Multi-QoS-aware Fair Scheduling for LTE

Yasir Zaki[1], Thushara Weerawardane[1], Carmelita Görg[1], and Andreas Timm-Giel[2]

[1]TZI ComNets, University of Bremen, Otto-Hahn-Allee NW1, 28359 Bremen, Germany, {yzaki/tlw/cg}@comnets.uni-bremen.de
[2]Institute of Communication Networks, Hamburg University of Technology, 21073 Hamburg, Germany, {timm-giel@tuhh.de}

Abstract—**the MAC scheduler is an important and crucial entity of the Long Term Evolution (LTE) and is responsible for efficiently allocating the radio resources among mobile users who have different QoS demands. The scheduler takes different considerations into account such as throughput and fairness when deciding the allocation of the scarce radio resources. LTE is an all IP packet system in which guaranteeing QoS is a real challenge. Therefore the LTE MAC scheduler should consider not only the throughput optimization but also the QoS differentiations in an effective manner. In this paper, we propose a novel LTE downlink MAC scheduling algorithm. The proposed scheduler differentiates between the different QoS classes and their requirements. Two different QoS classifications are considered: Guaranteed Bit Rate (GBR) and non Guaranteed Bit Rate (non-GBR). The proposed scheduler also considers the different users channel conditions and tries to create a balance between the QoS guarantees and the multi-user diversity in a proportional fair manner. The simulation analysis confirms that guaranteeing the different QoS requirements is possible.**

Keywords-component; LTE; simulation model; OPNET; QoS aware scheduler

I. Introduction

LTE is the latest standard of the Third Generation Partnership Project (3GPP). It is also known as the fourth generation mobile network (4G) or 3GPP release 8. The LTE specification started at the 3GPP Radio Access Network (RAN) Evolution Workshop in Toronto, Canada in November 2004. The goals of the 3GPP was to design a system that can surpass the older mobile standards (e.g. UMTS, HSPA …) and that can still stay competitive at least for the next 10 years.

LTE uses a simplified flat network infrastructure that consists of only two nodes: the enhanced NodeB (eNB) and the mobile management entity/serving gateway (MME/S-GW). This is also one of the main factors that LTE can achieve a reduced latency compared to UMTS/HSPA. Another design goal was to increase spectrum efficiency. LTE uses a combination of multi-antenna techniques and Orthogonal Frequency Division Multiplexing (OFDM) to achieve higher data rates and offer the required spectrum deployment and flexibility.

LTE MAC scheduling is one of the key elements of the overall system performance. There are many publications that widely address this topic. In [1] the advantages of the joint time and frequency schedulers are presented. In [4] it is shown that for VoIP traffic dynamic scheduling is preferred over semi-persistent scheduling. [2] and [3] have made different proposals for time and frequency domain scheduling, pointing out the advantages and disadvantages of each of them. In [5] a principle of metric decoupling between time and frequency domain schedulers is investigated, showing to provide coverage gains of up to 60% for a cell throughput loss of 5% over the well known time-frequency domain proportional fair scheduler for Best Effort (BE) and Constant Bit Rate (CBR) users. The authors in [6] show that the Differentiated Services (DiffServ) polices within a scheduler can provide an effective way of dividing bandwidth between QoS classes. In [7] a scheduler with a VoIP priority mode is proposed to satisfy the QoS requirements, keeping the overall system degradation to a minimum. In this paper, a service aware MAC scheduler algorithm that defines five new QoS classes within the MAC layer is proposed. The scheduling is performed taking into account the different QoS classes as well as the different channel conditions. The goal is to perform a fair QoS aware scheduling together with cell throughput optimization.

II. LTE System Architecture and Protocols

Research of the 3GPP has resulted in the development of the Evolved Packet System (EPS). The EPS consists of: Core Network (CN), Evolved Packet Core (EPC), and Evolved UTRAN (E-UTRAN), it is also referred to as LTE. The architecture in EPC/LTE is much simpler than in UTRAN release 6. The EPC consists of one control plane, which is called Mobility Management Entity (MME) and two user plane nodes, which are called serving gateway (S-GW) and Packet Data Network Gateway (PDN-GW). The LTE radio access network consists of the enhanced NodeBs (eNB). The eNBs are connected with each other through the X2 interface and with the EPC through the S1 interface. An eNB may be served by more than one MME as can be seen in Figure 1.

