SPEAKEASY QUALITY OF SERVICE: VQ TECHNOLOGY

January 2009

Nationwide
800.556.5829

Seattle Office
206.728.9770

Partner Program
888.886.6116

1201 Western Ave
Seattle, WA 98101

www.speakeasy.net

TABLE OF CONTENTS

OVERVIEW

Speakeasy's business voice service provides an integrated voice and data solution that delivers voice and data over a single broadband connection. With Voice over Internet Protocol (VoIP), voice is digitized and transmitted in Internet Protocol (IP) packets rather than using traditional circuit-switched Time Division Multiplexed (TDM) resources of the public switched telephone network (PSTN).

As a further advance, Speakeasy's Voice Quality (VQ) technology prioritizes voice packets over data. Because Speakeasy has a private, dedicated IP network, it provides dependability and performance, which allows Speakeasy to maintain a voice Quality of Service (QoS) guarantee from the desktop to the PSTN in a manner not possible for most providers.

This paper addresses issues related to Quality of Service and Speakeasy's mechanism for handling each related issue.

MAINTAINING VOICE QUALITY OF SERVICE

As business managers think about cutting costs and gaining efficiencies with Hosted Voice, IT managers think about quality and reliability of service. One of the biggest components in choosing a reliable Hosted Voice vendor is determining how they will maintain Quality of Service (QoS).

Speakeasy's Voice Quality (VQ) technology solves the problems of keeping voice service over the Internet reliable. By prioritizing voice traffic at all the constraint points of WAN links between Speakeasy and customers, Speakeasy is able to provide a high quality and reliable service. This is done by smoothing traffic at the traditional bottlenecks: Toward the customer WAN and through use of router prioritization. Application Layer Gateways (ALG) are used in a similar way to manage the traffic flow from the customer to the network.

Speakeasy addresses the issue of customer WAN circuit congestion with a variety of techniques. Outbound voice traffic from the customer is protected with the use of an Application Layer Gateway (ALG) that is capable of ensuring that voice traffic receives priority through the LAN router. Additionally, Speakeasy edge routers are configured to prioritize voice traffic more favorably than data traffic when crossing the WAN link to our customers, frequently, the largest and most common bottleneck. Finally, the Speakeasy Network directly connects to all the voice network equipment through private interconnects so no voice traffic destined for the PSTN is passed over a third party network or any public peering points. This ensures consistent quality voice service.

THE BUILDING BLOCKS OF VOICE QUALITY (VQ) TECHNOLOGY

Historically in IP networks, there have been two major fields of thought when it comes to maintaining Quality of Service: Differentiatied Services (DiffServ) and Integrated Services (IntServ). Speakeasy's Voice Quality technology is based upon the DiffServ model rather than the IntServ model. The primary difference between the two is how they deal with the problem of resource allocation.

IntServ very closely resembles the way the Public Switched Telephone Network (PSTN) works. An attempt is made to reserve capacity from end-to-end and then to deliver service if that reservation is successful—your call either goes through or you get a fast-busy signal. Speakeasy ruled this out as a viable option for two reasons: signaling and the economics involved.

For IntServ to essentially 'guarantee' that resources will be available from end-to-end, there needs to be a way for the set-up of the call to occur end-to-end. The Resource Reservation Protocol (RSVP) was created to address this problem. In an IP network without end-to-end signaling such as that which RSVP provides, maintaining the 'guarantee' that IntServ requires is difficult. While access networks such as Speakeasy could deploy RSVP, it would also have to extend the protocol all the way through partner networks who provide the IP to PSTN gateways. Today, such cooperative examples of interprovider signaling for purposes of traffic prioritization are nearly nonexistent. Additionally, IntServ essentially recreates a circuit switched network on top of a packet switched network, which removes a lot of the statistical multiplexing benefits of the packet infrastructure. This overlay quickly erodes many of the benefits of a converged network.

The primary advantage of DiffServ is that the signaling protocol is not required. DiffServ does not attempt to solve the problem from end-to-end, but rather on a per-hop basis. Speakeasy considers the DiffServ model to represent a better vehicle to implement converged services across a single broadband connection, as the statistical multiplexing characteristics of the packet network are not eroded

DIFFERENTIATED SERVICES INFRASTRUCTURE

Fundamentally, there are two main concepts in any sort of a Differentiated Services infrastructure:

- Classification
- Scheduling

Classification takes place on the ingress to the domain and ultimately is only an attempt to identify and mark the different forms of traffic so that it can be mapped into the correct per-hop behavior.

Scheduling is where the resource allocation actually takes place; i.e., "I am going to give more buffer capacity to per-hop behavior 'A'." That is, Speakeasy is going to ensure that it will attempt to service the queue for per-hop behavior 'B' at a higher frequency than 'C' or 'D', etc.

HOW VOICE WORKS FOR SOHO (SMALL OFFICE/HOME OFFICE)

Before any sort of per-hop behavior can be applied to a given packet, it must be classified. For Speakeasy OneLink Voice customers, both the core routers and Analog Terminal Adaper (ATA) perform classification. Speakeasy's core routers classify traffic coming from its voice gateways and mark the header in such a way so that the voice per-hop behavior can be queued appropriately through Speakeasy's network. The customer's ATA also performs classification. Any voice traffic that is converted to IP is considered to be in the voice per-hop behavior.

