The 17th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'06)

REALIZATION AND PERFORMANCE EVALUATION OF IMS MULTIMEDIA TELEPHONY FOR HSPA

Stefan Wänstedt, Mårten Ericson, Kristofer Sandlund, Mats Nordberg and Tomas Frankkila
Ericsson Research
PO Box 920, S-971 28
Luleå , Sweden
mats.nordberg@ericsson.com

ABSTRACT

The proposed Voice over IP (VoIP) over HSPA solution targets high capacity as well as coverage and speech quality comparable to circuit switched (CS) speech. The VoIP over HSPA radio bearer realization includes one radio link control (RLC) un-acknowledged mode for the speech media. Parallel signaling radio bearers are also transmitted using HSPA access. Due to the flexibility of IP and HSPA, both narrow-band and wide-band speech codecs work well. A delay sensitive scheduler is selected for the downlink transmission. For the uplink, the non-scheduled mode is selected.

Coverage and quality can, with simulations, be shown to be at least as good as CS speech according to 3GPP Rel. 99 specifications. Similarly, capacity evaluations show that VoIP over HSPA has the potential of matching or exceeding CS speech capacity, depending on scenario. Finally, end to end simulations show that the maximum allowed delay of single links seldom occurs for uplink and downlink simultaneously, resulting in a lower perceived end-to-end delay than expected most of the time.

I. INTRODUCTION

The introduction of an all-IP cellular network [1] enables a seamless cellular user experience. In the all-IP cellular network users will be provided with an application flexibility superior to that of the circuit switched infrastructure offered in the first 3G networks [2]. These applications are often referred to as MultiMedia Telephony (MMTel) and belong to a set of services, including speech, real-time video, and image sharing. To facilitate the introduction of new services the mobile industry introduced the IP Multimedia Subsystem (IMS), a service layer that allows operators to efficiently deploy and manage the above-mentioned IMS services

The advent of a reliable, efficient and flexible VoIP service is one important prerequisite for the transition from today's networks into a single all-IP network. The convergence into one all-IP network can lead to lower overall operational cost for the operator.

One price to pay for the greatly enhanced service flexibility is the need to transfer the IP headers all the way to the receiving client. This results in higher total bit rate requirements for VoIP than for CS speech. Consequently, it is a challenging task to make VoIP over cellular into an efficient service, considering the CS speech capacity, quality and coverage that users have become accustomed to. With the introduction of HSPA, i.e. downlink High Speed Data Packet Access (HSDPA) [3] and Enhanced Up-Link (EUL) [4], the capability of the transport is vastly increased compared to 3GPP Rel. 99. Actually, these improvements of the radio interface are prerequisites of a resource efficient deployment of flexible real-time IP services.

In this paper we demonstrate that it is possible to deploy efficient VoIP over HSPA with capacity, coverage and quality that equals or exceeds that of CS speech.

II. SERVICE REQUIREMENTS FOR VOIP

Service requirements for VoIP over cellular are set by the mobile CS services available today. The most important service requirements are quality, coverage and capacity. A high quality CS speech service usually has an end-to-end delay of no more than 220 ms, speech frame loss of less than 2% (the sum of UL and DL) and almost no delay jitter at all. Reaching competitive efficiency for VoIP is not just a matter of optimizing the radio network performance; all nodes in the chain must be involved.

Delay and its consequences are some of the most important aspects affecting VoIP speech quality. The most important nodes in this context (that is, those that contribute the most to delay) manage speech encoding and packetization, the radio network, and possibly IP transport if the geographical distance between users is large, see Figure 1 and Table 1.

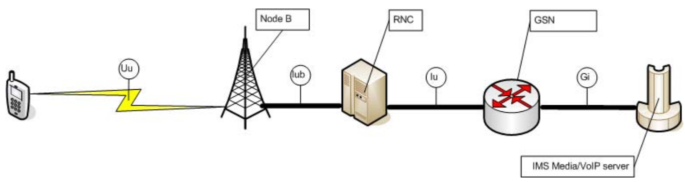

Figure 1: Network nodes.

