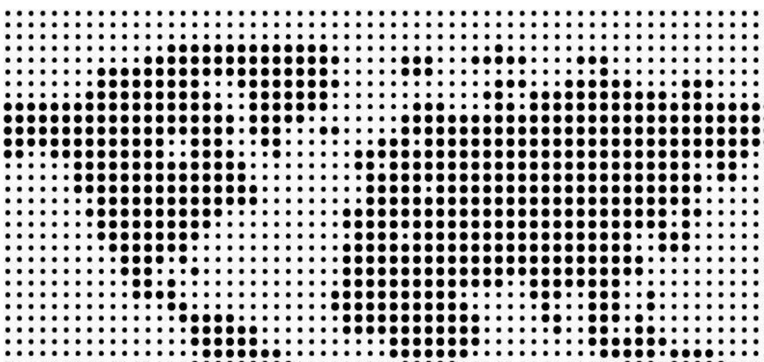
VoIP Overview



Qualcomm Incorporated
August 2008

Table of Contents

[1] Introduction

Consumers and business professionals alike desire rich and simultaneous multimedia services. Voice is an integral part of these services, and its high performance criteria require special attention to be paid to Quality of Service (QoS). With the advancement of the Internet and the increasing penetration of broadband cellular networks, it is now possible to offer voice services on packet-switched cellular networks with no compromise in voice quality. Voice over IP (VoIP) enables tight integration of all services, devices, and networks with a common IP platform.

3G and Evolved 3G networks, such as EV-DO, and HSPA+ have been designed to handle VoIP traffic efficiently. Furthermore, emergence of new multimedia applications causes the nature of data traffic to be more diverse. Besides general file downloads and uploads, applications like video telephony, Push-to-Multimedia, and simultaneous VoIP and gaming are also gaining in popularity. VoIP can easily be integrated with other rich multimedia applications. When all voice traffic is carried on an IP-based network, an operator is no longer required to maintain a separate network for circuit-switched voice, leading to savings in capital and operating expenses.

Evolved 3G also offers other standardized voice enhancements such as Circuit Switched Voice over HSPA (CS over HS) and 1X Advanced. CS over HS more than doubles the voice capacity of WCDMA R99 networks by leveraging benefits of VoIP without impacting the core network infrastructure. 1X Advanced is a natural next step for today's CDMA2000 1X networks that quadruples the voice capacity.

In this document, the benefits of VoIP are illustrated using EV-DO Rev. A. and HSPA+ as examples of cellular networks.

[2] Characteristics of Voice Traffic

The data rate of voice traffic depends on the codec, but is typically within 10 kbps. Speech patterns are discontinuous in nature, as speakers pause within and between sentences. In addition, since voice call is a two-way application (more if there are over 2 parties in a call) and only one person speaks at a time, the voice activity factor of each party in a call is less than 50% on average.

[3] Nature of 3G Cellular Networks

Early generations of cellular systems (1G and 2G) were designed to primarily handle voice traffic. On the other hand, 3G and Evolved 3G systems are expected to handle both data and voice traffic. Different applications have different Quality of Service (QoS) requirement. A well-designed system will prioritize these applications based on their respective QoS requirement.

[4] Challenges

The interactive nature of voice conversation implies that users are sensitive to latency and jitter (variance in latency). Excessive latency and jitter will degrade voice quality. End-to-end latency within 300 ms is generally considered acceptable. Circuit-switched systems are immune to jitter as a channel is dedicated for each call. For voice quality to be acceptable in packet-switched systems, latency and jitter both have to be well controlled.

Given that VoIP service will be provided in the presence of other data services, the network has to efficiently handle both kinds of traffic. The challenge is to ensure that the latency of VoIP traffic can be maintained even when the network is loaded with other data traffic.

Operators will expect VoIP to provide comparable, if not better, capacity than traditional circuit-switched voice service. A cost of carrying voice over the IP-based network is the extra overhead from the various protocols. The extra overhead is significant and, if not optimized, can consume a large part of the airlink capacity that can otherwise be used to accommodate more VoIP users. The design of VoIP system also needs to address users' desire for longer battery life, which translates to less frequent recharging of the phones.

