, we assume that the mobile operator network is decentralized, i.e., data anchor gateways (such as PDN GW in case of EPS) are not placed in the same geographical location in a centralized fashion, but rather distributed over the mobile network coverage. In the second assumption, we assume that user equipment is capable of establishing multiple connections via different data anchor gateways, i.e., UEs supporting multiple APNs in 3GPP networks, and that is via the same or different wireless access technologies.

As mentioned earlier, in this paper, we propose that in addition to load balancing and geographical/topological proximity, network should select a data anchor gateway based also on E2E connection and/or application type, and that is by reflecting connection and application type in anchor point selection. Regarding the latter, it shall be mentioned that the presented solutions can benefit from the outcome of the Data Identification work conducted in 3GPP [5] and other relevant work attempting to define unique identifiers for known and widely used applications. A UE can be preconfigured with a list of such application names mapped to unique application IDs. This list could be regularly updated by the network via one or more suitable nodes such as ANDSF. As explained below, when a UE attempts launching an application, it sends the application ID to DNS, MME or ANDSF to get instructions on whether a new/optimal PDN connectivity shall be established to accommodate the traffic of the application.

In the following, we describe and discuss twosolution variants. In the description of the two solutions, we consider the scenario of Fig. 1. Whilst the following description could easily apply to the case when a UE wants to attach to the network and to get its first PDN connection, in the envisioned scenario, we assume that a UE is having one existing connection via a particular PDN-GW1. At a later instance, the UE wants to connect to a target server identified with a URL, e.g., www.server.com, and an IP address, IP@. A different PDN-GW, PDN-GW2, is assumed to be optimal for establishing the connection between the UE and the target server.

### A. Solution Variant 1

The flow of main messages exchanged in solution variant 1 is shown in Fig. 2. In solution variant 1, the UE sends a query to a name resolution server (e.g., DNS) indicating the target server's URL. In response, the name resolution (i.e., DNS) server provides the IP address of the server. It shall be

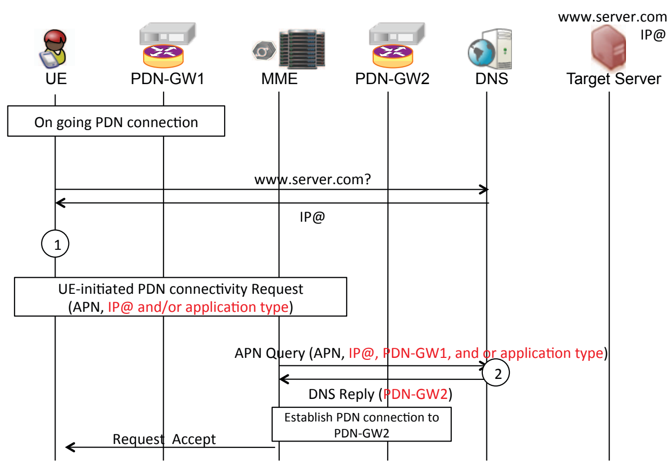


Fig. 2. Message exchange flow of solution variant 1.

noted that usually a UE would not issue a name resolution request for a server with an entry being still available in the UE's local DNS cache. This could impact the envisioned gateway selection mechanism. As a solution, the Time to Live (TTL) value of a DNS reply can be set to a small value. Upon receiving the IP address of the new target server, the UE runs a logic that decides whether the UE should consult MME for a new PDN connection to establish the new IP session to the target server (circle 1 in Fig. 2). Triggers for this logic could be based on the range of the server's IP address, content size, application type (e.g., video, email, audio, etc.), location and/or time difference since the setup of the current PDN connection, etc. If the UE decides to ask for a new PDN connectivity, it issues the "UE-initiated PDN connectivity Request" to the MME as part of the "UE-requested PDN connectivity" procedure in [10]. In the PDN connectivity request, the UE indicates the IP address of the target server to MME and/or the application type/ID. Alternatively, the UE could also send the URL of the target server, its Fully Qualified Domain Name (FQDN), or any other known identifier of the target server to the MME. In return, MME sends an APN query to DNS indicating the APN, IP@ of the target server, optionally PDN-GW1 – the current PDN-GW (or list of PDN-GWs) the UE is connected to, and/or application type. The DNS server is assumed to acquire a logic to decide an optimal anchor point based on UE location, target server location, and/or application type (e.g., video applications could be handled by PDN-GWs with cache support) (Circle 2 in Fig. 2). If the retrieved optimal anchor point is different than the currently used PDN-GW1, the DNS server inserts the new PDN-GW (PDN-GW2 in Fig. 2) in the DNS reply and sends it to MME. MME then uses this information and establishes for the UE the PDN connection to PDN-GW2. The procedure terminates with a request accept message as in the usual UE-requested PDN connectivity procedure [10]. Alternatively, in response to an APN query from the MME as described above, the DNS may suggest a list of anchor points (i.e., different than the current PDN-GW 1) ranked in order of preferences following a certain intelligent ranking scheme. This ranking method could be determined by the DNS or explicitly requested/specified by the MME. In turn, the MME follows the ordered list of suggested PDN-GWs to select the optimal PDN-GW for the requesting UE.

