Application Aware Mechanisms in HSPA Systems

Péter Szilágyi and Csaba Vulkán
Nokia Siemens Networks
Budapest, Hungary
{*peter.1.szilagyi, csaba.vulkan*}*@nsn.com*

***Abstract*—Nowadays, the focus of network operation is no longer on the coverage and basic services but it is rather centered around the quality of experience that can be provided to the subscribers and on the capability of smoothly operating complex, interactive and increasingly data hungry applications on mobile platforms. Despite the increased system capacity, improved efficiency and sophisticated quality of service (QoS) architectures, the right level of quality of experience (QoE) requires application level differentiation. Using a single data bearer for each user equipment (UE), which is a common setup due to system and equipment limitation and the bearer centric QoS architectures, represents a barrier in providing true differentiation between simultaneously used applications. This paper discusses possible application level differentiation mechanisms either assuming a single data bearer per UE or utilizing the potential of secondary bearers to prioritize selected applications. The mechanisms were evaluated and compared by simulations focusing promoting web browsing over bulk data transfer in a High Speed Packet Access (HSPA) network. Results show that application differentiation mechanisms are able to significantly improve the quality of experience.**

***Keywords*-application awareness; HSPA; quality of experience; quality of service; simulation and modeling**

I. INTRODUCTION

The increasing prevalence of mobile devices with enhanced capabilities of running multimedia and web-based applications requires network-side evolution to fully serve the traffic demand. The nowadays spreading smart phones give access to the full spectrum of Internet-based applications already familiar from desktop computers, such as streaming multimedia, web browsing, mail, instant messaging, micro blogging, etc., encouraging the usage of multiple applications and services at the same time. A natural expectation of the users is to have reasonably good access to all applications even if they are run simultaneously, regardless of their different QoS requirements. However, despite the increased system capacity, high data rates and low latency provided by the evolved systems such as HSPA [1] and Long Term Evolution (LTE), there are still not enough resources in the mobile networks (especially considering the capacity limited last mile) to be able to smoothly support this user behavior without active QoS management on the network side. The end-user quality of experience greatly depends on how well the network is able to fulfill the QoS requirements of the applications [2]. Currently, due to network and equipment limitations, the entire data traffic of the users generated by the various applications is served by one data bearer. The QoS architecture is bearer centric, therefore all applications of the user receive the same service regardless of their different quality requirements; this makes it difficult for operators to enforce policies such as separately demoting bulk traffic or promoting premium services or applications. A possible solution is to use application aware mechanisms that are able to provide differentiation among the simultaneous applications run by the users. The requirement for application aware QoS has been raised not only in mobile networks [3] but in the context of transport network provisioning as well [4]. Research towards enhancing QoE is important not only in future network architectures such as LTE and beyond but also in HSPA networks, which today serve the vast majority of mobile broadband users.

In HSPA systems, bearers are used to deliver traffic according to a predefined set of QoS parameters over the radio access network (RAN) between the UE and the Radio Network Controller (RNC), referred to as the radio access bearer (RAB) service, and between the RNC and the core network (CN), referred to as the CN bearer service. A one-to-one mapping between RABs and CN bearers is done at the RNC to provide the Universal Mobile Telecommunications System (UMTS) bearer service [5]. At bearer setup, the UE can request certain QoS parameters such as guaranteed bit rate (GBR) or traffic class (TC). Based on that and operator policy settings, the Gateway GPRS Support Node (GGSN) determines additional parameters such as traffic handling priority (THP) and allocation/retention priority (ARP) and signals them to the RNC. The RAB specific QoS parameters, such as scheduling priority indicator (SPI) and discard timer (DT) are set by the RNC based on a mapping provided by the network operator or equipment vendor and signaled to the Node B along with the GBR [6]. The GBR parameter defines the target average bit rate that the air interface packet scheduler at the Node B should try to guarantee to the bearer. The SPI (an integer taking values from the range 0–15) specifies the priority of the data flow served by the bearer. DT gives the maximum allowed waiting time of the flow's packets (before being discarded) at the Node B buffers. These parameters are used by the Node B packet scheduler upon scheduling decisions. Once the active bearers receive their GBR, the packet scheduler is supposed to distribute the remaining air interface resources by considering the priority

