# Circuit-Switched Voice Services over HSPA

Qualcomm Incorporated, Corporate R&D
San Diego, USA

***Abstract*—Circuit-Switched (CS) Voice Services over HSPA (CSoHS) was recently introduced for 3GPP WCDMA Release 7/8 systems. The goals of this feature include improving voice and data system capacity by utilizing the improvements offered by the shared packet transport of HSPA air interface, while retaining the already widely deployed core networks. In this paper, we discuss the implementation of CSoHS and analyze its performance via simulations. We show that the system capacity of CSoHS is significantly higher than the Release 99 CS voice under similar system conditions and voice quality. We also discuss and analyze the performance of best effort traffic in the presence of voice users.**

***Keywords-CSoHS; 3G; AMR; HSPA;VoIP***

## I. INTRODUCTION

Voice services over wireless networks have traditionally been provided by circuit-switched (CS) systems where a dedicated channel is used for each voice call. This provides guaranteed Quality of Service (QoS) in terms of end-to-end delay for the voice traffic; however the capacity is limited since the resources for a dedicated channel are always occupied even though they are only used when the voice traffic is being carried out. As more advanced 3G systems are being widely deployed, using VoIP instead becomes an option for operators. These 3G systems allow integration of data services with voice services and thus provide higher network bandwidth efficiency, better manageability and cost savings as well as richer services. Carrying voice over a shared packet transport provides better utilization since the shared resources are used by a voice user only when the user is active.

HSDPA and HSUPA (called HSPA together) have been standardized in 3GPP (Third Generation Partnership Project) Release 5 and 6 respectively. They provided significant improvement over Release 99, including data rate increase and packet delay reduction by using features such as Hybrid ARQ (HARQ), Adaptive Modulation and Coding, shorter frame sizes, opportunistic scheduling and other QoS enablers. Release 7 (HSPA+) introduced further improvements including overhead channel reduction and battery savings features [1]. All of these advancements made these systems well suited not only for data applications but also for delay sensitive applications such as Voice over IP (VoIP).

It has been shown that VoIP provides higher user capacity as well as higher spare data throughput for non-voice users [2]. However, although the HSPA radio-access networks can accommodate VoIP, an end-to-end all-IP VoIP implementation requires changes at the core network. These include call session setup, control and signaling and handover mechanisms from VoIP to CS voice under mobility conditions. In addition, support of features such as header compression at the mobile and access network is needed.

Recently, in order to realize the capacity benefits of HSPA+ air interface while still employing the current voice core networks, CSoHS was introduced in 3GPP Release 7/8 [3]. The main proposal of this feature is to carry the voice traffic over the new HSPA radio channels which have better efficiency. This provides the higher capacity of HSPA+ systems as in a VoIP deployment but, unlike VoIP, the voice traffic is not carried over an IP backbone. The implementation of CSoHS requires relatively minor changes to the Radio Access Network (RAN) and the User Equipment (UE), which are achievable through software upgrades.

In this paper, we discuss the feasibility and performance of CSoHS in current 3GPP networks. We analyze the system performance and the effect of system parameters and features, including interference cancellation and mobile receiver types. We show that the CSoHS capacity of a Release 7/8 network is much higher than that of Release 99 circuit switched network given similar quality requirements. The remaining data capacity of Release 7/8 networks is shown to be considerably higher when the same number of voice users is served by CSoHS instead of circuit-switched voice over dedicated channels, thus providing better resource utilization in a typical mixed traffic environment.

## II. 3GPP Voice Evolution

The voice services require a guaranteed QoS in the sense that a voice frame should be delivered to the receiver with a very small probability of error and within a certain delay bound. In 3GPP Release 99, this is achieved by a CS system where each voice call has a dedicated channel and thus can receive the required QoS all the time. In contrast, in a Release 7/8 VoIP type implementation, the voice traffic has to share the radio channel with other types of traffic. The required QoS can be provided by using intelligent scheduling which gives higher priority to voice traffic over other data traffic when necessary.