III. HSPA REALIZATION FOR VOIP

The objective with our optimized VoIP radio bearer realization is to achieve capacity matching or exceeding that of CS speech in an interference limited scenario, coverage comparable to that of CS speech and speech quality at least comparable to CS speech (AMR 12.2 kbps).

Codec types and PDU sizes

The radio bearer realization elaborated in [9] investigates possible optimizations for at least the following codec modes (besides the SID frames): AMR: 12.2, 7.95, 7.40, 5.90 and 4.75 kbps. AMR-WB: 12.65, 8.85 and 6.60 kbps.

Table 1: Delay budget for system with EUL TTI of 10 ms without delay for jitter buffer. The maximum delay budget is 280 ms and the minimum is 100 ms.

UL		**DL**	
AMR	35 ms	AMR	5 ms
UE L1/L2	4 ms	UE L1/L2	10 ms
TTI Alignment	0-10 ms	-	-
Uu Interleaving	10 ms	Uu Interleaving	2 ms
UL Tx + re-Tx	2-75 ms	DL Scheduling	5-100 ms
RNC/Iub	9 ms	RNC/Iub	9ms
Iu + Gi	5ms	Gi + Iu	5ms
Sum max UL	148 ms	Sum max DL	131 ms
Sum min UL	65 ms	Sum min DL	36 ms

Media transport

The AMR payload format encapsulating the AMR frames should use the bandwidth efficient mode to maximize efficiency. One AMR frame will be transmitted per RTP/UDP/IP packet.

Head compression - ROHC

The IP/UDP/RTP headers that need to be added to speech frames result in an overhead ratio of about 60-70% (header: 40 to 60 bytes; payload: 32 bytes for AMR12.2). Using header compression (ROHC) this ratio can be reduced to less than 10% (header: typically 3 bytes) [10].

Optimized PDU sizes

To gain efficiency by avoiding padding, the RLC PDUs are optimized for the VoIP frames, as far as possible. As a comparison, CS speech frames match RLC PDUs perfectly. The benefit of using RLC PDU sizes that match the payload from the speech codec is large. The capacity gain is about 15% for AMR 12.2 kbps (RLC PDU of 288 bits), compared to when using a RLC PDU of the best effort RAB (336 bits). An alternative to optimized PDU sizes is to use RLC concatenation. A drawback of concatenation is that it adds delay and that the coverage may decrease compared with optimized PDU sizes.

Application and media control signaling

The radio bearer realization assumes SIP for the session control, but also for other services such as presence. The main part of the SIP signaling is sent during the set-up phase, while the SIP activity during the conversation is expected to be low and therefore having minor impact on the VoIP capacity. Compression of SIP using SigComp [11] is assumed. In this study SIP is not included at all.

Scheduler

Scheduling, in general, and the maximum delay that is allowed for scheduling, in particular, are very important factors for VoIP over HSDPA. A delay scheduler is assumed for the downlink transmission, since evaluations show that it is beneficial for fairness among users as well as capacity to consider delay when scheduling VoIP traffic.

The delay scheduler is based on a Max CQI scheduler normalized with the average CQI of the user, with a delay barrier function. With the barrier function, the priority is increased as a function of the total time in queue, automatically giving higher priority to retransmissions and maintaining real-time characteristics in a resource efficient manner.

Tuning parameters of the scheduler correctly is very important for efficient performance. It is possible to design a delay scheduler that works well for a short maximum scheduling delay without compromising the performance at higher maximum delays, not meaning that all delay schedulers performing well at higher maximum delays are suitable for shorter delays.

Radio bearers

The actual VoIP solution includes two radio bearers, one over RLC UM for the speech media and one over RLC AM for the SIP application signaling. These two bearers are transmitted over HSPA.

The radio resource signaling (RRC) and non-access stratum (NAS) signaling are also transmitted over HSPA. Mapping the signaling over HSPA has the advantage that is reduces the setup and handover procedure time.