Users in a cellular network are not tied to any location and can move around freely. This places a requirement on the system to provide continuous service to the users with no noticeable impact on user experience as they move across cell-sites as well as when user moves out of VoIP coverage in to Circuit-switched voice coverage.

Techniques applied in 3G and Evolved 3G technologies to overcome these challenges will be introduced in this document.

[5] 3G Supports Efficient VoIP

3G and Evolved 3G incorporate features that make efficient commercial VoIP services a reality. They enable operators to over come the challenges introducing voice on a IP based network and provide consistent user experience

5.1 QoS

In order to provide a user experience comparable to that of circuit-switched voice, latency has to be well-controlled. Higher priority needs to be given to voice traffic so that the voice packets can be processed ahead of other traffic types. The network has to identify the QoS profile of each data flow and then prioritize them accordingly.

In the Forward Link (FL) or Downlink (DL), where a centralized scheduler is used, the scheduler can ensure that the delay of the voice packets meets their QoS requirement. In the Reverse Link (RL) or Uplink (UL), different cellular technologies have different ways of controlling RL airlink latency. For example, the RTC MAC protocol in the EV-DO Rev A system can assign a shorter Termination target to voice packets such that they can be received at the Access Network with a lower latency than packets for other best-effort applications such as email. The protocol can also guarantee that voice capacity and latency are unaffected by sector loading.

5.2 Adaptive De-Jitter and Time Warping

While the vocoder in the voice terminal of a caller generates a voice frame periodically (usually once every 20 ms), these frames will not arrive at the callee’s phone at the same rate or even in the same order due to the varying condition of the wireless channel and also the different routes taken by the different voice frames in a packet-switched network. This results in delay jitter. A de-jitter buffer smoothes out the jitter by introducing an additional delay that changes adaptively. Even if a particular packet incurs a higher delay than other packets, it can still be played back at the callee’s phone without any noticeable impact on voice quality. To handle the varying latency of different voice packets, another technique known as time warping is applied. Time warping adjusts the playback duration of a voice packet without changing its pitch. It reduces the delay that needs to be introduced by the de-jitter buffer.

5.3 Handoff and VCC

As users move around, they may cross the coverage boundary of different cells and handoff from one cell to another. In EV-DO Rev A, the terminal can indicate its desire to handoff in advance to the access network through the DSC Channel so that the new cell can prepare the resources for redirection of the call.

Moving users can enter geographic areas with no VoIP coverage, but where circuit-switched voice service is still available. Voice Call Continuity (VCC) enables a smooth handoff of voice calls from VoIP to the circuit-switched network. This is important as VoIP may be deployed in phases with gradually expanding coverage, while the legacy circuit-switched network has ubiquitous coverage. A number of supporting standards that enable VCC are being published in the 3GPP and 3GPP2 standard bodies.

[6] Areas of Optimization

For VoIP deployment to provide a cost benefit to operators, convenience to users, and acceptable voice quality, VoIP systems must both be designed with the requirements listed in the previous sections and be optimized for the following goals.

6.1 Capacity

The cost of providing voice service can be reduced if more voice calls (users) can be accommodated simultaneously. Several techniques have been introduced in different cellular systems to increase voice capacity.

6.1.1 Header Compression

One capacity-increasing technique that can be applied to any cellular technology is Robust Header Compression (RoHC). A typical voice packet carries a very small payload, usually 32 bytes or less. On the other hand, the packet header (IPv4 header, UDP and RTP) has a size of 40 bytes. Thus, only 44% of the data is useful payload. Such inefficiency may not affect voice quality but will reduce voice capacity, as a great deal of power is spent on overhead, which in turn becomes interference to other users. Since much of the information in the header does not change from one packet to the next, it is possible to compress the entire header to 2 to 3 bytes, hence significantly increasing the efficiency to 90%. The benefit of this is a significant improvement in voice capacity.

6.1.2 Smart Blanking

Human callers are used to hearing some background noise in phone conversations. This does not mean the phone has to constantly send background noise to the other party during silence periods. Instead, the phone can transmit a background noise packet at the beginning of a silence period, and only update the background noise information when there is a significant change. The receiving phone can repeatedly play back the last packet of background noise until a new packet is received.

