The 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'07)

COMPARISON OF TRANSPORT CAPACITY REQUIREMENTS IN 3GPP R99 AND HSDPA IP-BASED RADIO ACCESS NETWORKS

Hiram Galeana, Ramon Ferrús, Juan J. Olmos
Universitat Politècnica de Catalunya
Barcelona, Spain

ABSTRACT

Growing number of subscribers and deployments of high speed air interfaces (e.g. HSDPA) are imposing stringent requirements for mobile backhaul. In this paper we perform a comparative study of the transport capacity requirements in an IP-based UTRAN when using either release'99 (i.e. DCH) or high speed channels (i.e. HSDPA) in the air interface. The analysis is focused in the interface between base stations and radio network controllers, referred to in the 3GPP as Iub interface. For both scenarios, the minimum bandwidth to be provisioned in the Iub interface is calculated for voice and data services, under different levels of traffic aggregation and delay requirements. Over such a basis, the influence of the protocol overhead within the bandwidth requirements is identified and a sensitivity analysis to factors such as the network size, traffic model parameters and channel rates is conducted.

I. INTRODUCTION

The transition of GSM/GPRS networks to third generation (3G) high speed wireless systems represents an evolution from second generation (2G) mobile networks in terms of capacity, data speeds and service capabilities. This migration, however, comes at a price of exponential growth of bandwidth in the Iub interface of the UMTS radio access network (UTRAN) in order to handle 3G cellular traffic. This interface lies between the Radio Network Controller (RNC) and the base station (Node B).

Backhaul is a growing concern among operators as traffic continues increasing with data applications like video streaming and high speed web surfing. Likewise, the launch of new technologies such as HSDPA (High Speed Downlink Packet Access) will force operators to add more capacity in the mobile backhaul network. HSDPA was defined and introduced in release 5 of 3GPP specifications, and has been designed to support a peak user data rate of over 10 Mb/s. In consequence, HSDPA will have a direct impact on resource utilization of backhaul transmission and thus will require large bandwidth in the radio access network [1].

Moreover, the advent of IP as a de facto networking technology and its presence on the transport network layer (TNL) of UTRAN facilitates the integration of different radio access technologies operating over a unique backbone and therefore enables the development of heterogeneous networks. However, it also represents a challenge due the stringent timing requirements that need to be satisfied by the TNL. In particular, advanced WCDMA radio control functions require that the transport of user traffic over the UTRAN must satisfy strict delay bounds, regardless if the traffic is real-time or non real-time. The IP-based radio access network (IP-RAN) should meet these requirements in a cost-effective way in terms of efficiency and utilization of the access network bandwidth, which consist in a vast infrastructure of point to point links (typically limited to an E1) between Node B and RNC.

In a previous paper [2], we presented an approach to estimate the amount of bandwidth required in the Iub so that traffic load conditions and stringent delay constraints of the TNL are fulfilled. Particularly, we restricted our attention on voice and web services to capture quite different dynamics in the resulting traffic pattern.

In this paper, we extend that study to compare capacity requirements of two possible scenarios that could be found in IP-based UTRAN deployments: dedicated channels and high speed channels. In the former, traffic is mainly supported by Dedicated Channels (DCH) according to UMTS release 99 (R99) specifications. On the other hand, in the second scenario traffic is offered through HSDPA channels in the air interface. Each scenario embraces particular conditions as well as delay requirements for the transport part of the network. Our target in this work is to quantify how an evolving technology such as HSDPA affects transport demands in an IP-based access network.

The paper is organized as follows. Section II, gives an overview of the evaluated scenarios, and give details of their corresponding delay requirements. In section III, we introduce the basis of our study, and briefly describe Iub interface modeling in addition to traffic models. Section IV presents some simulation results that will serve as a reference point. In section V, we perform a sensitivity analysis. Finally, some conclusions are presented in section VI.

II. EVALUATED SCENARIOS

This section gives an overview of R99 and HSDPA scenarios, including specific characteristics like protocol stacks and transport delay requirements.

A. *Release 99*

In this scenario is analyzed the impact on transport network resources when DCH channels are used in the air interface. DCH channels were introduced in R99 and currently they are commonly used in radio access bearers (RABs) for voice as well as for real time and non real-time data services. The user plane protocol stack for R99 scenario is depicted in Fig. 1a. The Radio Link Control (RLC) handles segmentation and retransmission of user data between the User Equipment (UE) and the RNC. The MAC layer handles the mapping between the logical channels and the transport channels as well as the selection of the data rates being used. At the output of the MAC layer bursts of Transport Blocks (TB) are generated every Transmission Time Interval (TTI) of the corresponding transport channel. For each DCH channel, the DCH Framing Protocol (DCH-FP) layer assembles the bursts transmitted in one TTI into one FP frame which is subsequently delivered to the IP TNL.