IV. THEORY

A. *UL Theory*

The maximum number of users, Mp_{ul}, in the UL can be written as

$$M_{ul} = \frac{1}{1+F}\frac{1+\gamma_{ul}L_{demod}(1-\alpha)}{\gamma_{ul}L_{demod}}, \tag{1}$$

where F is the interference factor from other cells and L_{demod} is the demodulator loss, i.e. the loss due to poor channel estimation, as a function of DPCCH carrier to interference ratio (CIR) target. Note that the received CIR at Node B is the combined CIR assuming 2 receiver antennas, i.e. $CIR = CIR_1 + CIR_2$. This gives roughly a gain of 3dB. This also means that the actual DPCCH CIR target used by the UE, γ_c, can be 3 dB lower. The total CIR, γ_{ul}, is calculated as

$$\gamma_{ul} = a_c\gamma_c + a_{ec}\beta_{ec}\gamma_c + a_{ed}\beta_{ed}\gamma_c, \tag{2}$$

β_{ec} and β_{ed} are the (linear) power offsets for the E-DPCCH and the E-DPDCH channels, respectively, with reference to the DPCCH CIR. The activity for the E-DPCCH and E-DPDCH channels are calculated as

$$a_{ec} = a_{ed} = a_{voice}\frac{tti}{amr_{frame}}R_{mean}. \tag{3}$$

a_{voice} is the voice activity factor, tti is the TTI length (2 or 10ms), amr_{frame} is the AMR frame length of 20ms and , finally, R_{mean} is the average number of transmissions per MAC-hs PDU. Figure 2 shows EUL VoIP capacity estimates as a function of the combined CIR (assuming 2 RX antennas). The figure shows the results for 2 and 10 ms TTI with and without demodulator loss. The resulting E-DPDCH CIR at the base station is constant for all DPCCH CIR targets, i.e. the power offset of E-DPDCH is adjusted to match a constant E-DPDCH CIR regardless of DPCCH CIR target.

The voice activity is 50% and 10% of the packets are assumed to be retransmitted. These retransmissions are only modeled as extra interference; the packet delay is not modeled in the equations.

Figure 2 shows that the demodulator loss lowers the capacity compared to the ideal case. When the DPCCH CIR target becomes lower than -24dB the demodulator loss is so high that the capacity cannot increase more.

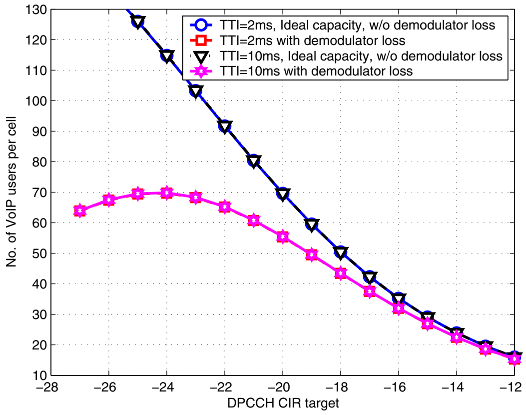

Figure 2: EUL VoIP capacity estimations example by simple equations.

Note that other issues that may occur due to low DPCCH CIR target, like TPC errors and UL synchronization problems, are not considered here, nor are they considered in the simulations.

B. *DL Theory*

The theoretical VoIP over HSDPA capacity is derived in [5] and [6], and is not repeated here due to space constraints.

It is found that the capacity function is rather complex, and depends on many parameters. The overhead plays an important role. The common channel (CCH) overhead transmit power and the power of the HS-SCCH result in a reduction of capacity proportional with the square root of their transmit power. At the same time, the capacity increases proportionally with the square root of the number of code multiplexed users. The optimal number of simultaneous code multiplexed users is around 4-6, but obviously depends on the required transmit power of the HS-SCCH.

The A-DPCH, in contrast to CCH and HS-SCCH transmit power overhead, decreases the capacity almost *linearly* with the average transmit power per user, and therefore affects the capacity to quite a large extent compared to the common channels. Exchanging A-DPCH with F-DPCH results in a capacity gain of about 15%, [5].

