Whitepaper

Samsung Wireless Enterprise

Samsung
Smart WLAN Solution

Smart Capacity & Security
for Smarter Mobility

Voice Optimization

THE NEW BUSINESS EXPERIENCE

Introduction

In our modern world, enterprises are in constant need to provide their employees with Internet access and communication tools regardless of their location. By adding voice to their services, they can improve collaboration and responsiveness, increasing the efficiency of their businesses. Voice over Internet Protocol (VoIP) is a technology that allows users to make and receive calls using an Internet connection or other data connection, instead of traditional phone lines. VoIP services have reached an enormous popularity mainly due to the fact that they offer the possibility of voice communication at a relatively low cost and easy maintenance, using the existing Internet access networks.

Meanwhile, Wireless Local Area Networks (WLANs) have been widely deployed for ubiquitous Internet connectivity. They are completely pervasive in the enterprise, to the point that they have become the dominant wireless technology to access the Internet. The combination of VoIP and WLAN (VoWLAN) represents a very promising approach for voice communication in the enterprise. VoWLANs allow businesses to enjoy phone services regardless of their location at a considerably reduced cost.

Despite all the advantages VoWLANs have to offer, guaranteeing a satisfactory Quality of Service (QoS) for the enterprise user is a challenging but crucial task. Voice traffic is subject to very tight constraint on the system performance and the overall status of the WLAN. These unique performance requirements are vastly different from data traffic, for which WLANs were originally designed.

In order to overcome the challenges and support voice traffic in WLANs and maintain an acceptable QoS level, numerous efforts have been made. The IEEE 802.11e amendment to the legacy 802.11 MAC aimed to accommodate QoS-constrained traffic transmission allowing the use of multimedia traffic and improve the channel efficiency by introducing differentiated service mechanisms for different types of traffic and a new channel access function called Enhanced Distributed Channel Access (EDCA), which is built on top of the existing one.

Unfortunately, despite all these enhancements, satisfactory levels of QoS for the enterprise are hard to guarantee. The reasons behind this are, among others, the random nature of the channel access mechanisms, the inefficient overhead included in transmissions, and the wasted resources due to packet losses and excessive delay that forces the network to drop packets.

In this paper we present Samsung Electronics' efforts to support voice traffic over enterprise WLANs. Among the algorithms we can find: Voice -aware Traffic Scheduling (VaTS) and Network Controlled Voice Optimization (NCVO) that attempt to arrange the enterprise WLAN to achieve an optimized performance for voice traffic. VaTS aims to improve the utilization of network resources and make the overall network performance more efficient. It reduces the amount of overhead that hinders the correct operation of voice traffic and employs the network resources to transmit voice data packets. Similarly, NCVO's objective is to reduce the probability of packet collision by finely tuning the MAC parameters according to the network status. Eventually, the combined use of these two techniques provides an optimized QoS to the voice clients.

Voice-aware Traffic Scheduling

One of the most important issues when designing an enterprise WLAN with voice traffic is the call capacity, in other words, the maximum number of simultaneous voice connections that can be supported by the WLAN. Besides the constraints on delay and packet loss that the voice traffic is subject to due to its real time nature, voice networks experience a capacity upper bound considerably lower than the capacity one might expect for such a small load.

Let's study this phenomenon in more detail. First of all, it is worth mentioning the characteristics of the voice encoder. The G.711 is the international ITU-T standard for encoding telephone audio, which has a fixed bit rate of 64 kbps, which imposes a very small load on the overall network capacity. The common value of the packet generation interval is 20 ms, which corresponds to a packet arrival rate of 50 packets/s. Hence, a data packet of 160 bytes is periodically generated.

A naïve analysis of the network capacity would consist on the theoretical throughput of the network divided by the load of a voice call. For example, a IEEE 802.11n Access Point (AP) serving smartphones using MCS 7 would have a transmission rate of 150 Mbps, so theoretically it could accommodate 150 Mbps/ 64 Kbps ≈2343 calls. The real value experienced by voice WLANs is below

39 calls, shockingly lower than expected.

Voice Optimization

The reason behind this huge difference lies in the amount of overhead the network has to add to the voice data in order to be transmitted. Digital voice streams, at the output of the codec, are packed into constant-bit-rate (CBR) voice packets with small payloads. Each voice packet is appended to a 12 bytes long Real-Time Transport protocol (RTP) header, 8 bytes long User Datagram Protocol (UDP) header, 20 bytes long Internet Protocol (IP) header, 8 bytes long Link Layer Control (LLC) header and finally, 36 bytes for the IEEE 802.11n MAC. A final physical layer header (36 µs for IEEE 802.11n Mixed-Mode configuration) is incurred for every single packet transmission. To put things into perspective the transmission duration of a single voice data packet would take 48 µs out of which only 8 µs are actual voice data.