The wireless environment and the mobility of users present additional challenges to VoIP. The radio channel conditions between the UE and RAN are usually time varying so that the channel can become better or worse over time. In Release 99, the solution to this problem is to adjust the transmission power according to the channel fluctuations in order to keep a constant level of quality as well as to take advantage of robustness provided by macro diversity. However, for VoIP other types of traffic which have varying degrees of tolerance to delay, the network can allocate power and other resources when the channel conditions are good. It has been shown that this is a more bandwidth efficient mechanism and has been adapted in HSPA (as well as other 3G systems). Furthermore, it is possible to balance this allocation based on a traffic flow's required QoS. Moreover VoIP allows tradeoff between delay (quality) and capacity, which can be exploited to some degree. HSPA systems considered here provide this flexibility and make it possible for the RAN to support high quality voice services over shared packet channels.

The VoIP evolution for wireless networks requires other changes at both the RAN and core network. The RTP/UDP/IP and other headers contribute a significant overhead for VoIP packets. Therefore it is necessary to support header compression at both the UE and the network. The core network also has to support QoS, voice call control and switching. These features are well established for CS system elements and an equivalent structure is necessary for efficient support of an IP-based voice service. To this end, 3GPP has introduced IP Multimedia Subsystem (IMS) which provides an architecture to support combined multimedia and voice services as well as to facilitate wireline and wireless convergence for these services. IMS relies on Session Initiation Protocol (SIP) for VoIP call setup and control. It also includes Voice Call Continuity (VCC) for seamless handovers between IMS and CS systems which is necessary for gradual deployment of VoIP. Even though IMS/SIP provides a unified framework to support VoIP and other richer services and QoS, its adoption has been slow in wireless networks.

Since a full deployment of VoIP is dependent on the other enablers mentioned above, 3GPP has recently introduced CSoHS where voice traffic is carried over the HSPA data channels at the RAN but the CS core network remains intact. The implementation of CSoHS in current 3G HSPA networks requires relatively minor changes at the radio access as shown in Figure 1. These modifications can be introduced by software upgrades. The main new functional elements are de-jitter buffers at both the RNC and UE as well as a delay sensitive scheduler at the NodeB. In addition, both the RNC and UE now map the circuit-switched call to the packet data radio bearers with appropriate QoS.