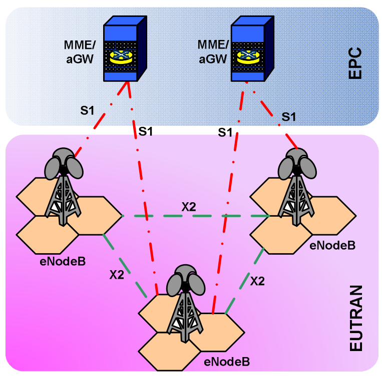

Figure 1: E-UTRAN architecture

Figure 2 shows the end-to-end user-plane protocol architecture from the application layer of the LTE mobile terminal to the application layer of the remote server. The complete architecture uses IP as the transport network protocol. Depending on the services, different end user protocols are involved during data transmissions between end-nodes.

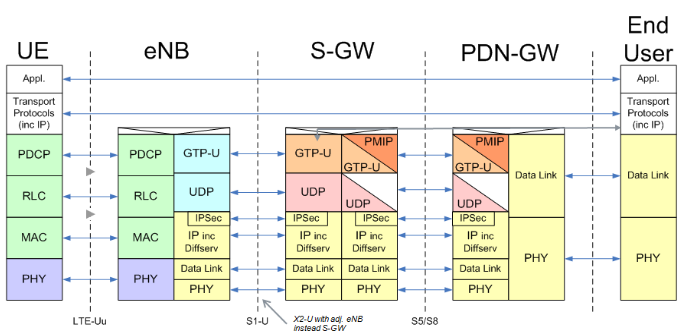

Figure 2: LTE user-plane protocol architecture

978-1-4244-8331-0/11/$26.00 ©2011 IEEE

III. LTE TRANSPORT SIMULATION MODEL

The design overview of the OPNET LTE simulation model can be seen in Figure 3. Each eNB can be configured with a variable numbers of users. The simulation model can be divided into a number of nodes that carries the LTE protocols and functions.

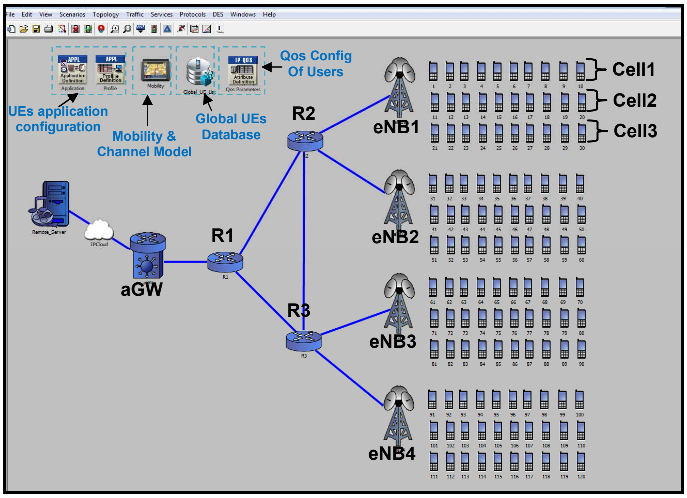

Figure 3: OPNET LTE simulation model

The link between the remote server and aGW is configured as an Ethernet link which emulates an average delay of 20 ms between these two nodes. The aGW and eNBs are connected through the transport network which consists of three IP routers named R1, R2 and R3. All these routers are configured with standard OPNET diffserv models and routing protocols. Standard QoS configuration for traffic differentiation at the transport network is set at the QoS configuration node (upper right node, is also called QoS parameters) and is then linked to the corresponding router which has to be activated for service differentiation.

A. Mobility and Channel Model Node

This node includes the implementation of the mobility and channel models. The mobility model emulates the user movement within the cell by updating the user location at every sampling interval which has been adaptively configured according to the UE speed. Currently two mobility models are implemented: Random Way Point (RWP) and Random Directional (RD). All UE mobility information is stored and updated regularly in the global user database shown in Figure 3 where it can be accessed at any time by any node in the system.

