

Implementing a Session Aware Policy Based Mechanism for QoS Control in LTE

Zouhair Bazzal, AbdelMehsen Ahmad, Ibrahim El Bitar, Ali Ghouwayel, and Hussein Hijazi

School of Engineering - Department of Computer and Communications Engineering Lebanese International University (LIU) P.O. Box: 146404 Mazraa, Beirut, Lebanon

ABSTRACT

Quality of Service (QoS) provisioning has become significant with the widely growth of multimedia applications and high increase in the number of users in both wireless and wired networks. In this paper, we implemented a session-aware policy based mechanism for QoS provisioning and control in LTE (Long Term Evolution) networks. Policies are a set of rules identifying the QoS parameters for users. Implementation included DiffServ (Differentiated Services) configuration and setting policies inside the PCRF (Policy Charging Rules Function) which is the brain entity in LTE, then mapping from QCI (QoS Class ID) to DiffServ. Moreover, the dialogue between PCEF (Policy Charging Enforcement Function) and PCRF was implemented. Simulations on four different traffic application types: VoIP (voice over IP), video, web, and ftp (file transfer protocol) were performed under the network simulator (ns2). Results showed that applying PCEF over the different traffic applications has a great effect in controlling these applications and specifically UDP (User Datagram Protocol) based applications such as video.

Index Terms: QoS, LTE, PCRF, PCEF, EPS.

I. INTRODUCTION

Multimedia applications have been widely used in the recent wired and wireless networks [1]. They include advanced features such as high rate, timing constraints, and high computations. Developing new multimedia applications requires specific network resources. The biggest challenge is for the wireless networks where resources are rare and shared among users. Thus, the complexity of multimedia networking is a natural result of integrating a diversity of the traffic patterns and the performance requirements of different multimedia sources (video, voice, and data) over a single communication channel. Thus, the QoS provisioning is being strongly required to support such applications. Foundations of QoS include throughput, reliability, timeliness, and perceived quality. The best utilization of resources in a network environment is only possible by first characterizing the traffic, then determining the required QoS parameters such as tolerated delay, jitter, packet loss ratio, and required bandwidth [2]. Hence, controlling the traffic is an efficient way to ensure that user traffic fit within his constraints. Hence to provision appropriate QoS, LTE operators [3] had to implement policy-based mechanisms that allow controlling the traffic so that user traffic fits within pre-defined constraints.

As users have different QoS profiles over the internet, providers are interested in assigning users controlled bandwidth according to their QoS

profiles so that gold users are assigned the maximum bandwidth they paid for. Thus, our main objective is to implement a policy based mechanism to control QoS in LTE networks; however, it is not a simple task. This includes mapping of QoS from LTE to QoS in DiffServ with setting the policies in the PCRF node and implementing the dialogue between the PCRF and PCEF which is the function that enforces policies. PCRF decides the level of resources to be assigned to users based on policy manner [4][5].

This paper aims to implement in ns2 a policy based mechanism to control QoS in LTE networks. To sum up the following contributions have been added:

- Create and configure the policies.
- Implement the mapping from QoS in LTE to QoS in DiffServ.
- Implement the PCRF and PCEF entities of LTE.
- Implement the dialogue between PCRF and PCEF.
- Apply the policies over a session.
- Extract and analyze results.

The paper is organized as follows. In Section II, we review the evolved packet system as proposed in LTE. Section III describes the implementation environment. Section IV presents the simulation results. Finally, Section V concludes the paper.

II. EVOLVED PACKET SYSTEM (EPS) IN LTE

A. Policy control concept

The policy control can be defined as an application of rules to tune access and usage for network resources. In the Evolved Packet System (EPS) of LTE, a network controlled QoS approach has been proposed, it aims to remedy the shortcomings of the previously adopted user initiated QoS architecture. This approach considers that the resource reservation can be not directly triggered by the subscriber, but by a brain entity PCRF to which the request has to be signaled. This entity has the only responsibility to specify the charging functions, and to decide of the level of resources that have to be reserved based on a policy oriented manner while taking into account the momentary resources available on the network elements [4][5].

The Evolved Packet Core (EPC) [6] is responsible for the whole control of the UE (User Equipment) and consists of three main logical nodes: packet data network (PDN) Gateway (PGW), Serving Gateway (SGW), and Mobility Management Entity (MME). Additionally, it contains other logical nodes and functions: Home Subscriber Server (HSS) and the PCRF.

The PCRF is responsible for policy control decision-making in which it analyzes how data flows are to be treated in the PCEF located in the PGW, and ensures that they meet with the users subscription profiles located in the HSS.