With respect to delay requirements, the tolerable delay bounds in the TNL for DCH channels are dependent on (1) delay requirements of the user traffic itself but also on (2) requirements derived from supporting radio control functions such as outer-loop power control and soft handover. These requirements result in particularly tight delay budgets to be satisfied. We assume that the acceptable delay value considered for voice services is around 5 ms and around 50 ms for data services. In the case of voice, the service itself is the limiting factor, while in the case of data services the radio functions are the limiting factor. Along with the previous values, and in order to assess the sensitivity of the obtained results with the delay constraints being considered, a softer delay restriction (20 ms) is also considered for voice as well as tighter delay restriction (5 ms) for data.

B. High Speed Channels

The launch of HSDPA radio channels leads to higher data rates in Iub interface, as well as different traffic patterns characteristics in the transport due to the fact that radio packet scheduling is moved to the Node B. HSDPA introduces new elements in the protocol architecture that have a direct impact on TNL requirements on the Iub interface. Fig. 1b depicts the user plane protocol stack for high speed channels. Unlike R99 where MAC layer was completely located at the RNC, a fast packet scheduling functionality is now introduced at the Node B (MAC-hs for HSDPA). The RNC retains only part of the MAC (MAC-d) mainly to handle logic channel multiplexing. It is worth noting that the RLC layer stays mainly unchanged except for some optimizations for real-time services such as VoIP. The use of buffering in the Node B permits a peak rate for the connection as high as the terminal and Node B capabilities allow, while keeping the maximum bit rate over Iub in line with the QoS parameters received from the packet core. In fact, having the transmission buffer at the Node B also requires flow control mechanisms to be applied, so that Node B buffer does not overload if radio conditions in the downlink make data to be retained at the Node B. Also in the downlink direction, the Node B buffer shouldn’t get empty as long as there are still user data pending for transmission at the RNC. The FP protocol specified to carry HSDPA data in the Iub interface is called the high-speed downlink shared channel FP (HS-DSCH FP).

Under this scenario, delay requirements for HS-DSCH FP frames are mainly due to the service itself since neither outer-loop power control nor soft handover are supported on these channels. According to this, softer delay restrictions can be considered for voice and data traffic (e.g. 50 ms and 150 ms, respectively). However, attending to potential delay values given in [3] UTRAN Long Term Evolution (LTE), values ranging between 1 ms and 15 ms are accounted for packet transmissions in the transport part of the RAN. Thus, in accordance with previous arguments, delay upper bounds of 5ms/50ms for voice and of 5ms/150ms for data have been considered in this scenario.

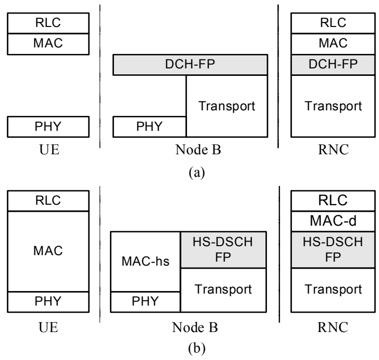

Figure 1: Protocol stacks: (a) R99 and (b) HSDPA

III. SIMULATION MODEL

In this section, we described the procedure used to estimate capacity requirements in the IP-RAN in order to meet the given delay requirements and mean traffic volumes. In addition, a brief review of the structure of the Iub interface modeling along with two service traffic models are described.

A. Capacity Estimation Approach

To determine the amount of bandwidth needed in a specific UTRAN deployment it is neccessary to know traffic demands, and the path each of these traffic demands follows through the network. With this information, the aggregated of traffic traversing each link in the transport network can be determined. In particular, we find out the mean traffic rate supported in a given link as the sum of mean values of the traffic generated by concurrent connections traversing that given link. Thus, for a particular link we compute the minimum link capacity required to meet a given delay constraint. If the network were to be designed without any delay guarantees, it would need enough bandwidth just to support the average data rate of the traffic. It is not possible to reduce the bandwidth beyond this point.