6.1.3 DTX

The voice activity factor of any party in a call is less than 50%, since a caller is typically silent for more than half of the time. When there is no voice or control traffic, the transmitter can be turned off so that it generates no interference to other users. This is known as Discontinuous Transmission (DTX). Without DTX, pilot and other overhead channels would still be sent during silence periods. This feature has been introduced in HSPA+ and EV-DO Rev. B.

6.1.4 Interference Cancellation

In technologies where different users transmit in the same time and frequency, thus causing interference to each other, Interference Cancellation (IC) can be introduced to mitigate the impact of interference. IC can provide about 45% increase in VoIP capacity.

6.1.5 4-Way RX Diversity

At the base station, if four receive antennas (4-way RX diversity) are deployed, RL capacity proportionately improves as each user can transmit less power, thus generating less interference to other users. For VoIP, due to the large number of simultaneous users, the full potential of 4-way RX diversity may not be realizable, as other constraints may begin to limit the number of users in a sector. For EV-DO Rev A, the voice capacity in a single 1.25 MHz carrier increases from 41[1] to 66 per sector with 4-way RX diversity. In this case, the number of MAC indices is the limiting factor for voice capacity. Aside from the 66 users in a sector, when soft handoff users are included, the total number of users in each sector reaches 114, which is the maximum number of MAC indices available per sector.

[1] *EVRC, RL 2-way diversity, no IC, no DTX/DRX, inter-base station distance of 2 km.*

6.1.6 HS-SCCH-Less Operation

This technique is specific to HSPA+. In HSDPA, since DL HARQ transmissions are asynchronous, HS-SCCH has to be transmitted with every data (HS-PDSCH) transmission to notify the user equipment (UE) of the next data transmission. With HS-SCCH-less operation, it is no longer necessary to transmit HS-SCCH with the HS-PDSCH retransmissions. Instead the UE will blindly decode HS-PDSCH. This provides significant savings in power and increases DL data capacity accordingly.

6.2 Improved Battery Life

Since cellular phones run on batteries, it is desirable to maximize the standby and talk time, so users do not have to recharge their phones as often.

6.2.1 DTX and DRX

The DTX scheme described in the previous section also contributes to a longer talk time as the transmitter is turned off between talk bursts. Another scheme known as Discontinuous Reception (DRX) has a similar benefit by turning off the receiver when no information is expected to be received at the phone. For HSPA+, combination of DTX and DRX can increase talk time by up to 50%. DRX was also introduced in EV-DO Rev B.

Qualcomm's MSM and CSM chipsets currently support many of these optimization, and plan to support all of them in the future.

[7] Data Capacity in Mixed Traffic

3G systems support a rich mixture of voice and data traffic. Hence, not only does a well-designed system have to have a high VoIP capacity, it is also necessary to ensure the data users get reasonable data capacity in the presence of VoIP users. The introduction of VoIP users to a system takes up airlink resources, which leads to a drop in data throughput. A packet-based system has the advantage of being able to handle a mix of data and voice users efficiently, thus minimizing the impact on data throughput. The following three figures show that FL best-effort (BE) throughput for the three different technologies decreases slowly as the number of VoIP users increases. The leftmost points on all three figures are the data-only capacity in the absence of voice users.

The data capacity is shown in absolute terms in Figure 7-1 for HSPA+, and comparison is shown with HSDPA (R5), which uses circuit-switched voice. Although the data-only capacity is almost identical for both technologies, HSPA+ clearly outperforms HSDPA circuit-switched voice as more and more voice users are added.

In Figure 7-2 the data capacity is shown in relative terms for EV-DO Rev. A., and the trend is similar to that observed in Figure 7-1. In each figure, the data-only capacity of one of the configurations is used as reference, with normalized capacity set to 1. Also note that at full VoIP capacity, there is still some FL capacity remaining. This is due to the fact that VoIP capacity is often limited by RL, thus allowing FL-only applications such as best effort download or Platinum Broadcast (PB) to utilize the extra capacity on the FL.