ISBN: 978-1-61208-203-5

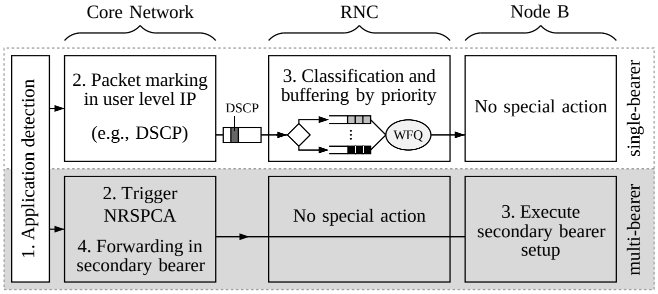

Figure 1. Concept of application aware mechanisms showing the entities involved in the discussed single- and multi-bearer alternatives.

of the bearers. An efficient packet scheduling discipline is the one implemented by the Proportional Fair with Required Activity Detection (PF-RAD) [7] scheme, that is capable of scheduling the bearers based on their weight. Accordingly, in addition to the GBR, a pre-configured parameter, the weight of the SPI class (wSPI) can be used, which is configured in the Node B for each SPI (not signaled from the RNC). Throughout this paper, we assume that the packet scheduling is based on the PF-RAD scheme. This QoS mechanism is not application aware as it is only able to differentiate among RABs but not among applications. In order to improve the situation, a couple of network-side techniques can be used, including single-bearer and multi-bearer mechanisms. Single-bearer means that the one data bearer per UE limitation is kept but the bandwidth available to the RAB is split between the applications in a predefined ratio (referred to as *in-bearer prioritization*), as proposed in [3]. Multi-bearer means to map the packets of applications with different QoS requirements to separate radio access and CN bearers, facilitated by the secondary PDP context activation procedure [8] standardized by the 3rd Generation Partnership Project (3GPP). This of course requires support from both device and network side.

In this paper, we discuss in-bearer prioritization and network requested secondary PDP context activation (NRSPCA) in detail, study their advantages and disadvantages and evaluate them based on web browsing user experience by conducting simulations in a HSPA network. The concept of single- and multi-bearer solutions is shown in Fig. 1. Results indicate that both mechanisms can considerably help enhance web page download performance; the apparatus required to implement the features can be the key differentiator in choosing the one selected for practical adoption in a real network.

The rest of this paper is organized as follows. In Section II, in-bearer prioritization is discussed. Section III provides an overview of NRSPCA and the related apparatus. In Section IV, the simulation models used in the performance evaluation of the proposed mechanisms are presented and Section V contains the simulation results and their interpretation. Finally, Section VI concludes the paper.

II. IN-BEARER PRIORITIZATION

The rationale behind single-bearer mechanisms is to maintain compatibility with such UEs and network-side implementations that are only capable of managing one data bearer per UE but still improve the QoE when different applications are simultaneously run by a user.

A plausible single-bearer mechanism capable of prioritizing traffic in the RAB is to mark packets in the CN according to the priority of the generating application and use per-priority packet buffering for each UE in the RNC Packet Data Convergence Protocol (PDCP) layer. The different buffers are served by Weighted Fair Queuing (WFQ) scheduler with configurable weights for the different priorities [3]. This feature requires an application detection facility in the CN, suitably in the GGSN as this is the node capable of intercepting packets arriving from external networks such as the Internet and investigate their application level content. One possible realization of application detection is Deep Packet Inspection (DPI), which can examine the TCP/IP headers of the user traffic and (or) apply pattern detection to recognize different applications. The result of the detection needs to be conveyed to the radio node where the RAB is originated, i.e., to the RNC, where the in-bearer prioritization takes place. The RNC is the best choice also as the next entity capable of accessing the application level data is the UE itself. Propagation of the application from the GGSN can be implemented by mapping the detected applications to priorities and marking the downlink (DL) packets accordingly, e.g., by utilizing the 6-bit DiffServ Code Point (DSCP) field of the inner IP header. The marking is thus encapsulated by GPRS Tunneling Protocol (GTP) and remains hidden from the transport mechanisms on the Gn and Iu-PS interfaces. For priority mapping, the following three levels may be used: middle priority for default traffic, i.e., traffic corresponding to the original QoS settings of the bearer; high priority for traffic to be prioritized; and low priority for traffic to be deprioritized. Generalization to additional priority levels is also possible.

In the RNC, the DL data packets are classified based on their priority marking and transferred to the corresponding per-priority PDCP buffer of the RAB. The amount of data sent from a given PDCP buffer to the RLC layer is determined by the WFQ mechanism and it is proportional to the weight of the buffer. The apparatus required by the mechanism is shown in Fig. 2. The solution is flexible as it allows the definition of different weight for each SPI class.