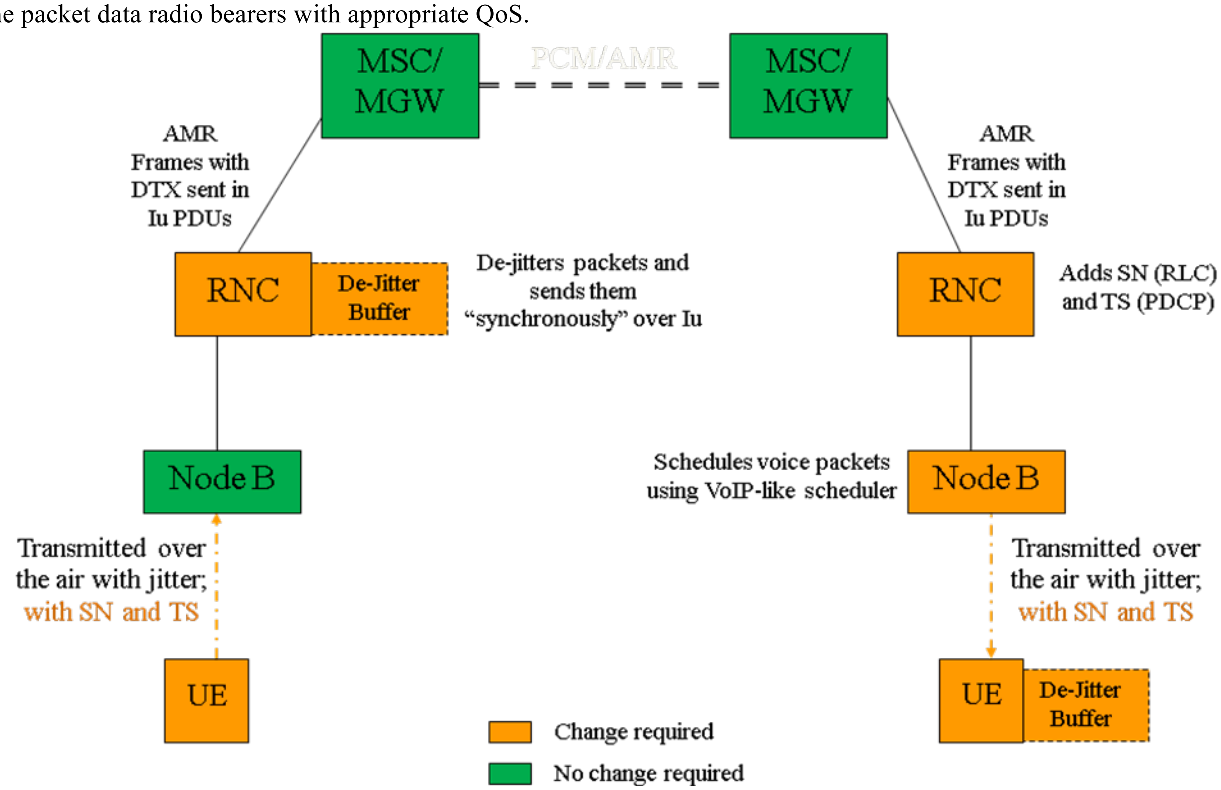

**Figure 1 Mobile to Mobile CSoHS Data Flow**

A de-jitter buffer at the RNC is necessary because the voice packets may arrive at the RNC from a UE with jitter (on the uplink) which means that the inter-arrival times of packets is not constant.. This is mostly due to the HARQ transmission on the uplink (UL) where the number of HARQ attempts for a successful transmission can vary for each packet. Another reason for jitter can occur in soft-handover situations where the transmission delay from each NodeB to RNC could be different. The RNC uses information elements in the headers to identify the correct order and timing of the voice frames. These are: Transmission Number (TN), Sequence Number (SN), and CS Counter. TN is a 7 bit number which is added to the MAC packet header to help with the re-ordering. The SN is carried in the RLC header and makes it possible, along with TN, to determine the correct order of received packets as well as detect missing packets. A new CS Counter was added to the PDCP header to indicate the timestamp of the voice frame. The RNC sends the output of the de-jitter buffer to the MSC synchronously over the IuCS interface as it is done for a CS call.

The UE also implements a de-jitter buffer to remove the jitter on the downlink (DL). The main causes of jitter on the DL are the HARQ transmissions and scheduling delays which may vary with loading. The UE de-jitter buffer sends the voice frames synchronously for play-out.

The main requirement at the NodeB for successful implementation of CSoHS is a scheduler which is aware of the delay sensitive nature of voice traffic. In HSPA, the traffic flows of all users share a common data channel on the downlink. The data flows can tolerate delays while the voice packets need to be delivered within certain delay bounds. Therefore the NodeB scheduler should give higher priority to voice flows, especially when they encounter large delays while waiting in the queue.

## III. SIMULATION MODEL

We study the capacity and performance of CSoHS via simulations. We consider a cellular wireless system where mobile terminals are served by base stations (NodeBs). The simulation framework is based on 3GPP simulation assumptions [5]. In particular, we consider a 57-cell deployment with wraparound and inter-site distance of 1 km. The channel model is a mix of Pedestrian A 3 km/h, Pedestrian B 3 km/h, Vehicular A 30 km/h and Vehicular A 120 km/h with respective ratios of 30%, 30%, 20% and 20%. We assume that TTI is 2ms on the uplink.

The voice sources used in our simulations are based on Markov models of the AMR 12.2 vocoder [6]. A full rate voice frame is generated every 20ms when the user is active (talking) and a silence insertion descriptor (SID) frame is generated every 160ms when the user is not talking. The size of a full rate voice frame is 244 bits for AMR 12.2. The SID frame size is 39 bits. The activity factor (the probability of being in active state) is 0.5 and the active and non-active times are exponentially distributed with a mean value of 2 seconds.