The delay threshold roughly increases the capacity with the square root of the allowed time for DL scheduling (using round robin scheduling). Furthermore, the arrival rate and the retransmission factor (due to e.g. bad CQI measurements) have a linear impact on the capacity. Another important observation concerns the required CIR for transmitting a speech packet successfully, which impacts the capacity inversely proportionally to the square root of the CIR.

V. Capacity Simulations

The capacity of VoIP over HSPA, based on the considerations described above, has been evaluated both for uplink (UL) and downlink (DL) separately, but also with end-to-end simulations. All the simulations are performed with a dynamic network simulator. As references to the above simulations, capacity evaluations are performed of CS speech on the same simulator and the capacity numbers presented in the figures are normalized with CS DL capacity. Throughout the evaluation, AMR 12.2 is assumed, since this is the most common CS speech codec rate. The results consider a normal point-to-point speech only conversation, in which case no synchronization with other RTP media streams or remote endpoint aliveness is needed. RTCP is therefore excluded.

A. *UL Simulations*

Figure 3 shows the VoIP over EUL capacity for a CIR target of -19dB and -22dB, respectively, for the control channel carrying the physical layer control information. A CIR target of -22dB is used since it gives a TPC error rate of approximately 4%, see e.g. [8].

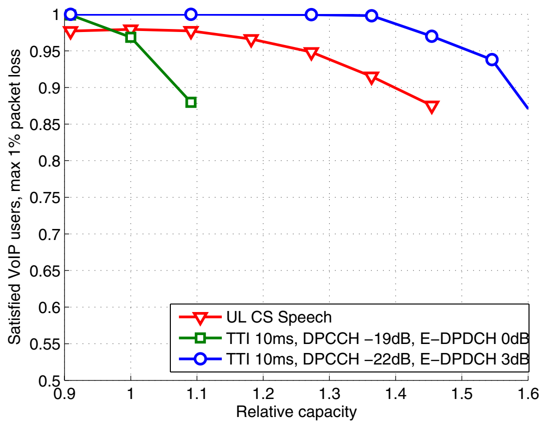

Figure 3: VoIP over HSPA UL relative capacity. Maximum delay in the UL is 75ms.

The higher CIR target is included to demonstrate the importance of the DPCCH CIR target for the VoIP capacity. The VoIP capacity (with CIR -22dB) in the figure is more than 40% higher than CS DL speech.

Note that results in Figure 3 are a bit optimistic since the effect of HS-DPCCH is not included . In short, the impact of the HS-DPCCH depends on the CQI update rate. An update every 4:th or 8:th TTI seems reasonable. The capacity loss is then 19% or 11%, respectively. Taking the HS-DPCCH into account gives a VoIP uplink capacity that equals that of UL CS if a DPCCH CIR target of -22dB is achievable.

B. DL Simulations

The simulations show that the capacity of VoIP over HSPA is at least as high as that of CS speech, given the assumptions that these simulations are based on, see Figure 4 where VoIP capacity is normalized with the DL CS, which limits CS capacity. The downlink VoIP capacity is almost 20% better than CS downlink capacity. The capacity results assume a fixed cell

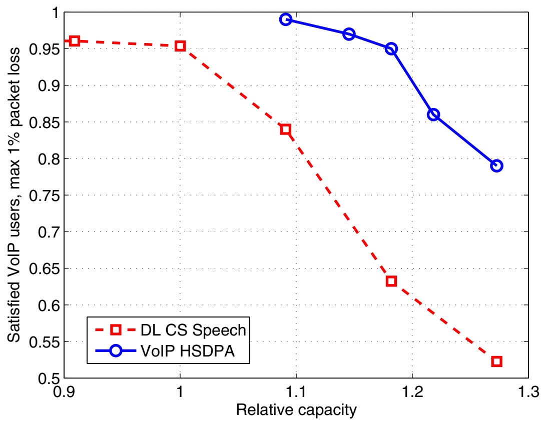

Figure 4: VoIP over HSPA DL capacity. Maximum delay in the DL is 100ms.