In-bearer prioritization can only be applied to the DL traffic as it is based on classification and WFQ scheduling mechanisms implemented at the RNC, where packets are multiplexed into the RAB. Such mechanisms are difficult to implement for uplink (UL) traffic as the other end of the bearer is at the UE. No other network element in between the UE and the RNC has access to the applica-

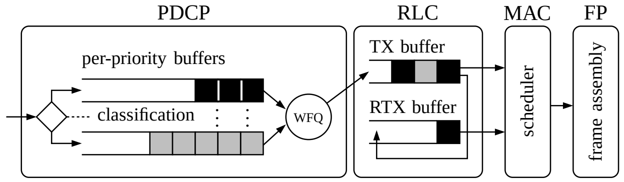

Figure 2. Implementation of in-bearer prioritization in the RNC radio protocol stack.

tion level, therefore UE side extra functionality would be required. Accordingly, network-side in-bearer prioritization transparent to the UE is feasible only in DL. The benefit of this solution is that it requires differentiation only at the radio network layer without involving any of the underlying transport network functionalities. This keeps the solution simple as the transport parameters and QoS mapping of the radio access and CN bearers are kept unchanged during their lifetime, thus no extra signaling is required.

III. SECONDARY PDP CONTEXT ACTIVATION

In the previous section, a single-bearer mechanism for application differentiation was discussed; in this section, a multi-bearer mechanism is presented that requires the use of multiple data bearers per UE. In the context of NRSPCA, a data bearer means both the radio access and the CN bearer, not only the RAB as with in-bearer prioritization. Multiple data bearers enable clean differentiation of applications with diverse QoS requirements without need to change the bearer-based QoS architecture. Additional bearer setup in 3GPP networks (such as HSPA or LTE) is a standardized procedure called secondary Packet Data Protocol (PDP) context activation, which can be requested either by the UE or by the network. Since application awareness requires that the network reacts to different applications of users, the secondary PDP context has to be requested by the network.

According to the 3GPP specifications, there is a PDP context for each data bearer that stores the service or Packet Data Network (PDN) the user connects to (e.g., the Internet); the IP address the UE uses in that PDN; and the QoS settings that apply to the PDP context. There are two types of PDP contexts: primary and secondary. Each different PDN to which a user is connected has an associated primary PDP context with default QoS profile attributes set according to operator policy. Users may have multiple active primary PDP contexts, one for each different PDN they connect to; however, the QoS profile of each PDP context applies to all traffic sent to or received from the corresponding PDN, i.e., although the access of different PDNs may be configured with different QoS settings, there is no means to further differentiate between traffic mapped to the same primary PDP context. The requirement of finer QoS configuration is the key motivation behind secondary PDP contexts, which allow QoS differentiation for applications and services (e.g., web browsing, FTP, P2P, streaming) over the same PDN, i.e., the Internet. Each secondary PDP context is associated with a primary PDP context, from which the PDN itself and the IP address of the UE are reused but the QoS profile can be different. A primary PDP context may have multiple secondary contexts assigned to it. Each PDP context, either primary or secondary, has a separate data bearer consisting of a RAB and a CN bearer for user plane data, i.e., the QoS configuration of the bearer is applied not only on the RAN (as with in-bearer prioritization) but consistently on the CN as well, both in UL and DL directions, which gives opportunity to prioritize applications end-to-end. Additionally, as the solution is compliant with the RAB-based QoS architecture, mapping of the bearers to the transport QoS is straightforward, which results in a compact harmonized end-to-end application aware QoS architecture.

The mapping of user-plane traffic to a certain PDP context is based on the Traffic Flow Templates (TFT). A TFT is created dynamically when a PDP context is activated and defines what kind of traffic belongs to the new context based on filters that can match, e.g., the IP address of the remote server or the source and destination TCP/UDP ports. DL TFTs are used in the GGSN for mapping DL user data to the correct GTP tunnel whereas UL TFTs are used in the UEs to implement the mapping in the opposite direction.