The main factors that determine the voice system capacity are the voice quality, uplink interference (system stability) and code and power limitations on the downlink. The voice quality is quantified here by the radio link delay and frame loss rate. A delay bound of 100ms is used on both the downlink and uplink such that the frames whose delays exceed this bound are discarded. It is assumed that a voice user is in outage if the total frame loss due to transmission errors and the delay requirement is more than 3%. The system capacity is reached when the number of such users is more than 5% of the total number of users. The system stability on the uplink is also determined by the total interference. We impose that the "Noise Rise" defined as the ratio of total received power to the thermal noise should not exceed 7 dB for more than 1% of the time. The system capacity on the uplink is taken as the minimum of this "Noise Rise" based capacity and the user outage capacity as defined above. For the downlink, the system capacity is determined only by the user outage which could also be caused by code limitation as will be discussed below.

We obtained the uplink and downlink system capacities by running simulations for each link where simultaneously active voice users are present in the system. An equal number of UEs are dropped in each cell randomly with uniform distribution and with interference from all UEs in the 57-cells, with the wrap-around model, considered. The number of UEs is varied and the performance metrics are computed to determine capacity according to the defined criteria.

For the downlink simulation, we also incorporate a receiver implementation margin. For a Rake receiver, 0.8dB is assumed to account for the loss due to RF front-end, channel estimation and other losses. For an equalizer receiver, SINR loss is added as a function of different channel types to the ideal equalizer output to accurately model the practical equalizer implementation.

It is assumed that the pilot channel on the uplink (DPCCH) uses the so called "slot 1" format where 8 pilot symbols are transmitted every 2ms.

The power gains of traffic and overhead channels have an impact on the spectral efficiency on the uplink. By simulation and analysis, we have determined the optimal settings which minimize the required Eb/No. These are tabulated in Table 1 as the ratio (offset) of a channel's power to the DPCCH power.

**Table 1 Channel Power Offsets on HSUPA**

| | EDPDCH 120 bits | EDPDCH 307 bits | HS-DPCCH | EDPCCH |
|---|---|---|---|---|
| Power Offset [dB] | 7.11 | 9.92 | 0 (SHO), -2 (non SHO) | 0 |

On the uplink, the number of HARQ retransmissions provides a trade-off between time diversity gain and delay. We simulated both a maximum of 3 and 4 attempts such that power control keeps the residual physical layer packet error rate at 1% after the maximum is reached. On the downlink, the code usage of extra HARQ attempts cancels the benefit of time diversity. Therefore the downlink packet error rate is adjusted so that 90% of packets are decoded successfully on the first attempt and a maximum of 4 attempts is used.

By using the Discontinuous Transmission (DTX) feature defined in CPC, a CSoHS UE now transmits the pilot channel only at the allowed DTX cycle times and data packet transmissions. This provides not only system capacity gains by reducing the total interference but also reduces the battery usage. Here we assume that DTX cycles are 4 and 8 TTIs which mean that the UE transmits data only every 4 TTIs when the user is talking and 8 TTIs when the user is listening. The switch from 4 to 8 TTIs happens when the user does not transmit on the uplink for more than 4 TTIs. The standard also defines preamble and postamble which are pilot transmissions before and after the DTX cycle times. These are used to help the channel detection and estimation at NodeB. We assumed that the preamble and postamble are 2 and 1 slots respectively.

We have analyzed the effect of Interference Cancellation (IC) at the NodeB on CSoH performance on the uplink. Our IC implementation is based on the design described in [8] when applied to HSUPA and described in more detail in [2].

A Channel Quality Indicator (CQI) value is reported by the User Equipment (UE) for downlink transmission selection and scheduling. We have observed by simulation that reporting CQI only every 16ms results in negligible reduction in the downlink CSoHS capacity. For the HS-DPCCH, we used a higher power offset for a user in soft-handover as shown in Table 1 since HS-DPCHH is decoded by the serving NodeB only.