### *B. Solution Variant 2*

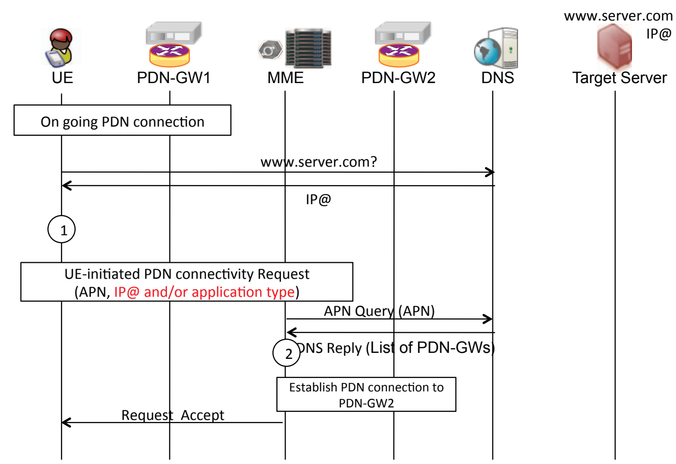


Fig. 3. Message exchange flow of solution variant 2.

The flow of main messages exchanged in solution variant 2 is shown in Fig. 3. Solution variant 2 resembles variant 1 till the APN query sent by MME to DNS. In this APN query, the MME indicates only the APN; and in response DNS replies with a list of appropriate PDN-GWs ranked, e.g., based on load information. This DNS query can be omitted if the MME is preconfigured with a list of suitable PDN-GWs for each APN and maintains a cache of previous APN query resolutions. After receiving a response from DNS, the MME runs a logic to decide the optimal anchor point from the list of PDN-GWs based on UE location, target server location, and/or application type (e.g., video applications could be handled by PDN-GWs with collocated caches). In case of solution variants 1 and 2, if the network decides that the PDN-GW does not need to change, the network, namely MME, sends a Request Reject message in response to the UE-initiated PDN connectivity request to indicate to the UE to establish the IP session to the target server via the current PDN-GW.

## IV. Performance Evaluation

### *A. Scenario*

In order to evaluate the proposed solutions, we used the NS3 simulator. The conducted simulations were run for 100s; a duration long enough to ensure that the system has reached its stable state. The simulated topology is shown in Fig. 4. We simulate 20 UEs, distributed uniformly over the coverage area of an eNB, simulated over a surface of 1000 x 1000 m2. All UEs are downloading an H.264 video stream at a rate of 320 Kbps from a remote server. The simulated 20 UEs start

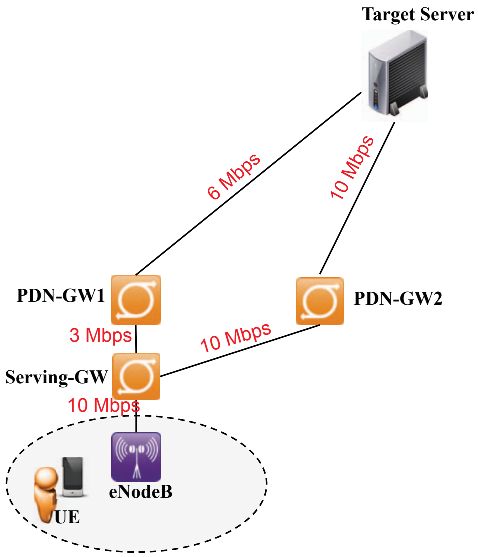


Fig. 4. Simulated topology.

downloading the video stream within the first 20s time interval of the simulation duration. When the proposed solution is used, all new connections are set up to the optimal data anchor gateway in terms of bandwidth, namely P-GW2 in Fig. 4. In contrast, when the conventional approach is used, the geographically nearest gateway, namely P-GW1 in Fig. 4, is selected.