Once we have found the minimum link bandwidth, we evaluate the required capacity in the transport network by defining the excess of bandwidth with a parameter known as *over-provisioning factor (β)*, also refer to as beta factor. This parameter represents the percentage difference between the mean rate of the traffic and the minimum link capacity found to support a desired level of performance. This performance requirement is specified in terms of probabilistic delay requirement since average values are not enough for our purpose because variations of delay are non-negligible. We assume that a delay constraint is met if it is met for 99.9 % of the packets.

B. Iub Interface Modeling

The implementation structure of the IP-based transport is defined according to the guidelines provided in [4]. Such structure is separated into following modules that correspond to the functionalities of the Iub interface: link, IP transport, Radio Protocols/FP, and traffic sources. For more details about the Iub reference models used for the considered scenarios, please refer to [2].

On the other hand, the Iub interface model also involves input parameters like the amount of overheads of the modeled protocol stacks. Table I and II, respectively summaries the overheads considered for R99 and HSDPA scenarios.

Table 1: Protocol stack overheads in R99 scenario

Module	Component	Overheads	
		Voice	*Web*
Radio Protocols/ FP	PDCP/RLC /MAC	0 bytes	2 bytes
	DCH-FP	8 bytes	5 bytes
IP Transport	Stream / Overhead	3 bytes	3 bytes
	Container/ Overhead	8 bytes	8 bytes
	UDP/IP overhead	28 bytes	28 bytes

Table 2: Protocol stack overheads in HSDPA scenario

Module	Component	Overheads	
		Voice	*Web*
Traffic Source	IP/UDP/RTP [a]	4 bytes	N/A
Radio Protocols/ FP	PDCP/RLC /MAC	2 bytes	2 bytes
	HS-DSCH FP	10 bytes	10 bytes
IP Transport	Stream / Overhead	3 bytes	3 bytes
	Container/ Overhead	8 bytes	8 bytes
	UDP/IP overhead	28 bytes	28 bytes

a. We assume that after compression the header size is 4 bytes

C. *Traffic Models*

Two different traffic models showing quite different dynamics are analyzed: voice and data traffic. For each type of traffic, a detailed characterization of the complete Iub protocol stack is addressed so that the mechanisms used there are reflected into the traffic patterns observed at the transport network. The analysis of the two types of services is done separately, without mixing services.

The voice model consists of series of ON and OFF periods with a service rate of 12.2 kb/s, which corresponds to one of the bit rates achieved by the AMR codec specified by 3GPP. ON and OFF states are exponentially distributed with a mean duration of 3 sec. We assume that all users' sessions are kept active during the simulation elapsed time.

For data traffic we consider a web browsing model. A web session is modelled as a sequence of packets corresponding to the download of pages. The number of pages in a session is a geometrically distributed random variable with a mean of 5 pages. A truncated Pareto distribution is used to model the packet size of web traffic, resulting in a mean packet size of 366 bytes. The number of packets per downloaded page is modelled by a geometrically distributed random variable with a mean of 25 packets. Packet-calls are separated by a reading time which is geometrically distributed with a mean of 10 sec.

IV. Reference Results

With aforementioned traffic models and protocol stacks, an initial assessment of capacity requirement for both scenarios is performed using OPNET Modeler. For each scenario we assume two different maximum delay requirements in order to capture the sensitivity of the beta factor to the imposed QoS constraints.

In this sense, Fig. 2 plots the degree of over-provisioning required in R99 and HSDPA scenarios for different mean voice traffic load conditions. Part of the capacity corresponds to the amount of overhead introduced by each protocol stack. In the case of HSDPA scenario the overhead is more pronounced with a value around 60%, while for R99 scenario is approximately of 40%. In both cases, the resulting capacity requirements are mainly impacted by the overheads added at each layer of the Iub interface model. However, HSDPA incurs in higher overhead values due to the support of VoIP. On the other hand, Fig. 3 shows again both scenarios but now considering that web traffic is injected into the network. As can be seen, web service produces a significant increase in the over-provisioning factor with respect to voice service, as well as drastic changes between different mean traffic loads. The reason behind this is that web traffic requires large amount of bandwidth to cope with high traffic fluctuations. The impact of overheads is small in both cases with values around 10%, and consequently the most part of beta exclusively depend on the traffic pattern characteristics of web browsing traffic.