The channel model includes the well known factors: path loss, slow fading and fast fading. By using the link budget the different Signal to Interference plus Noise Ratio (SINR) per each Physical Resource Block (PRB) is calculated for each connection. This represents the user channel conditions that differ between the different PRBs due to the frequency and time selectivity. The path loss is calculated as follows [13]:

$$L = 128.1 + 37.6 Log_{10}(d) \qquad (1)$$

The slow fading is typically modeled using a log normal distribution with zero mean and a variance, but the time correlation between the slow fading values needs to be considered. For such a model, consider a moving mobile user starting at an initial point P where the slow fading value is to be randomly generated using the log-normal distribution equal to S (0). The shadowing at points which are at distance δ, 2δ, 3δ away from P, can be determined according to [8] as follows:

$$S(n\delta) = e^{-\delta/Xc} \cdot S((n-1)\delta) + V_i \qquad (2)$$

Where the V_i are independent and identically distributed normal random variables with zero mean, and a variance $\sigma_2^2=\sigma^2$ [1-exp (-2δ/X_c)] in dB. X_c is the de-correlation distance.

As for the fast fading, the implementation of the ChSim [9] was used to model the Doppler spread for the time selectivity and the delay spread for frequency selectivity.

B. UE Node

The UE advanced node model is created according to the layered hierarchy as shown in Figure 4. It includes two sets of protocols: LTE air interface related protocols and end user application protocols. The LTE Uu protocols include PDCP, RLC, MAC and PHY layers. Except for the PHY layer protocol, all other protocols have been implemented according to the 3GPP specification [10]. The physical layer is modeled just to transmit the transport blocks of user data between eNB and UE entities. The enhanced MAC scheduler is designed in order to emulate the PHY layer characteristics.

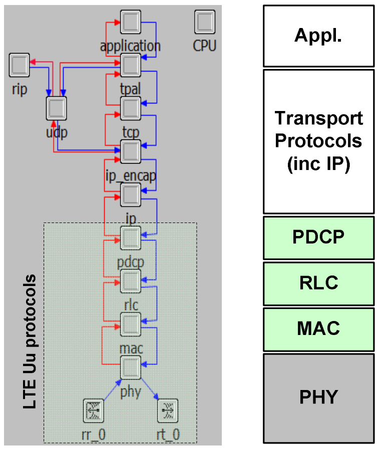

Figure 4: UE advanced node model

C. eNB Node

The protocol layers of the eNB advanced node model are shown in Figure 5. The right hand side protocols of the node model PDCP, RLC and MAC, this represent the Uu related functionalities. Whereas the left hand side protocols of the node model GTP, UDP, IP and Ethernet this represent the transport related functionalities. The MAC scheduler is located at the MAC layer in the eNB.

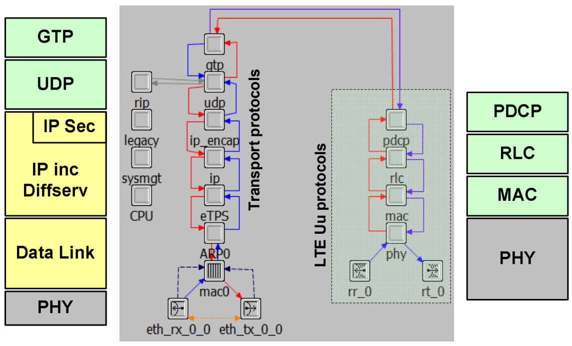

Figure 5: eNB advanced node model

IPsec and the service differentiation are performed at the IP layer in the eNB. The diffserv model is implemented in the eNB advanced node model. The model includes traffic differentiation and scheduling in the IP layer as well as shaping functionalities at the Ethernet layer.

D. aGW Node

The aGW user plane transport protocols are shown in Figure 6. One part is the peer to peer transport protocols: GTP, UDP, IP and Ethernet, these are located towards the transport network; whereas the other part is located towards the remote server side. In addition to the diffserv model, the IPSec functionality is also implemented within the IP layer in both aGW and eNB node entities in order to provide data security over the transport network.