According to the standardized procedure of Network Requested Secondary PDP Context Activation [8], the GGSN triggers the UE to initiate the Secondary PDP Context Activation procedure with the QoS parameters and UL TFT specified by the GGSN in the first message. Thus, a functionality located in the CN is able to trigger the setup of secondary bearers with a predefined QoS configuration and, what is also important, the mapping of UL traffic can also be specified by the network. Using NRSPCA as an application aware feature is possible in a way that in case an application is detected in the CN (possibly via the same DPI mechanism also used for single-bearer mechanisms) that should be prioritized according to operator policy (e.g., HTTP traffic), the NRSPCA procedure is initiated to establish a secondary bearer with the desired QoS settings and the corresponding DL and UL TFTs are created in order to map the traffic belonging to the application into the new secondary bearer. Since the UL TFT is also created by the network and signaled to the UE, NRSPCA is suitable for treating both DL and UL traffic in a uniform way. In either case, the detection is done by the DPI located in the CN. After the application that triggered the NRSPCA finishes, which can be noticed, e.g., by activity detection, the secondary bearer should be terminated.

IV. SIMULATION MODELS FOR EVALUATION

In-bearer prioritization and NRSPCA were evaluated by examining web page downloads in a simulated multi-cell HSPA network. The radio network layout consisted of a central cell surrounded by six other cells placed at 250 m

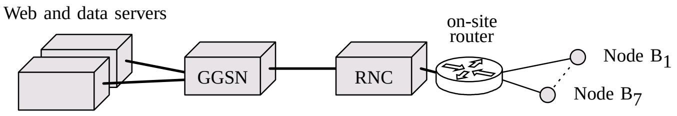

Figure 3. Simulation topology

inter-site distance. Users were distributed in the 7 HSPA cells, moving at an average of 3 kmph according to the random way-point mobility model. Wideband Code Division Multiple Access (WCDMA) air interface and handover procedures between cells were modeled in detail. The simulation topology is shown in Fig. 3; on the Iub interface, the Iub/IP/Ethernet protocol stack was used with different capacity configurations of 5, 10 and 100 Mbps, covering the range from heavy to no Iub congestion. Base stations were connected to the RNC in a star topology and implemented a Congestion Control (CC) algorithm [9].

Applications modeled in the simulation were TCP-based bulk data transfer and web browsing, the latter consisting of the complete HTTP/1.1 [10] protocol suite including Domain Name System (DNS) queries over UDP for name resolution. That is, users were executing file downloads and web page retrievals during the simulation.

The web browsing quality of experience was studied through the two most prominent quality measures: responsiveness, measured by web page download latency; and speed, measured by the page download rate. The download latency was the time between the user sending the request and receiving the first data byte of the web page. The download rate is the aggregated rate of TCP connections measured over the interval between receiving the first data byte of the main page and the download completion. Web surfing was implemented so that a random web site was selected from the list of top web sites [11] such as Google, Facebook, Wikipedia, etc., and the objects of a web page from that site were downloaded. For the simulation of web traffic, a profile was built for each modeled web page to record its main page size, the number and size of embedded objects and the server name for each object (in order to decide whether a DNS query was needed before establishing a TCP connection to the server). After a page had been downloaded, there was a random reading time in which no web traffic was generated. Then, the user visited another page from the same or another randomly chosen site.

For the sake of simplicity, two distinct user behaviors were simulated: background users having one bulk data transfer of continuous data download and multi-flow users with a similar bulk data transfer and additional web surfing. At the start of a simulation, there were 11 background users in each cell (total of 77 background users in the system) and there was one multi-flow user in the central cell. This setup was created in order to show the maximum achievable gain. With more multi-flow users, the gain would be smaller due to the increased amount of concurrent HTTP connections; however, relatively better service can be still provided to the HTTP connections since they are prioritized over the parallel applications.

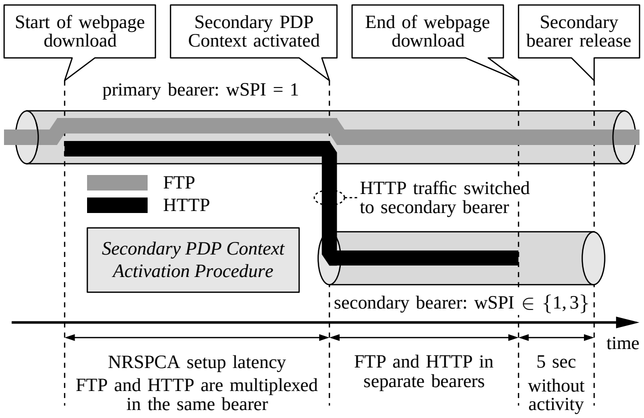

Figure 4. Modeling of Network Requested Secondary PDP Context Activation (NRSPCA) in the simulations.