size and quality comparable to CS speech. Furthermore, the system is optimized for VoIP. Naturally, given a different scenario the results may change. Since the system is interference limited, increasing the cell size will decrease the capacity. This can to some extent be mitigated by allowing longer time for scheduling, since retransmissions partly make up for not having enough power. Increased scheduling delay may lead to decreased quality, since the end-to-end delay will become much longer than that of CS. However, since the delay on the UL usually is relatively low up to the capacity limit, the DL can potentially use this time to do at least one more retransmission. This would enable maintaining the maximum capacity even though the cell size increases, without exceeding the delay budget

C. End-to-end Simulations

Probably the most interesting case is to simulate VoIP sessions where two different UEs located in different cells communicate with each other, thereby simulating a VoIP call between two different mobile users. The results were obtained using both a 2 ms TTI and a 10 ms TTI for EUL. In this setup, the simulator is able to combine the delay of both links traversed and in this way, the total end-to-end delay can be measured. Also, the overhead of downlink signaling on the uplink and uplink signaling on the downlink can easily be applied to the results, since both links are simultaneously active in the simulation. It should be noted that the users' talk bursts are not correlated, and therefore it is as likely that both users are speaking at the same time as it would be that they are silent simultaneously. Figure 5 shows the resulting delay at high load. It is clear that a user in the system seldom requires the maximum delay in UL and DL, which means that the resulting delay most of the time is much lower than the sum of the maximum allowed delay for the links.

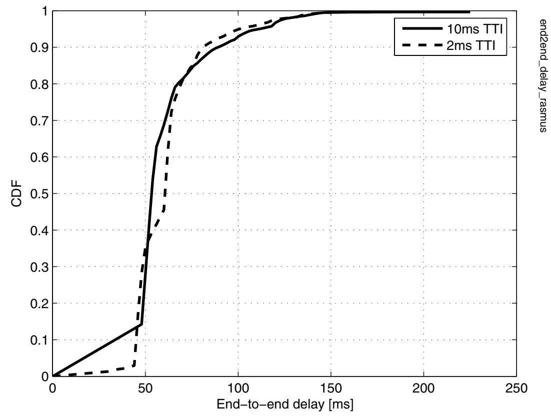

Figure 5: Resulting delay from e2e simulation at high load

VI. QUALITY EVALUATION

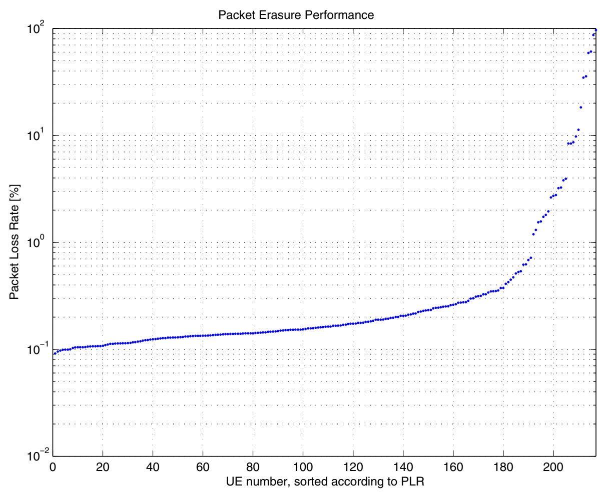

Figure 6: Packet losses in system with high load.

When evaluating the capacity in cellular systems, and when comparing the capacity with circuit switched systems, it is important to verify that the speech will have comparable quality. This is needed to ensure that the evaluation is made on equal terms.

A quantitative objective speech quality evaluation was carried out by compiling delay and packet statistics per call, imposing the quality constraints described in section II., for a system with high load. The analyses show that 85% of the users experience no packets losses at all, and that almost all users (99%) experience good quality (less than 2% losses), see Figure 6. But when the packet loss rate begins to increase, it increases rapidly to very high (unacceptable) levels. Analyzing the packet loss distribution for the satisfied users shows that most losses are single or double losses. There are occurrences with 3-5 packet losses in-a-row, or longer, but these loss bursts are few and should have no major impact on the speech quality. A similar analysis of delay shows that 98% of the users have a maximum packet delay that is less than 220 ms.