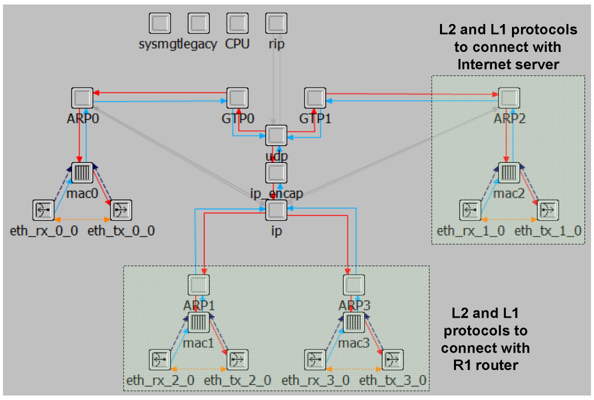

Figure 6: aGW advanced node model

E. Global UEs Database

This entity holds all of the user's information. This information can be accessed from anywhere in the simulation model. For example, the mobility related information like for example the user position, channel conditions … etc. are all stored within this node, and the mobility and channel model entity only access this information and update it.

IV. MAC Scheduler

The MAC scheduler is responsible for scheduling the air interface resources among the users in both the downlink and the uplink. Since OFDM technology is used in LTE, the scheduler effectively distributes the radio resources in both time and frequency domain. The smallest scheduling resources unit is called Physical Resource Block (PRB). In order to simplify the LTE MAC scheduling, two stages have been defined: Time Domain (TD) and Frequency Domain (FD) schedulers as shown in Figure 7. The TD scheduler is used to differentiate the users according to their QoS characteristics. The FD scheduler is responsible for assigning the radio resources (i.e. PRBs) between the priority users.

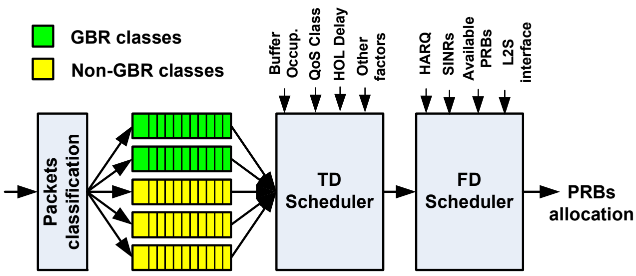

Figure 7: General LTE MAC scheduler framework

The proposed MAC scheduler considers two main types of QoS bearers: Guaranteed Bit Rate (GBR) and Non Guaranteed Bit Rate (non-GBR). Based on the QoS Class Identifier (QCI), the incoming packets are categorized upon their priority order (see Table 1) into five different MAC QoS classes that are defined by the MAC scheduler. The two highest MAC QoS classes (1 and 2) represent the GBR bearers, whereas the other three represent the non-GBR bearers. The MAC scheduler performs strict priority scheduling for all GBR bearers (i.e. served first) and then scheduling the non-GBR bearers.

Table 1: QCI to MAC QoS class mapping

	Service Type (QCI)	MAC QoS Class
GBR	QCI1	M-QoS1
	QCI2	M-QoS2
Non-GBR	QCI7	M-QoS3
	QCI8	M-QoS4
	QCI9	M-QoS5

A. Time-Domain (TD) Scheduler

The TD scheduler creates two candidate lists (one for GBR and the other for non-GBR). The GBR candidate list is created by adding all bearers in MAC-QoS1 class in the top of the candidate list and then adding the MAC-QoS2 class bearers. As for the non-GBR bearer list, the TD scheduler adds the entire non-GBR MAC-QoS class bearers into the candidate list and then prioritizes them based on their priority factor as follows:

$$PF_0 = \frac{MQoS_{weight}}{R_{accum}} \tag{3}$$

Where PF_0 is the priority factor at iteration 0, $M\text{-}QoS_{weight}$ is the weight of the MAC QoS class, and R_{accum} is the accumulated data rate of that bearer from the past. R_{accum} can be calculated as follows:

$$R_{accum}(n) = \alpha \cdot R_{accum}(n-1) + (1-\alpha) \cdot R_{inst} \tag{4}$$

Where $R_{inst}(n)$ is the accumulated bearer data rate at the nth TTI, α is the smoothing factor, and R_{inst} is the instantaneous bearer data rate at nth TTI.