For in-bearer prioritization, the DPI functionality was modeled in the GGSN so that DL packets were marked according to the application: HTTP packets were high and FTP packets were middle priority. DPI was assumed to be perfect so that all packets were marked correctly according to the corresponding application. This is possible as the DPI mechanisms available are efficiently detecting the applications with practically no latency, thus even the first HTTP packet can be treated according to the predefined QoS differentiation strategy. In-bearer prioritization was implemented in the RNC according to Section II with WFQ weights configured so that $w_{\text{high}} : w_{\text{middle}} = 9 : 1$. Due to the specified marking of HTTP and FTP, the weight of the applications was also $w_{\text{HTTP}} : w_{\text{FTP}} = 9 : 1$.

On the transport, all user-plane traffic was mapped to the same Per-Hop Behavior (PHB) group. On the radio, the default wSPI of all bearers was set to 1. There was no GBR configured to any of the bearers.

The simulation model of NRSPCA is illustrated in Fig. 4. In this case, the application detection was also done by the DPI in the GGSN but instead of marking the packets according to the application, the DPI triggered the activation of a secondary PDP context whenever a starting web page download was detected. Once the secondary bearer was established, the HTTP traffic sent in either DL or UL was mapped to that new bearer, whose wSPI was either set to 1 (i.e., the same QoS profile was used for primary and secondary bearers) or 3 (i.e., a new, better QoS profile was used for the secondary bearer in order to prioritize it over the other bearers). The signaling messages of the secondary PDP context activation were not simulated; instead, when a HTTP packet was detected for a user in the gateway, a timer was started that modeled the NRSPCA latency, i.e., the time required for completing the NRSPCA Procedure. When the timer expired, a secondary bearer was created for the HTTP

Table I
SHORT LABEL AND DESCRIPTION OF SIMULATION SCENARIOS

label	description
ref	reference case (no application aware feature)
wfq	in-bearer prioritization
nrspca-Z nrspca-L	NRSPCA with secondary bearer wSPI = 1; '-Z' denotes zero NRSPCA latency; '-L' denotes random NRSPCA latency between 0.8 – 1 seconds
nrspca-Z-pro nrspca-L-pro	NRSPCA with secondary bearer wSPI = 3; the meaning of '-Z' and '-L' are the same as with nrspca

packets. During the NRSPCA setup, HTTP packets were multiplexed with FTP in the primary bearer. After the web page download was complete, which was detected in the GGSN as 5 s without activity in the secondary bearer, the release of the secondary bearer was triggered.

Since the transmission of HTTP packets on the Iub interface is slower in the primary bearer during the NRSPCA setup than later in the dedicated secondary bearer (as the simultaneous applications of the users are still competing for the resources during this time), the first HTTP packets sent in the secondary bearer may arrive at the UE earlier than some of those sent through the primary bearer, potentially causing out-of-order delivery at the UE that would eventually trigger TCP Fast Recovery mechanism at the sender. In case the secondary bearer is prioritized as well, this effect is even more pronounced. In order to prevent this potential problem, we propose that after the secondary bearer setup is complete, the GGSN sends an end marker (GTP-U packet) [12] in the primary bearer and starts forwarding subsequent DL HTTP packets in the secondary bearer. Packets received in the secondary bearer are buffered at the RNC until the end marker arrives; on that occasion, the RNC transmits all packets it has buffered in the secondary bearer in the order of their arrival. After that, subsequent packets arriving in the secondary bearer are transmitted instantly. The same mechanism can be applied in UL as well, with the RNC sending the end marker and the GGSN buffering the packets in the secondary bearer. This mechanism is transparent to the UE and requires only network-side modification; it was implemented in all simulations presented in this paper.

For the NRSPCA setup latency, two configurations were used in the simulation: it was either set to zero, modeling an ideal case to assess the maximum achievable performance of this technique or it was chosen randomly between 0.8 – 1 seconds at each bearer setup to study the effects of a long and variable setup latency on web browsing performance.