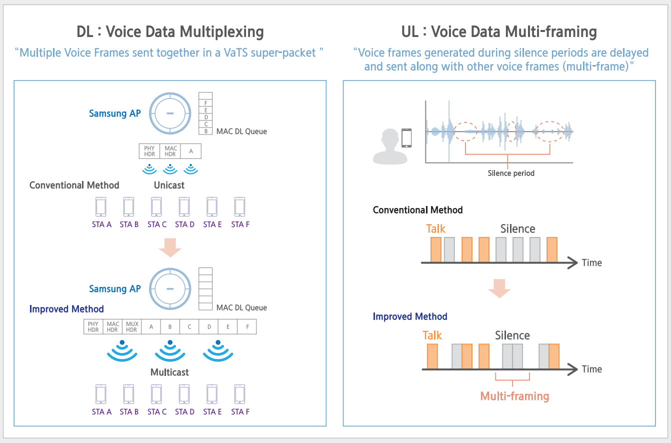
We define wireless resources as the period of time the network allocates for clients to transmit their data packets, also known as airtime resource. Since all the stations in the network have to share the medium and therefore can only transmit one at a time, the resource is the airtime that they consume in the transmission of their data packets when they access the channel. Due to the small payload size of the voice packets, we have observed that most of the airtime resources are being used on over-head transmission rather than effective voice data. Voice-aware Traffic Scheduler (VaTS) aims to reduce the redundant overhead in voice frame transmission increasing the efficiency of the network resource utilization and maximizing the system capacity.

The principles of VaTS lie in the nature of IEEE 802.11 WLANs and their inherent broadcasting nature. Voice-aware Traffic Scheduler assesses the quality of the environment and quantifies the number of active voice clients in the network. VaTS tries to reduce the network overhead by creating a super-packet constructed with voice packets addressed to different voice clients, thus increasing the voice payload and effectively reducing the unnecessary overhead. By doing so, the physical layer head incurred for every packet transmission only has to be transmitted once to forward voice packets in the super-packet to different voice clients instead of with each one individually. Note that by attaching multiple voice packets together all the airtime resources employed for inter frame spacing and random back-offs are also saved and can be used for effective transmissions.

The goals of the Voice-aware Traffic Scheduler algorithm are:

- Reduce unnecessary overhead
- Minimize the wasted resources in silence periods
- Maintain the performance level
- Operate accordingly to the IEEE standard

Let's compare the standard transmission protocol and VaTS in the diagram below:



VaTS operation diagram: Multiplexing and multi-framing

Unlike the standard downlink transmission protocol, whereby packets destined to different stations queue to be individually transmitted, VaTS has a multiplexing stage prior to transmission. The VaTS groups the voice packets destined for different clients and creates a super-packet that is eventually sent to the medium for all clients to hear.

After being enqueued in the AC_VO queue, data packets wait for the multiplexer entity to create a VaTS packet every scheduling period (i.e., 2 or 4 ms). The delay introduced is small enough so it does not affect the delay constraints of the voice packet, and consequently, the voice performance is not degraded.

Let's observe the structure of the super-packet formed by the multiplexer after aggregating multiple voice data packets:

IP	UDP	SEQ	MUX	RTP	Voice data	RTP	Voice data	RTP	voice data

VaTS multiplexed packet structure

Voice Optimization

The information contained in the RTP header and the voice data fields is used to create the VaTS packet. The structure shows a single header (IIP, UDP, SEQ and MUX) for all the aggregated voice data frames. Since we only have to send the PHY header once for the entire aggregated packet, we can see the overhead considerably reduced when VaTS is used.
Packets multiplexed into the super-packet are destined to multiple stations; therefore instead of being forwarded in a unicast fashion, VaTS packets with voice data are multicasted into the medium for all the STAs to hear. Multicast mechanisms allow for the conservation of bandwidth, making use of the inherently broadcast nature of the wireless medium.
No positive acknowledgement of correct reception is required, since receiving ACK frames from all the network clients would incur a large overhead and raise issues about scheduling and synchronization of receiving them.
Nevertheless, in order to guarantee satisfactory performance levels, other safety measures to acknowledge failed transmissions need to be implemented such as negative acknowledgement, where stations only send an acknowledgement packet requesting the retransmission of an unsuccessful frame. Negative feedback saves the network unnecessary acknowledgement transmissions when the conditions of the medium are favorable.
In the uplink scenario VaTS uses a voice data multi-framing procedure. Effectively, the voice frames generated during silence periods are delayed and sent along with other voice frames (multi-frames).
In summary, the VaTS multiplexing stage is ultimately in charge of assessing the quality of the medium and aggregating voice packets intended for different stations into a VaTS super-packet considerably reducing the overall consumption of airtime resources. Single header and no positive acknowledge effectively reduces the amount of network resources required for voice data transmission. VaTS successfully increases the simultaneous call capacity of the network up to 50%. This increase in call capacity allows the application of VoWLAN with low cost in high-density user environments with a satisfactory network performance.

Network Controlled Voice Optimization

The baseline IEEE 802.11 MAC was originally designed for data centric traffic such as email sending or web browsing applications. The majority of the IEEE 802.11 devices operate with the mandatory coordination function, known as Distribute Coordination Function (DCF), which is based on Carrier Sensing Multiple Access with Collision Avoidance (CSMA/CA) that will be described in detail as follows:
DCF works with a single First-In-First-Out (FIFO) transmission queue and does not distinguish between the different types of traffic in the network. Packets from all natures (real time voice, video or non-real time data traffic) reach the top of the transmission queue and get ready to be transmitted. The transmitter MAC only transmits when the channel is not being used by other stations, therefore it senses the state of the channel to ascertain whether it is in use or not. In order to reduce the chances of simultaneous channel access by the stations that are contending to transmit, the transmitter MAC uses a back-off counter, a random integer drawn from an uniform distribution over [0, CW], where CW is the Contention Window size ranging from CWmin to CWmax. If the channel is detected busy, the MAC waits until it becomes idle. Once the transition to idle occurs, it postpones the transmission attempt for Distributed Inter-Frame Spacing (DIFS). If the channel stays idle during this period, then the MAC checks its back-off counter value. The back-off counter is decremented every slot time interval that the sensed medium remains idle. When the counter reaches zero, the transmitter MACstarts the packet transmission. The larger this CW is, the lower the collision probability gets. However, the efficiency of the network also decreases when the CW value is increased since the channel is not utilized.

After each successful reception of a packet, the receiving station acknowledges the transmission with a reply message known as (ACK) frame after a Short Inter-Frame Spacing (SIFS). If the packet is not received correctly, then the ACK frame is not transmitted and the transmitting station increases the CW size reducing the collision probability. It is considerably difficult to provide delay and packet loss guarantees. Moreover, the performance of the MAC can easily become the bottleneck of the network due to factors like collisions and channel errors.
As an attempt to accommodate 802.11 MAC to support real QoS constrained traffic, IEEE 802.11e standard was published in September, 2005. It defines a new MAC protocol called Enhanced Distributed Channel Access (EDCA) built on top of DFC as an enhanced variant of the basic features provided in DCF. First of all, the FIFO queue system, in which all kinds of traffic share the same queue is replaced in EDCA by four different Access Category queues (ACs). The traffic stream is divided according to its QoS constraints into different priorities traffic groups referred to as Access Category. There are four different ACs:

- Voice (VO)
- Video (VI)
- Best Effort (BE)
- Background (BK)

The MAC assigns different access parameters to each specific AC. Among the channel access parameters that change their value according to the type of AC we can find:

- The Arbitrary Inter-Frame Spacing (AIFSN [AC]) which substitutes the functions of DIFS and controls the amount of time to be spent within contiguous transmissions attempts giving different priorities to the packet transmissions.
  AIFS = SIFS + AIFSN[AC] *SlotTime
- Contention Window size CW[AC] (CWmin, CWmax) to give different access priorities by controlling the size of the back-off counter.
- Transmission Opportunity (TXOP) which is the time interval during which a particular station is allowed to transmit once the station gets access to the medium.

The values of AIFSN, CWmin[AC], CWmax[AC] and TXOPLimit[AC] referred to as the EDCA parameter set are determined and advertised to stations by the AP via Beacon frames. The fundamental access method in IEEE 802.11e EDCA is still based on CSMA/CA protocol.

