INTEROP LABS VoIP and 802.11

The characteristics of voice traffic demand special attention on networks that were designed with data traffic in mind. Voice traffic consists of very short frames that require regular delivery. The telephone network was designed to offer minimal delay and have tightly controlled timing. The cost of optimizing the telephone network for voice transmission is that it is unsuited for applications that require different characteristics.

Transmitting voice traffic on a data network is a challenge because the network characteristics that lead to good VoIP performance are different from data. VoIP requires regular network service for very short packets; most of the methods for encoding voice will send data packets every 20 milliseconds. These packets must be transmitted almost immediately upon receipt. The best way to ensure that a packet can be transmitted immediately is to keep the transmit queues relatively empty so that waiting frames can be transmitted immediately. Good data throughput, however, depends on keeping transmit queues near maximum length so that long packet "trains" can be built. The best way to compromise is to build protocols that treat packets from different applications with appropriate priority.

Although prioritizing voice packets appropriately is the biggest hurdle for wireless LAN administrators, there are several other challenges as well. Users will expect that telephone calls will be smoothly handed off across access points, and that telephone calls receive the necessary share of limited network capacity of 802.11.

Prioritization in 802.11 with 802.11e and WMM

Telephone networks are engineered to provide high-quality service for just one application, the transmission of voice. Many people do not delve into the details of the research that created the telephone network and assume that voice on wireless LANs will "just work," especially since many of the newer wireless VoIP phones even look like their mobile phone brethren.

Without any special prioritization, a wireless LAN treats all frames equally. Frames go in a first-come, first-served transmission queue. The network does not distinguish between different types of data, which often means that voice frames will be delayed. Users perceive delayed frames as static, pops, or slight repetitions in the audio stream. To preserve voice quality across the wireless LAN, voice frames need to be treated as more important than frames carrying data for other applications. Fortunately, the priority of voice frames over other applications does not need to be absolute. Early experience indicates that if voice quality is as good as mobile phones, users will be satisfied.

Giving voice frames a higher priority than other applications is one of the major goals of 802.11e, which was published in 2005 [1]. Because of the interest in quality of service, the Wi-Fi Alliance worked to speed adoption by taking an early "snapshot" of the standard in progress and creating the Wi-Fi Multi-Media (WMM) certification program. (WMM was originally called WME, for Wi-Fi Multimedia Extensions.)

When an 802.11 network moves beyond a light load, stations divide up access to the radio through the congestion window. After a frame transmission completes, there is a gap before the next frame transmission begins. The size of the gap is dependent in part on a random number, with the lowest number "winning" access to the radio medium.

One early strategy for prioritizing voice was to "cheat" on the slot number in the congestion window. Rather than choosing a random number throughout the range as specified by 802.11, transmitters with voice would choose slot number zero, beating out any other frames waiting with data for other applications. Although effective, this strategy required careful timing coordination among multiple voice transmitters to scale beyond a few voice transmitters.

802.11e works by defining four "access classes" that have different priority. From highest priority to lowest priority, they are voice, video, best effort, and background. Each access class also has a defined congestion window, and the higher-priority classes have shorter delays. For example, in 802.11b networks, the voice class has a congestion window of 7 to 15, while the data frames transmitted at "best effort" will have a window size of 31 to 1023. When frames are transmitted out of a radio, the higher-priority voice and video queues are drained before moving on to the best effort queue.

Several different methods exist for mapping data traffic into a particular access class. The most common is to use the 6-bit differentiated services code point (DSCP) in the IP header[2] (formerly called the IP Type of Service (TOS) field). Higher values in the DSCP indicate that a packet should be given higher priority when forwarding. In a

common implementation, DSCP values from 48 to 63 are automatically placed into the WMM voice queue and given high priority access to the radio in both directions. On the Interop Labs Asterisk servers, we have configured Asterisk to send frames with DSCP value 56 to ensure high-priority treatment by the various wireless devices that support WMM. In the past year, support for WMM on SIP telephones has become widespread. The Interop Labs demonstrations have WMM-capable phones from Unex and Spectralink.

Security

Just as with prioritization, the familiar hand-held form factor can obscure some of the security threats posed by transmitting voice over an 802.11 wireless LAN. In effect, mobile phones have a form of security through obscurity because it is much harder to get tools that capture and analyze mobile phone traffic. Once that voice data moves on to a wireless LAN, though, it is readily available to anybody with an 802.11 card and packet capture tools like Ethereal.

Several attack tools are also readily available. Early 802.11 phones supported only manual WEP keys, which could easily be recovered by moderately sophisticated attackers. After recovering the WEP keys, an attacker could easily listen in to telephone calls in real time. If your telephone calls require better security than protection for a few hours (and we'd be surprised if they don't!), then the past year has been very good to you. Better encryption methods such as WPA have become widely available.

WPA's pre-shared key mode provides dynamically generated link layer keys. Note, however, that selecting a strong pre-shared key of adequate length is very important to resist common attacks [3]. Dictionary attack tools based on these attacks are widespread and easily available. These tools are effective only against WPA's pre-shared key mode; WPA Enterprise offers stronger security and is immune to these attacks. Many telephones also support the AES-based encryption algorithm from 802.11i (also referred to as WPA2). At a minimum, network administrators should use WPA or WPA2 pre-shared keys for authentication. More security-conscious networks may need to select telephones that support WPA Enterprise (RADIUS authentication).

Future Work for Voice on 802.11

Voice on 802.11 is still a work in progress. The industry is currently working on solutions to the limited capacity of 802.11. (With the most common codec, 802.11b networks can carry no more than 22 encrypted telephone calls[4].) An emerging component of 802.11e, the traffic specification (TSPEC), enables devices to specify the type and amount of traffic they will be sending once connected. A telephone attempting to join a wireless network can request 80 kbps, and the infrastructure can determine whether it can accommodate that traffic. This process is referred to as "call admission control" because it ensures that stations are only allowed to connect if there is sufficient network capacity for transmission.

The 802.11 power save protocol is quite simple, and was not designed for devices that send data regularly. 802.11e also introduced Automatic Power Save Delivery (APSD) in large part for telephones. In the original 802.11 power save protocol, stations had to periodically power up to check if data had accumulated during the period in which they were powered off. An APSD device will specify how often it will wake up, and the network infrastructure is responsible for keeping track of when a station is available to send and receive data. Because the wireless device can power on its transceiver at known intervals, it can power down more frequently and does not need to power on only to find out that no data is available.

Avoiding disruption to the telephone call requires that the any security context and quality of service parameters be moved rapidly between access points. Today, security information can be stored with 802.11i's pre-master key (PMK) caching. PMK caching is only effective when a telephone has previously associated to an access point, however. 802.11 Task Group R is developing protocols to quickly move both the security context and quality of service parameters between access points very quickly.

References

[1] IEEE 802.11e-2005, Amendment for Medium Access Control (MAC) Quality of Service (QoS) Enhancements
[2] RFC 2474 - Differentiated Services
[3] Moskowitz, Robert. "Weakness in Passphrase Choice in WPA Interface." November 4, 2003. http://wifinetnews.com/archives/002452.html
[4] Gast, Matthew. "How Many Voice Callers Fit on the Head of an Access Point?" December 13, 2005. http://www.oreillynet.com/pub/a/etel/2005/12/13/how-many-voice-callers-fit-on-the-head-of-an-access-point.html