V. SIMULATION RESULTS

The simulated in-bearer prioritization and NRSPCA scenarios are summarized in Table I. The web browsing experience measured by the download rate and latency is shown in Fig. 5, which displays the obtained results for the simulated scenarios at different Iub capacities. Each simulation was executed with five random seeds and the results were averaged

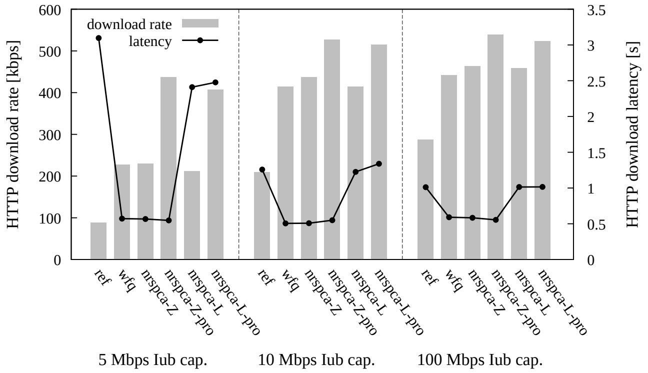

Figure 5. Simulation results showing the average HTTP download rate and download latency at different Iub capacities.

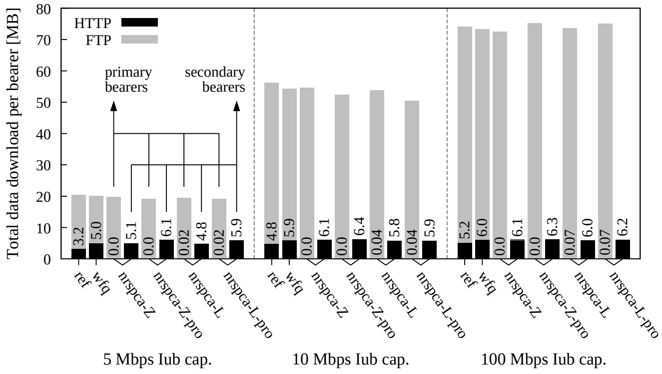

Figure 6. Total amount of data downloaded in the bearers, also visualizing the share between HTTP and FTP. For ref and wfq, there is only one bearer; for the NRSPCA cases, the primary and secondary bearers are both shown. Data labels represent the amount of HTTP traffic in MB.

to obtain the presented data. It is clear that, compared to the reference case, the mechanisms providing the highest HTTP download rate are those involving the setup of a prioritized secondary bearer for HTTP traffic, i.e., scenarios nrspca-Z-pro and nrspca-L-pro (regardless of the NRSPCA setup latency), which is due to the compact end-to-end QoS mechanism provided by the solution; in-bearer prioritization (scenario wfq) also results in considerably high download rate. The impact of the NRSPCA latency on download rate is not significant as despite the latency of setting up the bearer, the vast majority of HTTP traffic is still transmitted in the secondary bearer.

Regarding HTTP download latency, in-bearer HTTP prioritization and immediate secondary bearer setup for HTTP (nrspca-Z or nrspca-Z-pro) considerably reduce the download latency. This is due to that both types of mechanisms prioritize already even the first HTTP packets (either by allocating higher portion of the bandwidth to HTTP by PDCP WFQ or separating HTTP into a secondary bearer by NRSPCA), thereby reducing the queuing delay in the RNC and Node B radio buffers. By comparing nrspca-Z and nrspca-L, it is clear that higher NRSPCA setup latency results in increased HTTP latency that deteriorates user experience compared to the immediate bearer setup case; the reason is that it is the transmission of the first few HTTP packets that determine the HTTP latency and these packets are still transmitted in the primary bearer without any differentiation until the secondary bearer is established. It should be noted though that the HTTP latency is not worse than the one experienced in the reference case, thus in overall the user experience is improved. Prioritizing the secondary bearer (scenarios nrspca-Z-pro vs. nrspca-Z and nrspca-L-pro vs. nrspca-L) has no significant impact on the download latency as at the beginning of a HTTP session, the underlying TCP connection is still in slow start phase when the main page is requested and sent and there are not many packets on flight that would benefit from an increased wSPI in case of nrspca-Z-pro; also, in case the secondary bearer is set up with a latency, HTTP latency is determined by those packets still transmitted in the primary bearer, which has the same priority in the nrspca-L and nrspca-L-pro cases. Therefore, in case of NRSPCA, it is the separation of the HTTP packets into a secondary bearer and not the promotion of the secondary bearer that principally reduces the radio queuing delay and, consequently, the HTTP latency.