The PGW is a router located in the home domain of the user and directly interfaces with the SGW. It acts as a policy enforcement point, generates charging records at service level, and implements QCI to DSCP (DiffServ Code Point) mapping function. In addition, it is responsible of IP address allocation for the users, executes deep inspection and legal interception functions, and

also serves as the mobility anchor for interworking with non-3GPP. S-GW is a router located in the local domain of the user and connected to the eNB (evolved Node B). It acts as a policy enforcement point at the packet level for dynamic QoS policies, implements QCI to DSCP mapping functions, and performs some administrative functions in the visited network such as collecting information for charging and lawful interception. S-GW also serves as the mobility anchor for interworking with other 3GPP technologies.

B. Detailed LTE QoS model

QoS concept in EPS is class-based. In IP network terms, that means that there is a label (DSCP) which can be used for differentiation of services. The label in EPS is defined by QCI (QoS Class Identifier), which is a scalar that represents the QoS characteristics that the network elements are expected to provide for multiple SDFs (Service Data Flow) in an EPS bearer. In LTE, an EPS bearer corresponds to a one common QoS policy applied on the different sections of the network, this includes the UE to eNB section, the eNB to SGW section, and the SGW to PGW section as shown in Fig. 1. This means that all the SDF packets carried by the EPS bearer will be handled by applying the same scheduling algorithm, the same queuing policy and priority, the same RLC (Radio Link control) configurations on the radio access section, and the same policing and shaping mechanisms. Therefore, the user might need a separate EPS bearer if he/she requires different QoS levels. In TS23.401 (S5/S8 GTP-based mode), the EPS bearer terminates in the P-GW. The PCRF entity injects the decisions and policies to the P-GW, which will enforce their execution by applying them via a PCEF function through packet filters at the SDF (Service Data Flows) level within an EPS bearer [6].

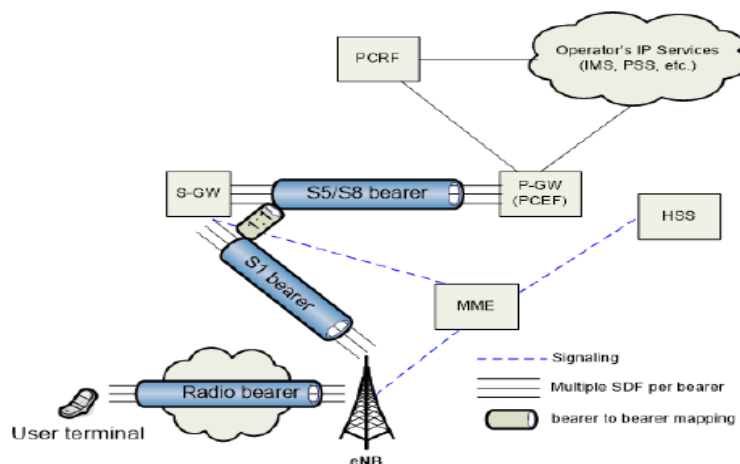


Fig. 1. EPS Bearers in TS23.401

The QCI will be mapped to a preconfigured QoS rules decided by the operator. It comprises the following elements:

- Bearer type: GBR (Guaranteed Bit Rate) or non-GBR, used to indicate whether the resources associated with that EPS bearer must be permanently allocated during the whole bearer lifetime (GBR), or the resources can be shared with other EPS bearers (non-GBR) during the bearer lifetime.
- L2PDB (Layer 2 Packet Delay Budget): used as an input to the layer 2 queue management system in order to control the maximum delay that a frame may reside within the link layer between the UE and the first hop (eNB).
- L2PLR (Layer 2 Packet Loss Rate): used as an input to the layer 2 configuration parameters in order to control the ratio of layer frames which have not successfully delivered to the eNB entity.

An EPS bearer is also characterized by an MBR (Maximum Bit Rate) which is used only for GBR EPS bearers, in order to limit the maximum bit rate of a GBR bearer. It works jointly with the GBR parameter. We can imagine an operator application (AF) sending a VoD (Video On Demand) stream which has a MBR greater than GBR, in such a way the operator can guarantee a GBR during the current network conditions, and a MBR (higher quality) where more network resources are available in the perfect conditions.

Finally, the EPS architecture has also introduced the concept of default bearers and dedicated bearers. A default bearer is only non-GBR EPS bearer, it is set up when the user attaches to the network, and remains established during the whole lifetime of the connection. This is the concept of the "Always ON" connectivity defined by EPS. Being "Always ON" contributes in eliminating the EPS bearer establishment delay once the user has attached the network for the first time. Any additional EPS bearer requested by the user or triggered by the network is called a dedicated bearer, which can be a GBR or non-GBR bearer.