B. Frequency-Domain (FD) Scheduler

The FD scheduler starts with the GBR candidate list provided by the TD scheduler. The PRBs allocation is done iteratively, where one PRB with the highest SINR value is allocated for one bearer each iteration. The bearers in the candidate list get orderly (based on their priority) a chance to select the next best PRB upon the highest SINR value, this PRB allocation process continues until all bearers in candidate list get one PRB. At the end of each iteration the achieved data rate of each bearer is calculated and checked if sufficient data is available in the bearer buffer to be served or if the sufficient guaranteed rate is achieved for that particular bearer. In case, one of the above conditions is satisfied, the bearer is removed from the candidate list and scheduled..

The data rate for each bearer is calculated through the allocated PRB's SINRs, where the effective SINR value of all the allocated PRB(s) can be calculated by using the Effective Exponential SINR Mapping (EESM) [11]. Then it is compared against the target SINR value that is calculated from the Additive White Gaussian Noise (AWGN) Block Error Rate (BLER) curves (by setting a target BLER value e.g. 10%). If the effective SINR is lower than the target one, then the Modulation and Coding Scheme (MCS) is lowered, and the effective SINR is recalculated. Otherwise, using the 3GPP tables defined in [12], the Transport Block Size (TBS) is determined and it represents the data rate of the bearer for this TTI. Once all GBR bearers are scheduled, the FD scheduler starts scheduling the non-GBR bearers. The FD scheduler picks the highest N non-GBR bearers from the non-GBR candidate list to be served this TTI. The same iterative procedure used in the GBR scheduling is also used here with one exception: after each iteration the non-GBR bearers are re-prioritized using the newly calculated priority factor as given below:

$$PF_j = PF_{j-1} \cdot TBS_j \quad (5)$$

Where PF_j is the priority factor at the j^{th} iteration, TBS_j is the transport block size at the j^{th} iteration.

The reason why the re-prioritization of the non-GBR bearers is done between the iterations is mainly to provide more chances to the non-GBR bearer who has better channel conditions compared to others to select its best PRBs first.

V. SIMULATION SCENARIOS AND CONFIGURATIONS

In order to show how the LTE MAC scheduler serves and handles the mixture of the different QoS services, three different scenarios are investigated as seen in Table 2. Users with four different QoS services (VoIP, Video, HTTP and FTP) are defined. Some of them must be served with higher priority because of their real time nature. We investigate what happens if we mix all these services within one QoS class without distinguishing between the GBR and non-GBR, and what happens if we separate them. The rest of the simulation configurations can be seen in Table 3.

Table 2: Simulation scenarios

scenario	M-QoS1	M-QoS2	M-QoS3	M-QoS4	M-QoS5
Sc1	-	-	-	-	VoIP, Video, HTTP & FTP
Sc2	-	VoIP and Video	-	-	HTTP and FTP
Sc3	VoIP	Video	HTTP	-	FTP

Table 3: Simulation parameters

Parameter	**Assumption**
Number of eNBs	1 eNB with one cell of radius = 375 meters
Spectrum	10 MHz (corresponds to about 50 PRBs)
Mobility model	Random Way Point (RWP)
Channel model	ITU Veh-A
Number of users	20 VoIP, 5 Video, 40 HTTP and 20 FTP users
Scheduler parameters	α = 0.7 and N = 5 (non-GBR bearers scheduled per TTI) M-QoS3$_{weight}$ = 10, M-QoS4$_{weight}$ = 5 & M-QoS5$_{weight}$ = 1
VoIP traffic model	Silence length (sec): exponential (0.65) Talk Spurt length (sec): exponential (0.35) Encoder Scheme: GSM EFR Call duration: constant 90 sec Time between calls (sec): exponential (30)
Video traffic model	24 Frames per second Frame size: 1562 bytes
HTTP traffic model	Number of pages per session = 5 pages Average page size = 200,000 bytes Number of objects in a page = 1 Objects size (frame) = 1 Kbyte and 200 Kbytes Reading time (default) = 12 sec
FTP traffic model	Inter-request time (sec): uniform (30,60) File size: constant 10 Mbytes
Simulation time	1000 sec

VI. SIMULATION RESULTS

The average air interface throughput for VoIP users can be seen in Figure 8. The results show that the VoIP users have the same throughputs in all different scenarios even when mapped together with non-GBR bearers. The reason is the proportional-fair characteristic which is represented by the R_{accum} value in equation (3). Since the VoIP bearer has a relative smaller accumulated data rate it tends to get higher priority factor and will always be scheduled first.