Besides HTTP download rate and latency, the total amount of data transmitted in each bearer was also measured; these results are shown in Fig. 6. In case of in-bearer prioritization, the total amount of data downloaded in the bearer is similar to that of the reference case, with the difference that HTTP represents a higher portion of the overall downloaded data due to the PDCP WFQ mechanism, and due to the fact that the web browsing session was not terminated during the simulation time; better circumstances resulted on increased amount of downloaded pages during the simulation time.

Among all scenarios, the total amount of downloaded HTTP data is the highest if NRSPCA is combined with secondary bearer prioritization (nrspca-Z-pro) since the prioritized secondary bearer does not have to share its bandwidth with other applications, i.e., it is fully allocated to HTTP traffic. With higher NRSPCA setup latency (nrspca-L and nrspca-L-pro), there is also some HTTP data in the primary bearer that is transmitted until the secondary bearer is established; however, the amount is not significant in comparison with the total HTTP data.

VI. CONCLUSION

In this paper, two alternative application aware mechanisms applicable in HSPA systems, namely the in-bearer prioritization and NRSPCA have been discussed and evaluated based on simulations. The evaluation was focusing on the use case of promoting web browsing traffic over bulk file transfer. Results indicate that NRSPCA is able to separate the applications efficiently. Together with its intrinsic, compact, bearer-based end-to-end QoS mechanisms it provides efficient differentiation and application specific services that outperform the in-bearer mechanisms. The advantage of the in-bearer prioritization is its transparency to the UE, making it a completely network-side solution only requiring support from the RNC; however, the fact that this solution is easily applicable only for DL traffic makes it less attractive. Nevertheless, the advantage of being transparent to the UE makes the in-bearer prioritization a competitive solution compared to NRSPCA-based solutions. Future work in the studied area can be devoted to the analysis of application aware methods in context of additional applications not considered in this paper.

REFERENCES

[1] H. Holma and A. Toskala, Eds., *HSDPA/HSUPA for UMTS*. John Wiley & Sons, 2006.

[2] M. Fiedler, T. Hossfeld, and P. Tran-Gia, “A generic quantitative relationship between quality of experience and quality of service,” *Network, IEEE*, vol. 24, no. 2, pp. 36–41, Mar. 2010.

[3] D. Soldani, H. X. Jun, and B. Luck, “Strategies for Mobile Broadband Growth: Traffic Segmentation for Better Customer Experience,” in *Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*, May 2011, pp. 1–5.

[4] J. Triay, D. Rousseau, C. Cervello-Pastor, and V. Vokkarane, “Dynamic Service-Aware Reservation Framework for Multi-Layer High-Speed Networks,” in *Computer Communications and Networks (ICCCN), 2011 Proceedings of 20th International Conference on*, Aug. 2011, pp. 1–7.

[5] H. Holma and A. Toskala, Eds., *WCDMA for UMTS*, 3rd ed. John Wiley & Sons, 2004.

[6] K. Pedersen, P. Mogensen, and T. Kolding, “Overview of QoS options for HSDPA,” *Communications Magazine, IEEE*, vol. 44, no. 7, pp. 100–105, Sep. 2006.

[7] D. Laselva, J. Steiner, F. Khokhar, T. Kolding, and J. Wigard, “Optimization of QoS-aware Packet Schedulers in Multi-Service Scenarios over HSDPA,” in *Wireless Communication Systems, 2007. ISWCS 2007. 4th International Symposium on*, Oct. 2007, pp. 123–127.

[8] 3GPP, “General Packet Radio Service (GPRS); Service description; Stage 2; R11,” 3rd Generation Partnership Project (3GPP), TS 23.060, Dec. 2011.

[9] L. Kőrössy and C. Vulkán, “QoS Aware HSDPA Congestion Control Algorithm,” in *Networking and Communications, 2008. WIMOB ’08. IEEE International Conference on Wireless and Mobile Computing*, Oct. 2008, pp. 404–409.

[10] R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, and T. Berners-Lee, “Hypertext Transfer Protocol – HTTP/1.1,” RFC 2616 (Draft Standard), Internet Engineering Task Force, Jun. 1999, updated by RFCs 2817, 5785, 6266.

[11] “Alexa Top 500 Global Sites,” Retrieved Sep. 2011. [Online]. Available: http://www.alexa.com/topsites

[12] 3GPP, “General Packet Radio System (GPRS) Tunnelling Protocol User Plane (GTPv1-U),” 3rd Generation Partnership Project (3GPP), TS 29.281, Dec. 2011.