III. IMPLEMENTATION ENVIRONMENT

A. Preliminaries and network model

Fig. 2 illustrates the targeted system model as well as the links parameters between the different nodes in terms of bandwidth and delay. Our main interest is in the packet core of LTE network. Recall that PCRF is the brain entity in LTE architecture; we implemented the network in a way the PGW can monitor the session type (VoIP, video, Web...) by inspecting the traffic that flows through it. For each new session, the PGW will ask

the PCRF for a new policy that the PCEF should apply over this specific session. The implementation ensures that the PCRF inject policies and decisions to the PGW which enforces them by PCEF through packet filters at the SDF level with an EPS bearer. We have created four types of traffic, VoIP, video, FTP, and web. Both FTP and web traffics are implemented over TCP, whereas VoIP and video traffic are implemented over UDP.

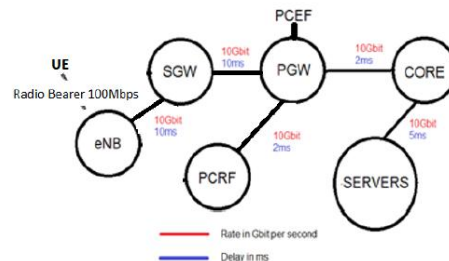


Fig-2 proposed system model

B. Implementation details

All implementations were conducted and developed in the network simulator (ns2) [7]. On the IP level, we assigned a set of DSCP [8] marking codes to the different types of traffic, then classified them into many queues as shown in Fig. 3. The LLQ (low latency queue) which is a combination of PQ (Priority Queuing) and CBWFQ (Class Based Weighted Fair Queuing) was also integrated. We classified the VoIP traffic in a PQ since it is not tolerable for delay and delay variations. The other types of traffic are classified into a CBWFQ with different weights (video of weight 0.4, web of weight 0.35, and FTP of weight 0.25). The total bandwidth is 1 Gbps; 85% of it is assigned to the CBWFQ, and the rest 15% is dedicated to the PQ (voice traffic). We defined four physical queues for the different types of traffic, the voice has 2 logical queues. The rest 3 queues (video, web, and FTP) each has 3 logical queues -3 drop precedences - (AF41-AF42-AF43, AF31-AF32-AF33, and AF21-AF22-AF23 respectively). We defined the applied policer and the shaping/dropping technique for the different types of traffic. For VoIP traffic, a token bucket policer is applied which uses CIR (committed information rate) and CBS (committed burst size) and two drop precedence EF (Expedited Forwarding) and AF42 (Assured Forwarding), that is an arriving packet is marked with the lower precedence (AF42) if and only if it is larger than the token bucket. For the rest of traffic, a two rate three color marker policer (trTCM) is applied that uses CIR, CBS, PIR (peak information rate), and PBS (peak burst size) to choose from three drop precedences. The shaping/dropping technique is Drop for VoIP and WRED for the rest of traffic.

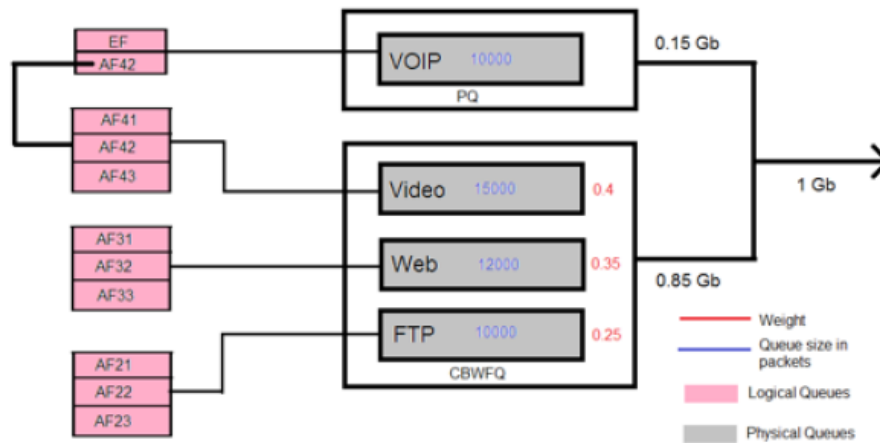


Fig. 3. Adopted QoS implementation at IP level

The main challenge was to implement under ns2 the QCI-DSCP mapping from eNB to SGW and from PGW to SGW and not vice versa as shown in

Fig. 4, the bearer type was only mapped from eNB to SGW and from PGW to SGW.