The VoIP average end-to-end delay can be seen in Figure 9. It could be noticed that "Sc2" and "Sc3" have slightly better end-to-end delay compared to "Sc1"; this is due to the fact that the former scenarios map the VoIP bearers to a higher MAC QoS class (M-QoS2 and M-QoS1).

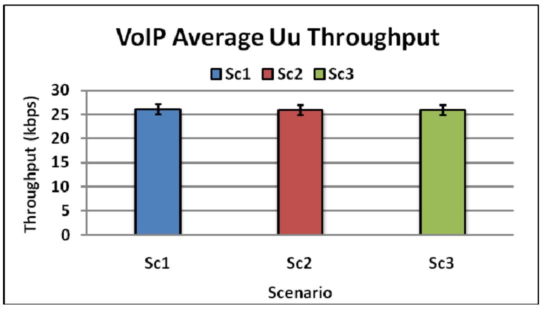

Figure 8: VoIP average Uu throughput

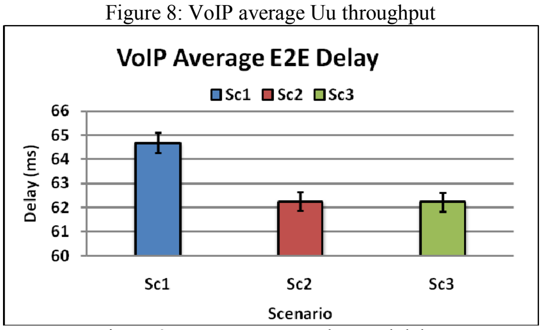

Figure 9: VoIP average end-to-end delay

Figure 10 shows the average air interface throughput for the video bearers. The result depicts that the video bearers have worse performance in the "Sc1" scenario compared to the other two scenarios where the video bearers are mapped into the GBR MAC classes. In the "Sc1" scenario, the video bearers share the same non-GBR MAC QoS class with VoIP, HTTP and FTP bearers. Since the accumulated data rate of the video bearers are significantly high (~ 300 kbps), they don't get served all the time.

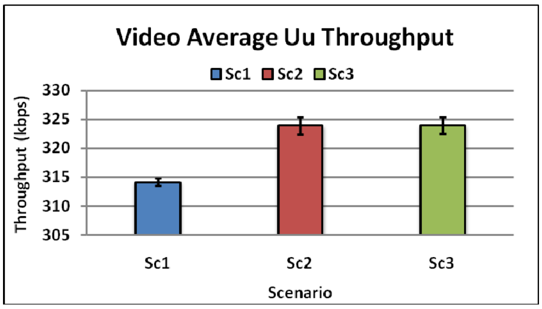

Figure 10: Video average Uu throughput

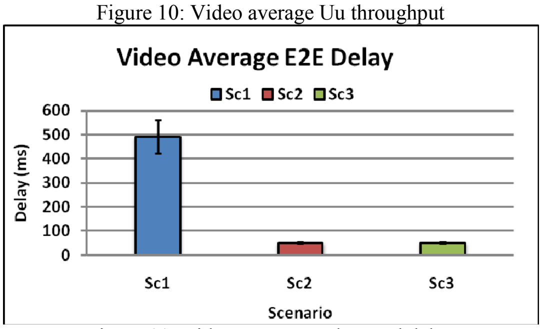

Figure 11: Video average end-to-end delay

The performance degradation of the video bearers in "Sc1" scenario can be clearly seen in Figure 11 with the average end-to-end delay. The video bearers suffer from significantly higher end-to-end delay compared to the other two scenarios where the video bearers get served with more strict priority. Looking at the HTTP bearers' results shown in Figure 12 and Figure 13, it could be noticed that the HTTP bearer has much better application performance when they are assigned on a higher MAC QoS class. Mainly when it's not mixed with the FTP bearers and is assigned to a higher MAC QoS class than FTP. This is because of the M-QoS$_{weight}$ in equation (3) which is set to be ten times higher compared to the other scenario.