## *B. Results*

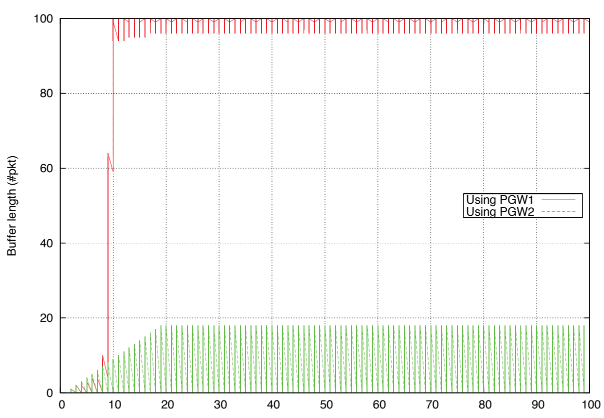


Fig. 5. The transmission buffer length at PGW1 and PGW2.

Fig. 5 shows the variation of the queue length of the gateways during the simulation.Queue length of P-GW1 refers to the conventional approach, whereas queue length of P-GW2 illustrates the benefits of the proposed approach when used. The obtained results are trivial as they demonstrate that a gateway selection mechanism based on only geographical proximity may result in several buffer overflows, followed by significant packet drops at the P-GW1 as shown in Fig. 6, and longer E2E delays and higher packet loss as depicted in Figs. 7 and 8 in case of an individual UE, UE1. This degradation in the network performance certainly impacts the QoE perceived at the UE, as shown in Fig. 9 in case of UE1. Note that UE1 is the first UE to establish a video stream session at the beginning of the simulation. The figures clearly indicate the benefit of using the proposed advanced gateway selection scheme in terms of high QoS perceived by the user. We note that the congestion level on the path from P-GW1 to the remote server increases along with the arrival of new session requests from other UEs, which ultimately increases the end-to-end delays for the UE1 session and causes it numerous packet losses. In contrast, when using our proposed gateway selection mechanisms, P-GW2 is selected ensuring low E2E delays and nearly no packet loss, which are ultimately translated into high QoE values as depicted in Fig. 9.

Indeed, Fig. 9 plots the instantaneous user QoE in terms of

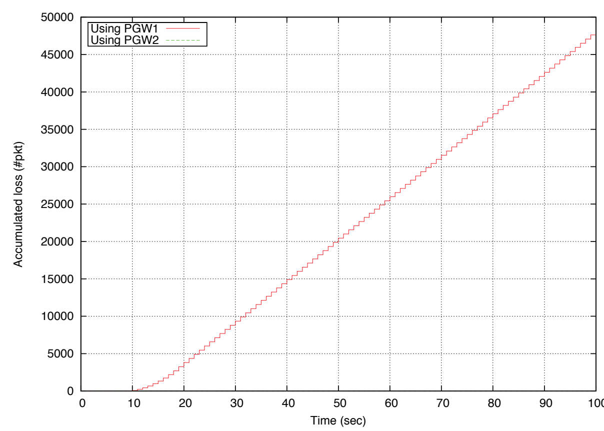


Fig. 6. The accumulated packet loss at PGW1 and PGW2.

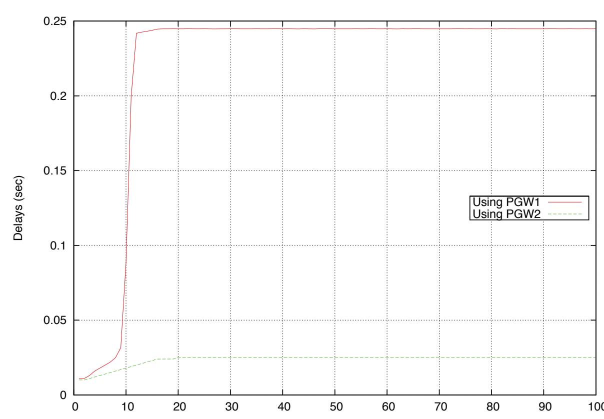


Fig. 7. Flow's packet delays for UE1.

the Mean Opinion Score (MOS). The MOS is a value between 0 and 10, representing the quality as perceived and given by users to a service. 10 and 0 represent the highest video quality and the worst video quality, respectively. These scores were obtained using the Pseudo Subjective Quality Assessment

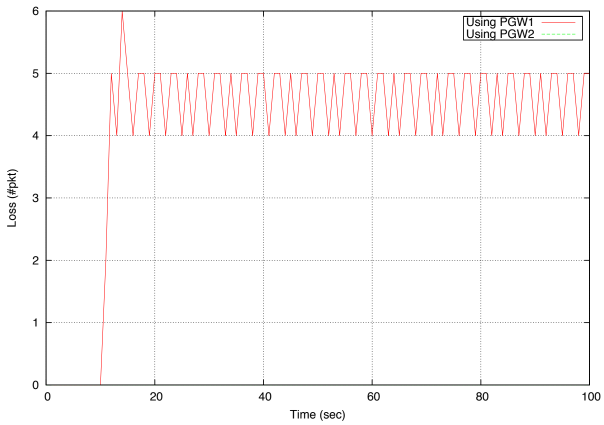


Fig. 8. Flow's packet loss for UE1.