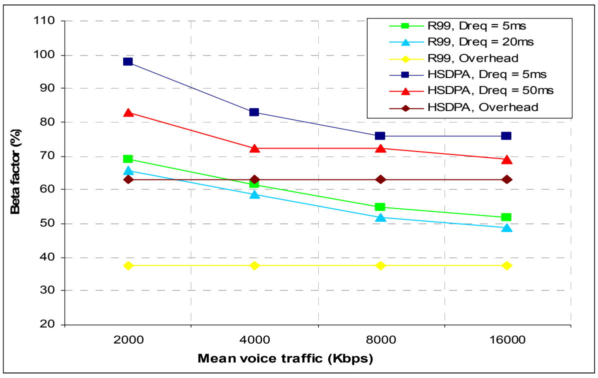

Figure 2: Beta factor in R99 and HSDPA for different mean voice traffic loads and delay requirements

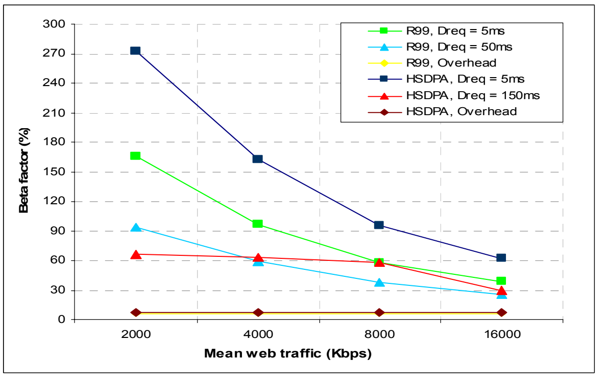

Figure 3: Beta factor in R99 and HSDPA for different mean web traffic loads and delay requirements

V. Sensitivity Analysis Results

In this section we explore the sensitivity of capacity requirements to a number of selected factors such as the network size, traffic model parameters and channel rates.

A. *Sensitivity of Capacity to the Network Diameter*

The traffic supported in the IP-RAN passes through a specific number of nodes (i.e. IP routers) in the path between RNC and Node B. This network path must accommodate enough capacity at each hop in order to fulfill end-to-end delay requirements. We refer to the number of hops a network path contains as *network diameter*. Thus, the total delay requirement has to be mainly distributed along the path components that incur in some type of delay. In general, delay components can be decomposed into per-hop delays and these in turn into per-link and per-node delay components. Per-link delay components imply the propagation delay, while per-node delay components involve three sub-components: serialization delay, processing delay, and queuing delay [5].

In this context, the amount of bandwidth that must be installed in each forwarding node to statistically guarantee a maximum tolerable end-to-end delay with a given probability can be approximated if the mean delay per-node is known. To calculate the mean value of the delay experienced by traffic in a node, it is necessary first to assemble the delay budget of an IP network path. Assuming that a path is composed by N identical nodes, the delay budget can be expressed in terms of the previously introduced delay components as follows:

$$D_{budget} = D_{prop} + \sum_{i=0}^{N} D_{n,i} \quad (1)$$

where the delay per-node i is given by

$$D_{n,i} = D_{ser,i} + D_{proc} + D_{q,i} \quad (2)$$

The first two components of (2), as well as propagation delay in (1), are likely to be the deterministic part of the delay budget and they are relatively easy to determine. On the other hand, queuing delay represents the stochastic part and it is more difficult to predict because its value depends on the congestion states of the network nodes.

A number of simplified conditions have been considered to analytically represent the delay experienced in a network path. We suppose that the total delay observed in the path exclusively depend on the waiting times in the queues, which are assumed to be exponentially distributed and independent. The queuing delay through N identical nodes is estimated using the closed-form formula presented in [6]. This formula expresses the (1-ε)-quantile of the total end-to-end delay requirement (D_{e2e}) for a Poisson traffic traversing N identical nodes as the sum of the average total queuing time μ_N and the number of times the standard deviation σ_N of the total queuing time. This formula can be written as:

$$D_{e2e} = \mu_N + \alpha_N(\varepsilon)\sigma_N \quad (3)$$

The values α_N solely depends on the diameter of the network and the value of ε, which is the portion of traffic that does not meet the delay requirements. With respect to the average waiting time, it can be expressed as:

$$\mu_N = N \times \overline{D}_{i,j} \quad (4)$$

where $\overline{D}_{i,j}$ is the mean delay in a network node. On the other hand, the standard deviation is determined by:

$$\sigma_N = \sqrt{N} \times \overline{D}_{i,j} \quad (5)$$

Using (4) and (5) in equation (3), and assuming a maximum tolerable value of D_{e2e}, the mean delay $\overline{D}_{i,j}$ in a hop can be obtained. From this mean delay and assuming that the delay is exponentially distributed, the (1-ε)-quantile of the delay incurred in a single hop can be calculated.