On the downlink, the MAC-hs scheduler at the NodeB needs to be both delay sensitive and channel sensitive. Thus it takes into account the delay constraints while still exploiting the multi-user diversity. For the results in the paper, we use a scheduler that prioritizes and allocates resources to CSoHS users according to the following priority metric:

$$priority(u) = \log_{10}\left(\frac{R^{u}{}_{requested}}{R_0}\right) + \min\left(\left(\frac{D^{u}{}_{HOQ}}{d_{bound}}\right)^2, (\delta)^2\right) \tag{1}$$

The first term of the metric in (1) is the channel sensitive part that represents the channel quality estimate for a user by its requested rate obtained from the CQI information. The second term is the delay sensitive part represented by the ratio of head-of-queue packet delay of the user and the delay bound. The scaling constants are used to adjust the tradeoff provided by the two non-linear functions.

The scheduler may serve multiple users in the same TTI, as well as bundle multiple vocoder packets for a user in a single MAC-hs packet whenever necessary. The scheduling algorithm is an empirically developed variant of a number of QOS aware channel sensitive scheduling algorithms discussed in literature [7].

## IV. SIMULATION RESULTS

Our simulations show that the bottleneck of the system capacity is the uplink interference if the UE has dual receive antennas; otherwise, it is the downlink code or power usage limit.

The overall CSoHS capacity is the minimum of the UL and DL system capacities for particular configurations. The system capacity on the uplink corresponds to the number of users for which the Noise Rise exceeds 7 dB 1% of the time. The "Noise Rise" for the IC cases is measured after the cancellation takes place. In other words, it is not measured at the antenna input but it is the "Effective Noise Rise" seen by the traffic packets [9]. The capacity numbers according to the "Noise Rise" criteria is shown in Table 2 and 3. We note that the system capacity according to the user outage definition was higher than the "Noise Rise" capacity shown here in all simulations and therefore are not tabulated separately.

**Table 2 Release 99 Voice Capacity**

| | Downlink | | Uplink |
|---|---|---|---|
| | 1Rx Rake | 1Rx Equalizer | |
| Capacity | 51 | 63 | 69 |

**Table 3 Uplink CSoHS Capacity**

| | Release 8 | Release 8 w/ DTX | Release 8 w/ IC | Release 8 w/ DTX, IC |
|---|---|---|---|---|
| 4 HARQ attempts | 93 | 121 | 118 | 165 |

On the downlink, the CSoHS system performance is significantly affected by the MAC-hs scheduler located at the NodeB. In addition to the variable delay introduced by Hybrid ARQ retransmissions, the scheduler introduces queuing delays for the downlink traffic.

The vocoder packets that cannot be played out in time are effectively lost even if they are transmitted and decoded correctly over the air. Hence, for supporting high CSoHS traffic load, the scheduler must trade-off channel sensitivity and delay sensitivity.

As it can be seen from Table 4 below, significant gains in downlink CSoHS capacity are obtained by the use of equalizers in the mobile receivers instead of RAKE receivers.

**Table 4 Downlink CSoHS Capacity**

| Deployment | Release 8 RAKE | Release 8 Equalizer |
|---|---|---|
| 1 Rx | 61 | 83 |
| 2 Rx | 172 | 232 |

The end-to-end (or mouth-to-ear) delay for voice is a very important metric of call quality. As mentioned above briefly, there are several factors which contribute to the delay. These include transmission delays over the air, scheduling delay at NodeB and delay introduced by the de-jitter buffers at the UE and the RNC. In order to determine the cumulative delay for a mobile-to-mobile (M2M) call correctly, we performed the simulations in two steps. First, uplink only simulations were carried out and packet traces were obtained in addition to other system wide performance metrics. Then these packet traces which contain timestamp and transmission delays were fed to a downlink only simulation to obtain the total delay. Figures 2 and 3 show the combined uplink and downlink over the air delay at the system capacity values. Here it was assumed that the total fixed delay not including the Radio Access part was 87ms. The components which contribute to this delay are shown in Table 5. The 98$^{th}$ percentile of total delay at the capacity of Release 8 system with IC is shown in Table 6. These show that the total radio-link delay will satisfy the ITU recommendations [10] for a satisfactory voice call.