Access Category	Traffic Priority	Traffic Type	ALFSN	CWmin	Cwmax	MAX TXOP
AC_VO	7,6	Voice	1 or 2	3	7	1.504ms
AC_VI	5,4	Video	2	7	15	3.008ms
AC_BE	3,2	Best Effort	3	15	1023	0
AC_BK	1,0	Back-ground	7	15	1023	0
Legacy DCF	N/A	Best Effort	2	15	1023	0

Default 802.11e EDCA parameters and values

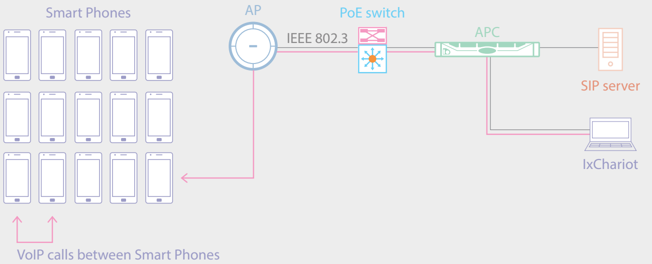
In order to give preference to higher priority ACs, they are allocated shorter AIFS and reduced contention window values that effectively increase the AC's channel access probability. However, the function of the back-off period is to avoid simultaneous channel access attempts by all the stations that need to transmit. Therefore, even though the stations with shorter back-off period access the channel medium sooner than others, they also experience higher collision probabilities between them, leading to the waste of network resources, potential packet drops and degraded voice performance.
Other improvements to WLANs that enable them to achieve higher throughput were brought by the IEEE 802.11n standard. The use of multiple antennas and the frame aggregation before transmission allowed the network to increase data rates and efficiency of transmission. Aggregated MAC protocol data units (A-MPDU) is a process of packing multiple MPDUs together to reduce overhead and average it over multiple frames, thereby increasing the user level data rate. This aggregation requires the use of block acknowledgement (BlockACK) which was introduced in the IEEE 802.11e standard.

Voice Optimization

AMPDU has proven an effective way to increase the network efficiency; however it imposes a new challenge on the voice traffic of the network. By aggregating transmission frames, clients take longer to transmit and therefore increase the contending times for other clients, which eventually result in higher competition for the channel access. A parameter of interest when dealing with voice traffic is the time required to transmit AMPDU. This parameter is known as the AMPDULimit, which describes the maximum time allowed for MPDU packet that can be aggregated and transmitted.
Samsung Network Controlled Voice Optimization algorithm treats Voice over IP as the most important application in wireless enterprise LANs. It aims to give priority to voice traffic above all other types of traffic and effectively guarantee a satisfactory quality of service for the voice user. It finely tunes the EDCA parameters and AMPDULimit according to the network status in order to make efficient use of the available resources and relax the collision and packet drop probabilities. Besides its main goal, NCVO also aims to:

- Optimize the performance and the Quality of Service of the voice calls
- Maximize the network call capacity

The operation of the Network Controlled Voice Optimization algorithm is straight-forward. After careful analysis of the voice enterprise scenario, Samsung has obtained the values of the MAC parameter set that translate into maximum quality of voice service (QoS) also allowing maximum throughput to lower priority traffic. The following scenario was studied:



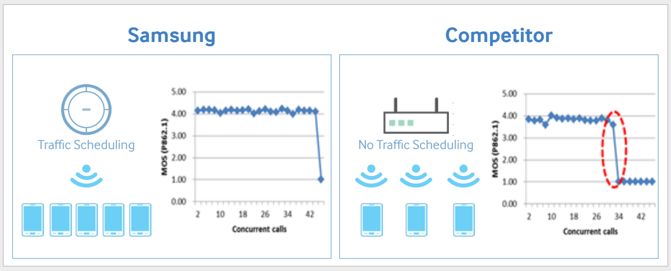
NCVO for enterprise WLANs test scenario

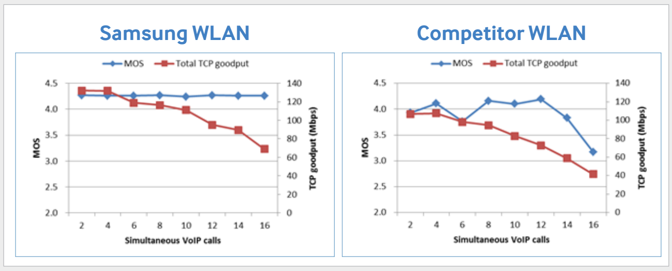
In this scenario, the number of simultaneous calls within the network was increased and the MAC parameter set values were modified. After a comprehensive test study and analysis the optimal parameters were identified according to the number of coexisting voice calls. Network Controlled Voice Optimization algorithm dynamically sets these values according to the traffic carried by the enterprise WLAN from an optimized parameter database.
Eventually, having the MAC parameters dynamically optimized according to the specific traffic carried by the network, translates into the use of an appropriate inter-frame spacing, random back-off settings and aggregation use according to the status of the network including the number of voice calls in service. In other words, the network MAC parameters are dynamically tuned in order to reduce the waste of resources, transmission jitter and frames loss by relaxing the contention in the network. This effectively allows voice calls to achieve maximum quality performance (i.e., average Mean opinion Score (MOS) of 4.0) and the TCP throughput of lower priority traffic is also maximized suffering the minimum impact possible from the voice calls.

Results

So far, we have presented the algorithms that Samsung enterprise Access Points utilizes to optimize the voice traffic performance in their WLANs and explained the principles behind their operation. In this section, we aim to illustrate the performance of Samsung enterprise networks and summarize the advantages of using Voice-aware Traffic Scheduling and Network Controlled Voice Optimization algorithms in order to maximize the efficiency of the network resources and provide optimal service for their voice clients.
In our practical scenario, let's imagine a generic office served by an IEEE 802.11n capable Access Point (i.e., Samsung WEA302i or an equivalent AP model from the competition). All the clients work on the 5GHz frequency band with 40MHz bandwidth. For the sake of the example, let's assume this particular office only carries voice traffic (e.g., a phone call or client service center). In this case, the performance of the enterprise WLAN is assessed with the Mean Opinion Score (MOS) of the calls. The MOS value ranges from 4.5 (optimal level) and below; we consider a value of 4.0 to be the minimum acceptable QoS for an enterprise WLANs.
The tests performed in this particular scenario show the following results when using the Samsung AP with the voice enhancements or an AP model of the competition:

Voice Optimization





From the graphs above we can observe the behavior of the voice traffic as we increase the number of simultaneous calls in the WLAN and the assessed MOS of the voice traffic. The voice capacity is reached when the MOS drops below 4.0.

It is made obvious from the graphs comparing the performance of the two AP models that the number of simultaneous voice calls cannot go over 34 when using the competitor's solution. However, when both Samsung enhancements are combined, a total of 44 simultaneous voice calls can be served with acceptable QoS for the voice clients. Samsung voice enhancements manage to reduce the overhead and minimize the number of collisions in the network and consequently it makes an efficient use of the network resources, effectively increasing the network capacity and allowing higher number of simultaneous calls with an acceptable QoS performance. It should be noted that when used individually, the algorithms are able to improve the voice quality of the network but it is when they are combined that the performance is optimized. In conclusion, the combined use of the VaTS and NCVO increases the overall voice call capacity.

In order to make our example more generic, let's observe the behavior in when voice and data services are used simultaneously in the enterprise WLAN. In this case, the performance of the network should also include the amount of downlink TCP goodput available for the background traffic. Increasing the number of downlink TCP traffic flows makes the strain on the voice calls more critical since the WLAN has to protect the high priority voice calls from all the lower traffic flows. The example below shows the behavior of voice calls and twelve downlink TCP flows to note-PCs.

The Samsung AP is able to guarantee the satisfactory QoS for the voice clients despite the presence of a heavy load of background TCP traffic. Since the EDCA and AMPDU parameters are optimized, collision and consequently packet drop probabilities are minimized and network resources can be used more efficiently. We can observe that Samsung AP can support higher number of simultaneous voice calls and higher goodput traffic for the TCP flows.

Conclusion

Samsung's efforts to protect the voice traffic in enterprise WLANs can be summarized in two algorithms: Voice-aware Traffic Scheduling that minimizes the unnecessary overhead and Network Controlled Voice Optimization that minimizes the collision and packet drop probabilities. These algorithms effectively improve the efficiency of the network and translate into the following of benefits for the customers.

Customer Value propositions:

- Mean Opinion Score is maintained at a satisfactory level under the presence of many concurrent calls.
- Samsung APs provide satisfactory voice service in enterprise environments where voice and data traffic coexist.

Competitive Strengths:

- Increases the simultaneous call capacity of the network up to 50%.
- This higher simultaneous call capacity in high-density user environments helps decrease the overall network cost.