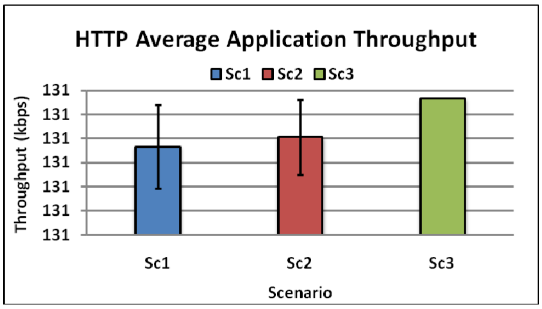

Figure 12: Http average application throughput

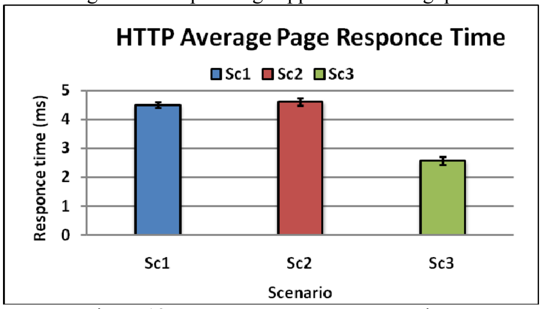

Figure 13: Http average page response time

Finally, the FTP bearer results are shown in Figure 14 and Figure 15. The FTP performance is reduced when going from fully mixed scenario (Sc1) to fully separate one (Sc3). Where the average throughput gets reduced and the average file download time increases. This is because the FTP bearer is mapped to the lowest MAC QoS class and is served with less priority as compared to the other services. Nevertheless, giving the FTP bearer lower priority is realistic because FTP is not a real time application and the FTP users are fine waiting a couple of more seconds for their files to be downloaded, whereas the same cannot be said for applications like VoIP.