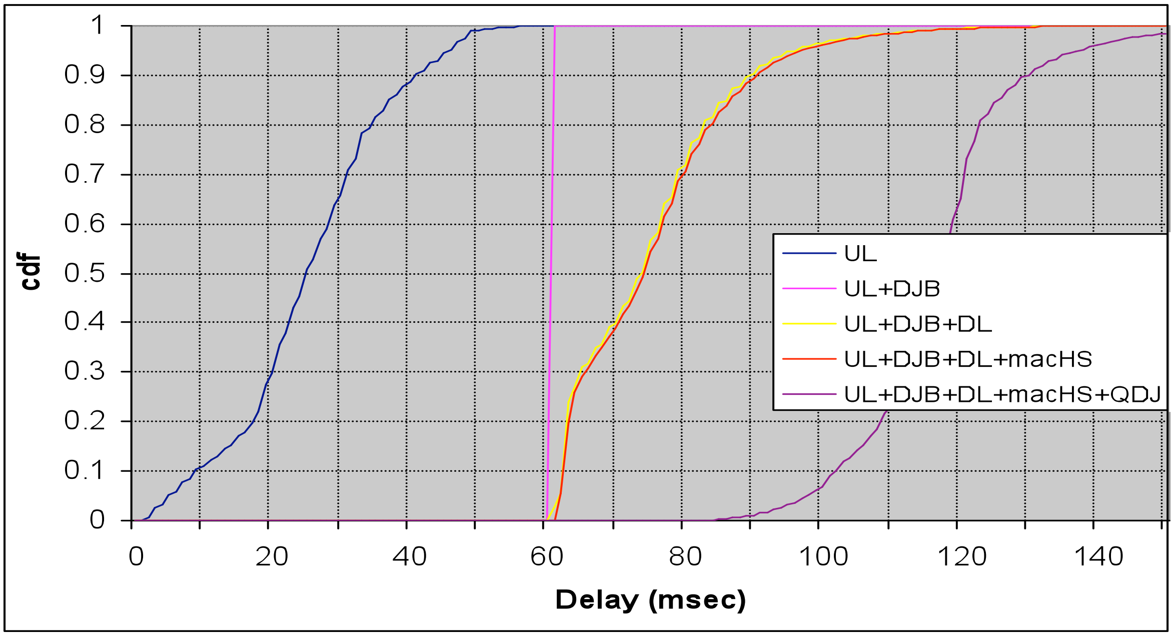


**Figure 2 Total Delay for UEs with Rake Receiver**

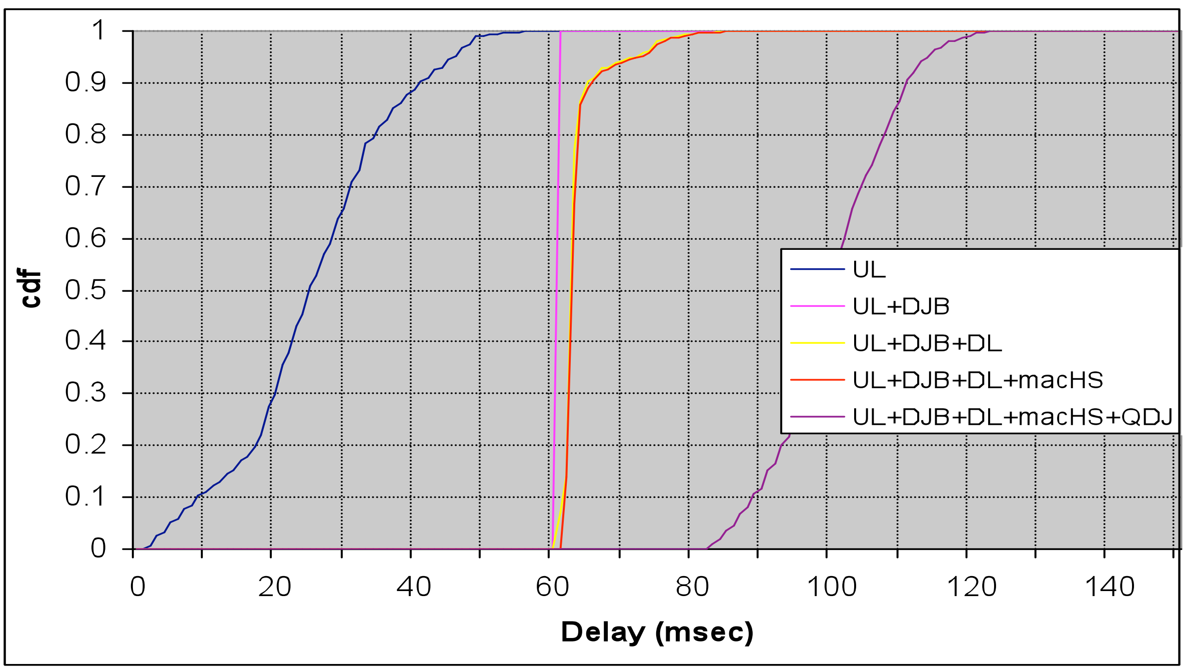


**Figure 3 Total Delay for UEs with Equalizer Receiver**

**Table 5 Fixed Delay Components**

| **Delay Componet** | **[ms]** |
|---|---|
| Source Vocoder (Accumulate 20ms, Lookahead 5ms, Processing 5ms) | 30 |
| Packet Processing (UE + NodeB) | 20 |
| RAN (Backhaul 10ms, RNC Processing 2ms, IuCS ~0) | 24 |
| Core Network Delay | 10 |
| Voice Frame Decoding | 3 |

| Total Fixed Delay | 87 |
|---|---|

**Table 6 System Capacity and M2M Delay for Release 8 w/ IC**

| Max UL transmissions | UE Receiver | System Capacity | M2M Delay |
|---|---|---|---|
| 3 | Rake | 153 | 258 |
| 3 | Equalizer | 153 | 235 |
| 4 | Rake | 165 | 283 |
| 4 | Equalizer | 165 | 256 |

### *A. Mixed Data and Voice Traffic*

In this section, we look at the performance of a typical operational system scenario when both voice users and best effort (BE) users are present. The BE traffic is generated by a full buffer traffic source. On the downlink, both types of traffic share the common channel. The NodeB HSDPA scheduler gives higher priority to delayed CSoHS users over BE users. The priority metric is a function of instantaneous channel quality and head of queue packet delay of the user. Among the BE users, the priority is decided by a proportional-fair metric which is the first term in Eq. (1). It is assumed that both HSPA BE users and Release 99 voice users have dual receive antennas. As the number of CSoHS users increases, the throughput served to BE users decreases because CSoHS traffic consumes both power and orthogonal code space. It is important to quantify the amount of BE throughput when there are voice users present in the system for evaluating the overall performance of the system. In particular, we look at the downlink throughput since the bottleneck for the voice capacity is the uplink (assuming mobile receive diversity). Figure 4 shows that significant amount of BE traffic can be served on HSDPA at the voice capacity of Release 99 CS system, if the voice calls were carried over CsoHS service.