To study the impact on bandwidth requirements due to different network sizes, we consider the following procedure. We define different values of maximum tolerable end-to-end delay requirements and network sizes. We then calculate the probabilistic delay requirement of a single hop. Then, these values are used as input constraints to estimate the bandwidth required in one network node in order to satisfy the end-to-end requirement and support a mean traffic load of 2 Mb/s. Under this framework, we perform a set of simulations to evaluate bandwidth demands for R99 and HSDPA scenarios. Figs. 4 and 5 respectively illustrate simulation results for voice and web services. It can be seen that varying the network diameter from 2 to 16, capacity requirements of network nodes in the path increase in order to fulfill end-to-end delay requirements. As expected this effect is higher for tight values of end-to-end delay requirements (e.g. 5ms). Focusing on Fig. 4, and comparing R99 and HSDPA beta results it is observed a relatively small difference of about 10%. Conversely, considerable variations between scenarios is observed in web traffic case for stringent delay requirements, due that in both scenarios beta factor is more affected by the bursty nature of web service.

B. Sensitivity of Capacity to Channel Rates

In this part we investigate Iub capacity requirements for web browsing traffic taking into account the following DCH channel rates: 128 Kb/s, 256 Kb/s and 384 Kb/s. Simulation results for these cases are illustrated in Fig. 6. The worst combination we found is when 384 Kbps DCH channels are used in order to fulfill hard delay restrictions such as 5ms. Higher channel rates offer elevated levels of burstiness to the transport network, and thus require more amounts of bandwidth to satisfy delay requirements. On the other hand, small DCH rates serve as a kind of "traffic shaping" at the RLC/MAC queuing for the traffic which permits to have a more relaxed IP transport and consequently demanding less transport resources.

C. Sensitivity of Capacity to Traffic Model

It has been inferred that the inherent characteristics of traffic substantially determines transport capacities. The dependence of the results to the level of burstiness of the traffic model can be evaluated by means of implementing other types of services that exhibit different peculiarities in the offered traffic pattern.

The selected services for this study are streaming and background. To simulate the behavior of these services the following adjustments have been included in the Pareto traffic model [7]. For streaming data, a still image service is considered where data blocks of 60 Kbytes are send each 2 sec. The minimum data block size is 22 Kbytes, while the maximum is set to 147 Kbytes. The shape parameter alpha is equal to 1.1. For background data, a fax service is considered with the following parameters: inter-arrival time of 10 sec, the mean data block size is 200 Kbytes, minimum is 56 Kbytes, and maximum is 1.1 Mbytes, with alpha equal to 1.1.

Some simulation results for this case are provided in Fig. 7. It is observed that web traffic is still being the most bandwidth consuming. This is because a higher number of sources are multiplexed which may lead to situations with pronounced spikes in the aggregated offered traffic and thus requiring more bandwidth. On the other hand, the possibility to have softer delay requirements (i.e. 150 ms) in HSDPA, considerable reduces the amount of bandwidth needed when comparing with tight delay restrictions (i.e. 5 ms).

VI. CONCLUSIONS

In this paper we have compared transport capacity requirements of R99 and HSDPA IP-based UTRAN. The introduction of new technologies like high speed channels lead to higher data rates in the Iub interface of UTRAN and consequently increase backhaul bandwidth demands.

Although HSDPA protocol stack incurs in higher overhead values, less stringent delay requirements can be imposed. In this sense, an important reduction of transport capacity requirements can be achieved, leading to beta factors close to those obtained with R99, if same mean traffic load conditions are considered.

In has been evaluated the effect of different factors on transport capacity demands. In this context, small channel rates for web traffic in R99 permit to have a more relaxed IP transport with less transport resources, while high network size values imposed more stringent per-hop delay requirements resulting in increased capacity requirements.

ACKNOWLEDGMENT

This work is part of the EU IST-AROMA project [8], partially funded by the European Commission under the IST framework. This work is also partially supported by Spanish National Science Council under grants TEC2006-09109 and TEC2006-26873-E.

REFERENCES

[1] A. Bolle, H. Herbertsson, "Backhaul must make room for HSDPA", *Wireless Europe*, February/March 2006.

[2] H. Galeana, R. Ferrús, J. Olmos, "Transport Capacity Estimations for Over-provisioned UTRAN IP based Networks", in *Proceeding of the IEEE WCNC*, March 2007.