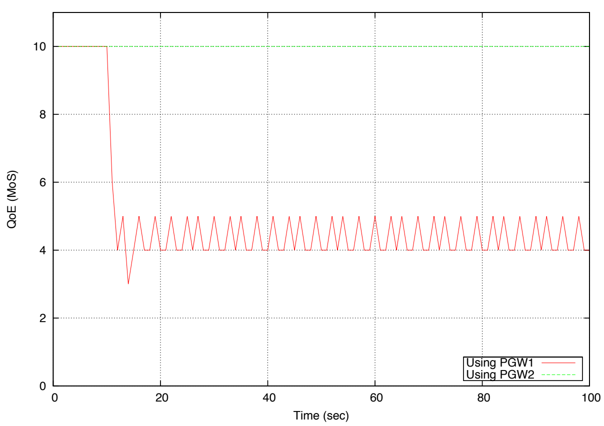


Fig. 9. QoE perceived by UE1.

(PSQA) tool [12] which is an automatic QoE evaluation tool for multimedia services based on Random Neuronal Network. It learns the non-linear relationship between parameters impacting the service quality and the user perceived QoE. It shall be noted that the PSQA version, used in the simulations, is dedicated to video quality evaluation. Fig. 9 shows clearly that the users QoE degrades heavily and frequently when the default gateway selection mechanism is used. In contrast, the figure shows that the proposed gateway selection mechanism ensures high values of MOS throughout the entire simulation.

## V. Concluding Remarks

In this paper, we demonstrated the need for considering other metrics than gateway loads and geographical proximity (of gateways to users) in gateway selection in distributed mobile operator networks. For that purpose, two solutions were proposed. The key feature of these solutions is that they consider the location/IP address of a mobile end host, the location/IP address/FQDN of a target corresponding node, application/service type/ID, task type (e.g., in case of MTC, emergency warning, delay tolerant measurement), and/or user class (i.e. user can be the mobile user or the owner/operator of the target node such as MTC server) in the selection of the optimal data anchor gateway to establish an IP session between the two end hosts, followed by optimal PDN connection establishment. Some of the benefits of the proposed solutions are manifested in efficient usage of the network resources, E2E connection and/or application type-oriented smart selection of data anchor gateways, and E2E route optimization along with the associated energy savings.

## References


[1] K. Samdanis, T. Taleb, and S. Schmid, "Traffic Offload Enhancements for eUTRAN", in IEEE Communications Surveys and Tutorials journal, Vol. 11, No. 3, Aug. 2012, pp. 884-896.

[2] T. Taleb, K. Samdanis, and F. Filali, "Towards Supporting Highly Mobile Nodes in Decentralized Mobile Operator Networks," in Proc. IEEE ICC 2012, Ottawa, Canada, Jun. 2012.

[3] T. Taleb, K. Samdanis, and S. Schmid, "DNS-based Solution for Operator Control of Selected IP Traffic Offload," in Proc. IEEE ICC, Kyoto, Japan, Jun. 2011.

[4] T. Taleb, Y. Hadjadj-Aoul, and S. Schmid, "Geographical Location and Load based Gateway Selection for Optimal Traffic Offload in Mobile Net-works," in Proc. IFIP Networking, Valencia, Spain, May 2011.

[5] 3rd Generation Partnership Project, "Access Network Discovery and Selection Function (ANDSF) Management Object (MO)," 3GPP TS 24.312 V9.1.0, Mar. 2010.

[6] 3rd Generation Partnership Project, "Operator Policies for IP Interface Selection (OPIIS)," 3GPP TS 23.853, V. 0.4.0, Jun. 2012.

[7] 3rd Generation Partnership Project, "Data Identification in Access Network Discovery and Selection Function (ANDSF)," (DIDA), 3GPP TR 23.855.

[8] 3rd Generation Partnership Project, "Multi Access PDN connectivity and IP flow mobility," 3GPP TR 23.861 V1.3.0, Feb. 2010.

[9] 3rd Generation Partnership Project, "IP flow mobility and seamless Wireless Local Area Network (WLAN) offload; Stage 2," 3GPP TS 23.261, Jun. 2010.

[10] 3rd Generation Partnership Project, "General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access," 3GPP TS 23.401 V11.2.0, Jun. 2012.

[11] T. Taleb and A. Kunz, "Machine Type Communications in 3GPP Networks: Potential, Challenges, and Solutions," in IEEE Communications Magazine, Vol. 50, No. 3, Mar. 2012.

[12] G. Rubino, 'The PSQA project", INRIA Rennes, URL: www.irisa.fr/armor/lesmembres/Rubino/myPages/psqa.html.