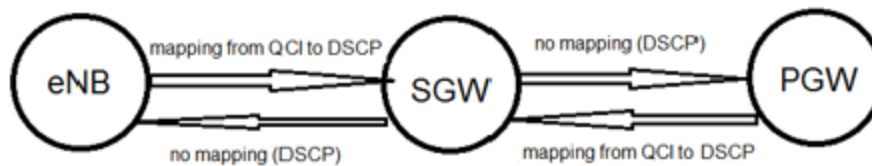


Fig. 4. Levels of mapping from QCI to DSCP

We configured the policies in the Service Profile Repository which is in our case located in the PCRF. In our implementation, we did not use all

the standardized QCI values; we only use five QCI values, and we mapped the DSCP marking to QCI as illustrated in Table I.

Table I Mapping DSCP marking to QCI

QCI	DSCP Marking
qci(0)	EF
qci(1)	AF41
qci(2)	AF31
qci(3)	AF21
qci(4)	BE

In our configuration each bearer has an SLA (Service Level Agreement); that is we cannot have two bearers with different SLAs and same QCI; in the following each bearer is assigned a

QCI, we have six bearers; control, default, VoIP, video, FTP, and web. For each bearer we can assign a new SLA that is QCI and MBR at UL (Uplink) and DL (Downlink) as shown in Table II.

Table II. Bearers and their assigned parameters

SLA	Type of Traffic	QCI	MBR(UL)	MBR(DL)
sla_Bearer_Control	Control Bearer	qci(0)	unlimited	unlimited
sla_Bearer_Default	Default Bearer	qci(4)	1 Mbps	1 Mbps
sla(0)	VoIP	qci(0)	10 Kbps	10 Kbps
sla(1)	Video	qci(1)	50 Kbps	50 Kbps
sla(2)	Web	qci(2)	50 Kbps	50 Kbps
sla(3)	FTP	qci(3)	35 Kbps	35 Kbps

IV. SIMULATION RESULTS

Simulations were performed using the four different types of traffic. However, to validate the implementation results, we decided to show how the PCEF is applying the policies over ftp and video sessions (TCP and UDP traffic respectively).

After assigning sla(1) to the video traffic, the throughput and goodput are clearly different and they both decrease. As depicted in Fig 5, the throughput starts at 55 Kbps then decreases slowly to reach an almost constant value of 50 Kbps, whereas the goodput (measured at UE) starts at 33 Kbps and decreases to reach 4 Kbps. This decrease is due to applying a policy in the PCEF that ensures that the corresponding session does not exceed the allowed MBR.

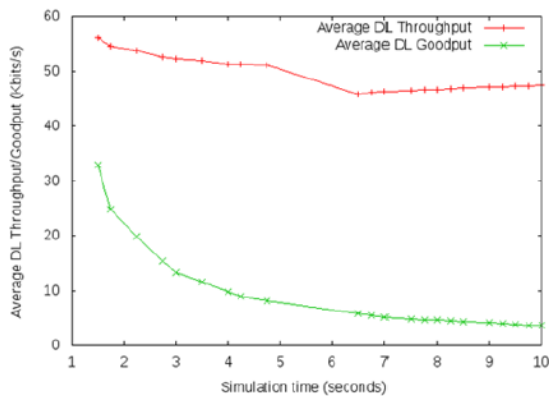


Fig. 5. Avg. DL throughput/goodput for video session at UE

In Fig. 6, the packet loss percentage increases to reach 90% at the end of simulation time interpreting the drop in the goodput and its difference with the throughput. Note that late packets will be discarded by the UE since they will be dropped by the PGW and will not be stored in the de-jitter buffer of the UE.

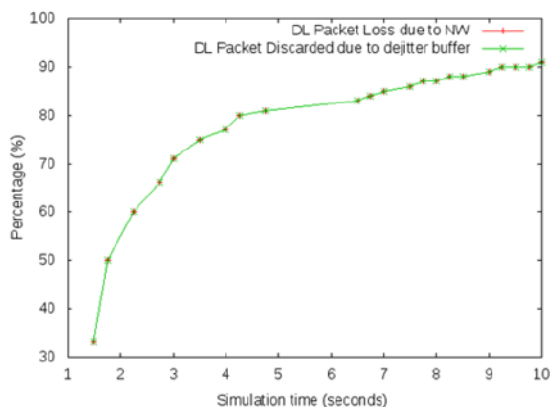


Fig. 6. DL packet lost/discarded rate at UE.