Combining the losses from the transport with the late losses due to delay shows that about 98% of the users would be satisfied. This quantitative speech quality analysis shows that the capacity simulations for VoIP over HSPA gives at least comparable speech quality to CS speech.

VII. Coverage

Similar to capacity, there is flexibility to trade delay for coverage both for UL and DL, but other settings also affect coverage. The UL coverage depends on how much power is allocated for the data channel (beta or power offset value), TTI length and number of retransmissions. The study [7] shows improved coverage for EUL using a long TTI (10 ms) compared to using a short TTI (2 ms), under comparable conditions, considering the same delay budget. By allocating more power to the EUL data channel and allowing more transmission attempts, VoIP over EUL has the potential to have a comparable or even better coverage than CS speech.

VIII. Evolution

The current improvements (e.g. F-DPCH, optimized RLC PDU size to AMR codec, optimized delay scheduler, optimized power offset in UL) have helped boost the VoIP capacity. It is of course possible to increase the VoIP over HSPA capacity even further.

One issue that can be improved is the overhead of the UL control channels (primarily the DPCCH and HS-DPCCH). In 3GPP [8] there is work ongoing on ways to reduce the overhead of these channels, which can substantially increase the UL capacity. It may also be possible to reduce overhead in the DL even further, especially for cases with many low bit-rate services in a cell. Other improvements include advanced receivers and new equalization methods, both for the UE and the base station. Some of these improvements can also be applied to CS speech, but not all.

IX. Conclusions

A prerequisite for MMTel over HSPA in an all-IP network is an efficient, high quality speech service. Evaluations of the proposed realization show that VoIP over HSPA has the potential of matching or exceeding CS speech capacity and coverage. For example, the downlink VoIP capacity is more than 10% better than CS DL capacity and the VoIP UL capacity equals the CS UL capacity with an EUL control channel CIR target of -22dB. The combined delay of UL and DL seldom reaches the maximum delay budget at the same time. The capacity results assume a fixed cell size, quality comparable to CS speech, and a system optimized for VoIP. Naturally, given a different scenario the results may change. Examples of enhancements that will further boost capacity comprise reduction of power on control channels, such as HS-SCCH, or even switching off control channels when no data are transmitted.

References

[1] 3GPP, *TS 22.258, Service Requirements for the All-IP Network (AIPN)* , 2005, Rel. 7

[2] Ulf Olsson, *Towards the All-IP Vision*, Ericsson Review, no 1,2005

[3] Stefan Parkvall, Eva Englund, Peter Malm, Tomas Hedberg, Magnus Persson and Janne Peisa, *WCDMA Evolved - High-Speed Packet-Data Services*, Ericsson Review, no 2, 2003

[4] Stefan Parkvall, Janne Peisa, Johan Torsner and Mats Sågfors, *WCDMA Enhanced Uplink - Principles and Basic Operation*, VTC 2005, Spring, May , Stockholm, Sweden

[5] Stefan Wänstedt, Mårten Ericson, László Hévizi, Jonas Pettersson and József Barta, *The effect of F-DPCH on VoIP over HSDPA Capacity*, VTC2006, Spring, Melbourne

[6] Mårten Ericson, Stefan Wänstedt and Jonas Pettersson, *Effects of simultaneous circuit and packet switched voice traffic on total capacity*, VTC2006, Spring, Melbourne

[7] Fan Rui, Fredrik Persson, Mats Nordberg, Stefan Wänstedt, Miao Qingyu and Gu Xingyu *Coverage study for VoIP over Enhanced Uplink*, VTC2006, Spring, Melbourne

[8] 3GPP, *TR 25.903 Continuous Connectivity for Packet Data Users*, 2005, Rel. 7

[9] *TR 25.993, Typical Examples of Radio Access Bearers (RABs) and Radio Bearers (RB) supported by the Universal Terrestial Radio Access (UTRA)*, 2006, Rel. 6

[10] Bormann et al. *RFC3095 RObust Header Compression (ROHC)*, IETF, 2001

[11] Price et al. *RFC3320 Signaling Compression (SigComp)*, IETF, 2003