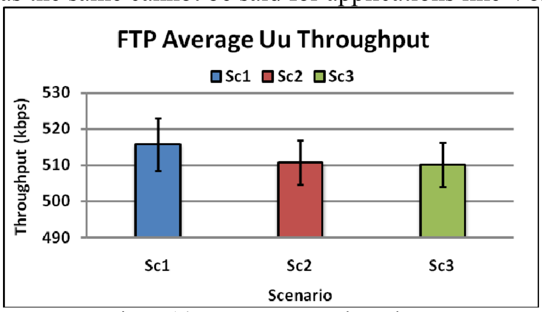

Figure 14: Ftp average Uu throughput

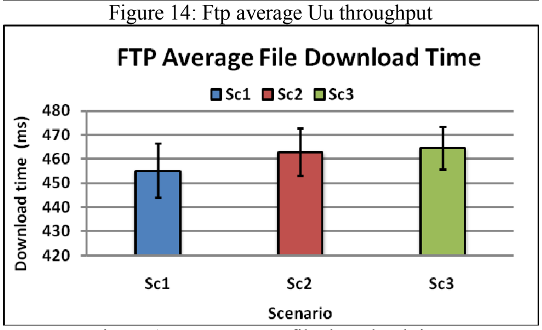

Figure 15: Ftp average file download time

VII. Conclusion and Outlook

In this paper, the impact of the multi QoS aware MAC scheduling on end user application performance has been studied. A QoS aware MAC scheduling algorithm has been proposed and implemented in the LTE downlink. The benefits of separating the bearers who are demanding different QoS requirements into GBR and non-GBR have been evaluated. Three different scenarios were investigated using simulations; each scenario represented a different service separation within the MAC scheduler. The goal was to show how to serve VoIP, Video, HTTP and FTP users while satisfying the different QoS requirements of each bearer. In the first scenario, all users have been assigned to one MAC QoS class with no prioritization. In the next two scenarios the GBR bearers (VoIP and Video) were separated from the non-GBR bearers (HTTP and FTP) and were given higher priority. The results showed that VoIP bearers survived even when they were mixed with the non-GBR bearers in one MAC QoS class. The reason was the proportional fair nature of the scheduler algorithm, because the VoIP bearers have relatively small data rate which inherently gave them higher priority. In contrast, the video bearers suffered significantly when mixed with the non-GBR bearers because of their higher data rate. The results also showed that since HTTP has a relatively lower data volume to transmit as compared to FTP, it is better not to mix with FTP in one MAC QoS class. Instead, it's better to assign the HTTP on a higher MAC QoS class than FTP.

References

[1] K. I. Pedersen P. E. Mogensen I. Z. Kovacs C. Rosa A. Pokhariyal, G. Monghal and T. E. Kolding. Frequency domain packet scheduling under fractional load for the utran lte downlink. In K. I. Pedersen P. E. Mogensen I. Z. Kovacs C. Rosa A. Pokhariyal, G. Monghal and T. E. Kolding, editors, Vehicular Technology Conference, 2007. VTC2007-Spring. IEEE 65th, pages 699-703, 2007.

[2] G. Monghal I. Z. Kovacs C. Rosa T. E. Kolding A. Pokhariyal, K. I. Pedersen and P. E. Mogensen. Harq aware frequency domain packet scheduler with different degrees of fairness for the utran long term evolution. In Vehicular Technology Conference, 2007. VTC2007-Spring. IEEE 65th, pages 2761-2765, 2007.

[3] S. A. Beh and A. Doufexi. Joint time-frequency domain proportional fair scheduler with harq for 3gpp lte systems. In Vehicular Technology Conference, 2008. VTC 2008-Fall. IEEE 68th, pages 1-5, 2008.

[4] E. Tuomaala, D. Jiang. Principle and performance of semi-persistent scheduling for voip in lte system. In International Conference on Wireless Communications, Networking and Mobile Computing, 2007. WiCom 2007., pages 2861-2864, 2007.

[5] I. Z. Kovacs G. Monghal, K. I. Pedersen and P. E. Mogensen. Qos oriented time and frequency domain packet schedulers for the utran long term evolution. In Vehicular Technology Conference, 2008. VTC Spring 2008. IEEE, pages 2532-2536, 2008.

[6] M. Gidlund and J-C Laneri. Scheduling algorithms for 3gpp long-term evolution systems: From a quality of service perspective. In IEEE 10th Inter national Symposium on Spread Spectrum Techniques and Applications, 2008.ISSSTA '08., pages 114-117, 2008.

[7] Y. Shin, S. Kang, B. Choi S. Choi, and K. Jun. Mac scheduling scheme for voip traffic service in 3g lte. In 2007 IEEE 66th Vehicular Technology Conference, 2007. VTC-2007 Fall., pages 1441-1445, 2007.

[8] C. Chuah, R. D. Yates, and D. J. Goodman, "Integrated Dynamic Radio Resource Management," in IEEE 45th Vehicular Technology Conference, vol. 2, Chicago, IL, USA, 25-28 Jul 1995, pp. 584 – 588.

[9] S. Valentin, "ChSim -- A wireless channel simulator for OMNeT++", TKN TU Berlin Simulation workshop, September 2006.

[10] 3GPP Technical Specification Group Radio Access Network, "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN): Radio Interface Protocol Aspects", 3GPP TR 25.813 v7.1.0, 2006.

[11] Y. Blankenship, P. Sartori, B. Classon and K. Baum, "Link Error Prediction Methods for Multicarrier Systems", IEEE VTC Fall , Los Angeles, 2004.

[12] 3GPP technical specification group radio access network, "Evolved Universal Terrestrial Radio Access (E-UTRA), Physical layer procedures (Release 9)", 3GPP TS 36.213 V9.3.0 (2010-09).

[13] Z. Wang, Y. Wang, D. Jiang, C. Tian, and D. Yang. "Scheduling and link adaptations for voip in tdd-lte uplink". In WiCom '09. 5th International Conference on Wireless Communications, Networking and Mobile Computing, 2009. ISBN: 978-1-4244-3692-7.