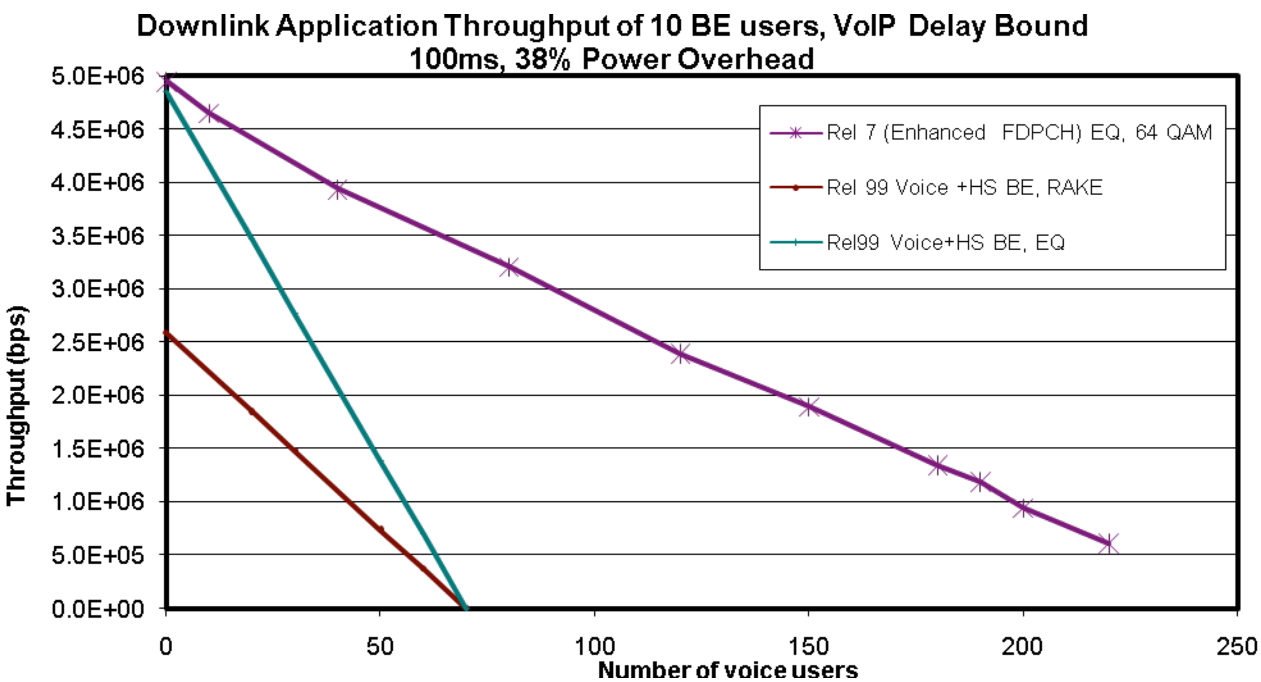


**Figure 4 Downlink Mixed Traffic Throughput**

## V. CONCLUSION

3GPP Release 7/8 provides efficient support for CSoHS voice. The simulations show that Release 7/8 CSoHS can support significantly higher number of voice users compared to Release 99 CS system. Using Interference Cancellation, the capacity can be further improved on the uplink. Simulations also show that a significant amount of BE traffic can still be served on the downlink at the CSoHS capacity operating point.

On the uplink, the capacity improvements of CSoHS are mainly due to time diversity gain by the use of HARQ and the DTX features which reduce the total transmitted power and interference.

On the downlink, the high CSoHS capacities obtained can be attributed mainly to the use of delay and channel sensitive MAC-hs scheduling and resource allocation, the use of F-DPCH channel, and the presence of receive diversity and advanced receivers/equalizers at the mobile.

## REFERENCES

[1] 3GPP 25.214, V.7.0.0, *"Physical layer procedures (FDD)"*
[2] Ozturk et. al, *"Performance of VoIP Services over 3GPP WCDMA Networks"*, PIMRC 2008, Cannes, France
[3] R2-074933, *"Introduction of CS voice over HSPA",* 3GPP RAN-WG2 #60
[4] R2-075403, *"Introduction of CS voice over HSPA",* 3GPP RAN-WG2 #60
[5] 3GPP 25.896, V.6.0.0, *"Feasibility Study for Enhanced Uplink for UTRA FDD"*
[6] 3GPP 26.071, V.7.0.1, *"AMR Speech Codec: General Description"*
[7] S. Shakkottai and A.L. Stolyar, *"Scheduling Algorithms for a Mixture of Real-Time and Non-Real-Time Data in HDR",* Bell Laboratories Technical Report, October 2000
[8] J. Hou, J.E. Smee, H.D Pfister, and S. Tomasin *"Implementing interference cancellation to increase the EV-DO Rev A reverse link capacity"*, IEEE Communications Magazine, Volume 44, Issue 2, Feb. 2006, pp.58-64.
[9] D. Zhang, S. Sambhwani, and B. Mohanty, "HSUPA Scheduling Algorithms Utilizing RoT Measurements and Interference Cancellations", ICC 2008, Beijing China
[10] ITU-T Rec. G.114, *"One-Way Transmission Time"*, Feb 1996\