[3] 3GPP TR 25.912 v7.0.0, "Feasibility study for evolved UTRA and UTRAN (Release 7)".

[4] 3GPP, TR 25.933 v5.4.0, "IP Transport in the UTRAN (Release 5)".

[5] E. A. Smith, "Understanding Performance Issues in IP Networks", *BT Technology Journal*, vol. 24, no. 4, October 2006.

[6] D. De Vleeschauwer, et al, "Closed-Form Formula to Calculate the Dejittering Delay in Packetised Voice Transport", in *Proceedings of the IFIP-TC6 / European Commission International Conference Networking*, May 2000.

[7] C.Y. Jung, et al., "Performance Comparison of ATM and IP based Transmission Schemes in the UTRAN", in *Proceedings of the IEEE WCNC*, March 2004.

[8] IST-AROMA project web page http://www.aroma-ist.upc.edu/.

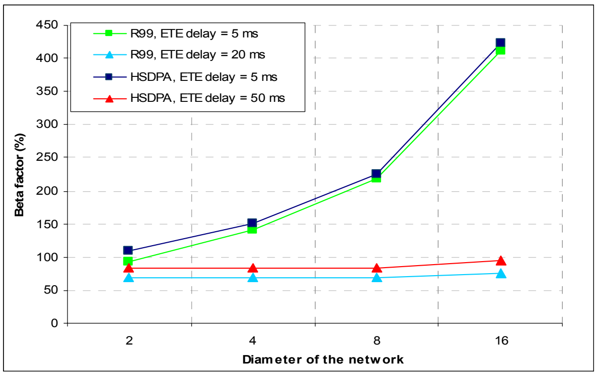

Figure 4: Beta factor for 2 Mb/s of mean voice traffic under different network diameters and end-to-end delays

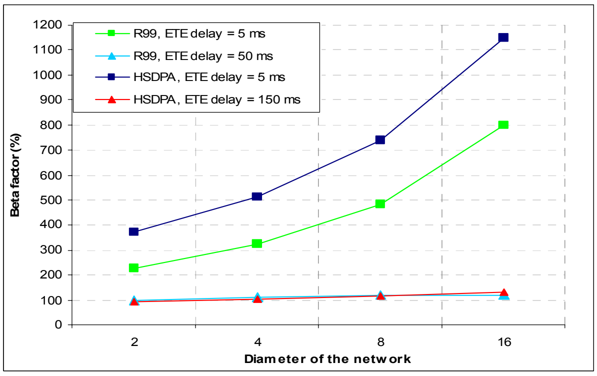

Figure 5: Beta factor for 2 Mb/s of mean web traffic under different network diameters and end-to-end delays

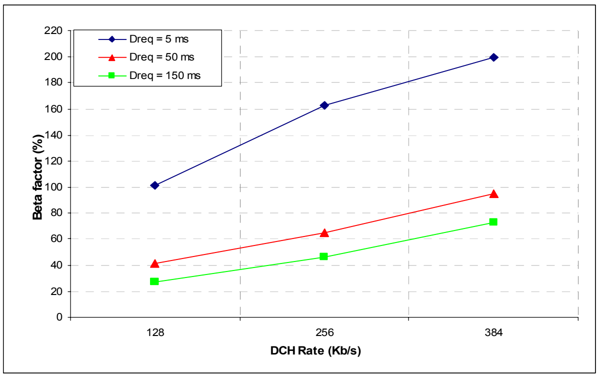

Figure 6: Beta factor for different channel rates and delays

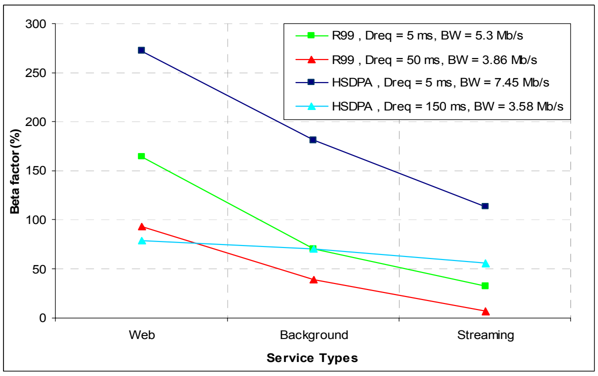

Figure 7: Beta factor for different services with link capacities that satisfy the 99.9% of delay requirements