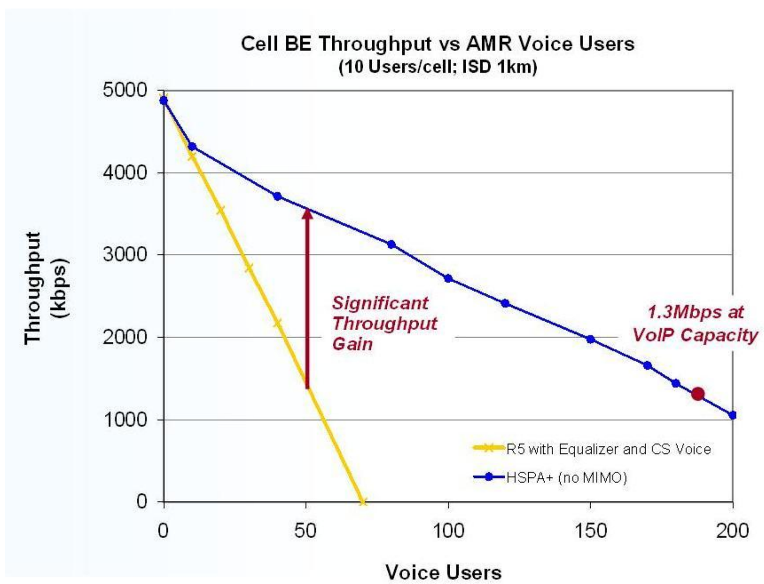


Figure 7-1. HSPA+ BE DL Capacity in Mixed Data and Voice Traffic

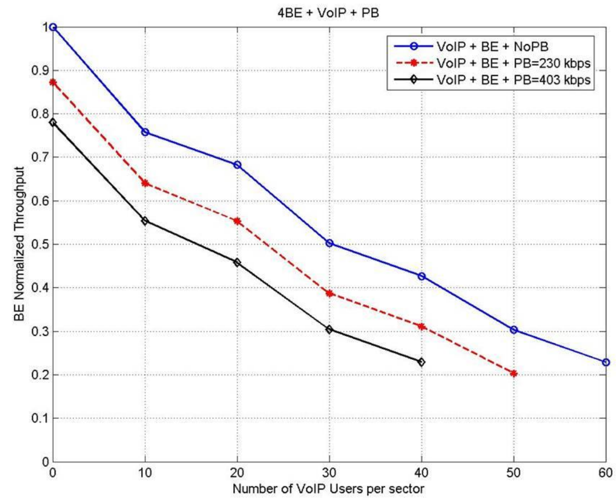


Figure 7-2. EV-DO Rev A BE FL Normalized Capacity in Mixed Traffic Scenario

[8] Conclusion

A number of design objectives and benefits of VoIP have been addressed in this paper. For a VoIP system to provide a similar user experience to existing circuit-switched service, it is necessary to have tight control of latency through QoS, adaptive de-jitter buffer and time warping. Moreover, VCC enables a smooth handoff from VoIP to circuit-switched networks, so that users can use their voice handsets wherever there is VoIP or circuit-switched coverage.

The benefit of VoIP for operator is the potential cost reduction, made possible by a single common IP-based network that handles both data and voice.

At the same time, enhancements such as 4-way RX diversity and IC help improve the capacity for VoIP. This capacity improvement is realized in both a loaded VoIP scenario, and a mix of VoIP and data traffic. The ability to efficiently handle simultaneous voice and data applications maximizes the data capacity as VoIP users are introduced in the system. At full VoIP capacity, there is additional FL capacity that can be allocated to applications such as BE download and multicast.

The standardized Evolved 3G voice enhancements such as CS over HS more than double the voice capacity of WCDMA R99 networks by leveraging benefits of VoIP without impacting the core network infrastructure, as well as 1X Advanced, a natural next step for today's CDMA2000 1X networks quadruples the voice capacity.

Voice over IP and the other voice enhancements clearly bring significant benefits for operators and end users alike. 3G and Evolved 3G systems such as EV-DO, and HSPA+ have specifically been designed to deliver high quality and high capacity voice services that offer the basis for many real time services with excellent voice and multimedia streams.