In Fig. 7, the average video delay starts at 500ms and increases to reach 4500ms which is considered bad. This affects the jitter and the packet loss rate which also increases drastically.

As for ftp traffic which was assigned an sla(3), Fig. 8 shows that the throughput and good put curves increase to reach a value of 32 Kbps which is set in the SLA. We can notice that the UE was able to receive all the ftp traffic. That is because the network was able to send the injected ftp traffic and to deliver all TCP segments as shown in Fig. 9 where the average retransmission rate measured at the ftp server was negligible.

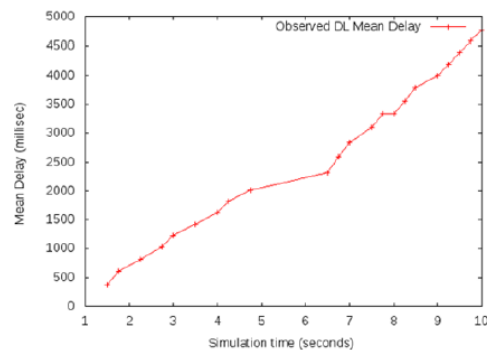


Fig. 7. Avg. video end to end delay in eNB

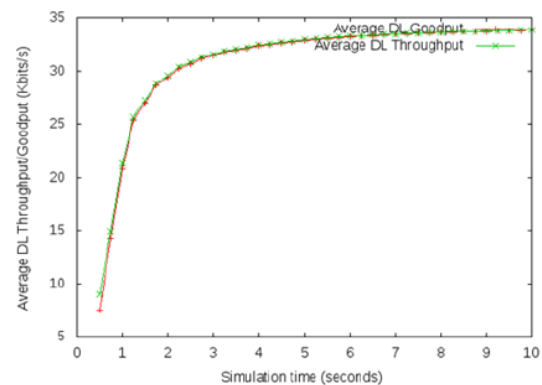


Fig. 8. Avg. DL throughput/goodput for ftp session at UE

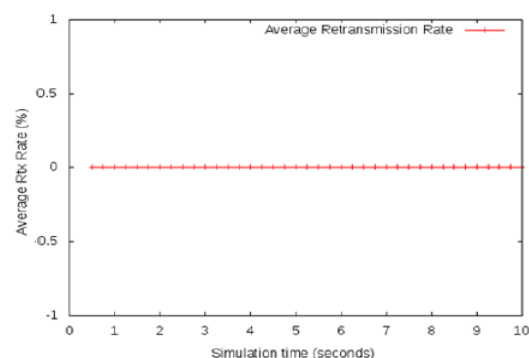


Fig. 9. Avg. retransmission rate at ftp server

V. CONCLUSION AND FUTURE WORK

As LTE networks involve advanced multimedia applications, controlling QoS based on policy mechanism is a significant challenging issue. In this paper, a policy based mechanism was implemented in ns2. It includes the mapping from QCI to DiffServ and the implementation of the dialogue between the PCRF and the PCEF in the PGW. Simulations were performed using four applications: VoIP, video, web, and ftp. Results have shown that the use of PCEF has a great influence on the UDP based applications (video). Hence, LTE Operators can rely on such mechanisms to control the QoS level of the sessions that consume their bandwidth depending on the assigned SLA.

As future work, we will integrate a dynamic policy based mechanism that is based on a prediction algorithm in which more than one policy can be assigned to users during a session. The algorithm will serve the PCRF to adapt the level of applied QoS based on accurate predictions related to the traffic behavior and network conditions.

REFERENCES

- [1]. "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015–2020", Cisco, [Online].
- [2]. T. Szigeti, C. Hattingh, K. Briley, R. Barton, "End-to-End QoS Network Design: Quality of Service for Rich-Media & Cloud Networks", 2nd Edition. Cisco Press, 2013.
- [3]. 3GPP LTE: <http://www.3gpp.org/LTE>, [Online].
- [4]. 3GPP TS 23.203 v.14.2.0, "PCC Architecture," 3GPP (Release 14), Dec 2016.
- [5]. 3GPP TS 29.213 v. 14.2.0, "PCC signalling Flows and QoS Parameter Mapping ," 3GPP (Release 14), Dec 2016
- [6]. 3GPP TS 23.401 v.14.2.0, "GPRS Enhancements for E-UTRAN Accesses," 3GPP (Release 14), Dec 2016.
- [7]. Information Sciences Institute. [Online]. <http://www.isi.edu/nsnam/ns/tutorial/>
- [8]. S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "An architecture for differentiated services," RFC 2475, Internet Engineering Task Force, Dec. 1998.