Once classified as a packet carrying voice data, Speakeasy 'protects' the voice signal from regular data flows. The customer's ATA shapes the traffic that is sent across the DSL line before it hits Speakeasy's network. It is important to note that ALL traffic must be passed through the ATA for it to have the desired effect of scheduling the voice traffic before data traffic.

The other area where scheduling comes into play before a packet is sent down a client's DSL line is on the Speakeasy Edge Router itself. Also called an access router, the edge router sits at the periphery of a network, in contrast to a core router that is in the middle of a network. Edge routers and core routers are somewhat relative terms, but may also indicate that different vendors or models of equipment are being used. All are routers, but of different size and capacity. They may have been built from Application Specific Integrated Circuits (ASICs), which are optimized for different tasks. The Edge/Access router is where the customer aggregation takes place.

The scheduler on the Speakeasy Edge Router in the default configuration divides all network traffic equally. This includes buffer capacity as well as how frequently a packet can be transmitted through the interface. For example, if we have an OC3c with 1000 customers bound to it, then each customer will be serviced 1/1000th of the time and have 1/1000th of the buffer pool available for his/her use. In this case, everyone is treated with classic best effort forwarding.

However, with Speakeasy VQ, we have defined four per-hop behaviors, so that the customer now has four differentiated behaviors. These four behaviors are:

- Network Internet routing protocols
- Voice—Packets from Speakeasy gateways
- OnNet—Packets from Speakeasy address space
- OffNet—Packets from the Internet

For every client that is VQ enabled, there are four different per-hop behaviors, each of which is modified from the default case for maintaining different service levels. For the purpose of preserving a quality voice conversation, the

most critical are voice, the valued resource to be protected, and OffNet, the large volume of traffic competing for resources. An earlier reference was made that should 1000 users be attached to a given aggregation circuit, at any given time, each would have access to 1/1000th of the resources. Should one of those users subscribe to Speakeasy's voice service, then the scheduling frequency is no longer completely equal, due to the protective mechanisms that exist to defend the voice traffic:

	Non-voice user	Voice user
Voice	-	strict
Data	1*	1*

The numbers reflect the frequency at which the interface scheduler will look at each client interface to see if there are any packets waiting to transmit. There is an asterisk in the data row to indicate that while the amount of scheduler resources that are allocated to the clients for data starts out the same, there are variables that play into the period in which the scheduler will visit the client interface, as well as the amount of data that is permitted to be sent.

One of the artifacts of handling voice packets with strict priority is that the exact frequency in which the scheduler will visit the client interfaces for data transmission is no longer as predictable. The reason for this is that the real-time nature of voice requires that it receive priority queuing, which then introduces a bit of randomness in the frequency that the data queues are serviced.

An Analogy of Per-Hop Behaviors

Imagine that you are student going on a school field-trip from Seattle to our nation's capitol. Because of the distance involved, flying is the obvious choice rather than driving. At school (a non-DiffServ enabled network), everyone is equal, so when it comes time to board the bus for the airport everybody lines up and files in one by one. Getting off the bus and heading into the main terminal is the same—the bus empties in the same order as it filled, First In, First Out. Of course, now that you are not on school property (a DiffServ enabled network), things are slightly different for some travelers.

Several have frequent flier miles to get an upgrade to business class—a different behavior aggregate that triggers an entirely different per-hop behavior. At the ticket counter, people flying business class are in a MUCH smaller queue (lower latency). Here, the ratio of airline agents to assist business class travelers vs. those flying economy is much more favorable (scheduler weighting). In the economy section, the lines are long and there are many directional ropes to traverse (larger buffer capacity).

The loud speaker announces, *"Last call for boarding flight DCA_1776 at gate J4."*

At this point, it is quite possible that those with business class tickets are already in the gate area, or perhaps even on the plane as they were processed through the lines much more quickly. Others, who are flying with an economy ticket, may miss their flights altogether due to having to wait in queue for so long (packet loss, similar to old world order TTL expiration). In this case, those with economy tickets have no choice other than to look for a later flight (re-transmit).

In Speakeasy's view, queues at the airport are very similar to the DiffServ model, particularly when considering the elasticity of resource allocation. Travelers may have noticed that the helpful airline attendants working with business-class customers will assist those in the economy if they are otherwise unoccupied. The same can happen in DiffServ where resources from one per-hop behavior can be "borrowed" from another if there is a momentary surplus.

BUSINESS VOICE—MORE COMPLEX, SAME QUALITY OF SERVICE

The principles that apply to Small Office/Home Office (SOHO) Voice hold true for business voice; however, another scenario exists. It is possible for the IP data from one phone handset to go to another directly without passing through a voice gateway. As a result, multiple classification possibilities exist. Additionally, there are times where a packet's per-hop behavior will change as it passes through the system.

From a DiffServ perspective, calls from our customers to and from the PSTN are handled in an identical fashion in both the Business and SOHO voice products. At the customer's location, there is still a need for a packet shaping device to assist in protecting the customer's voice traffic. As before, this device is primarily there to prioritize the voice packets that are sent to Speakeasy. However, the device used for business voice also has some common features of IntServ functionality. The device deployed at the edge of a customer's LAN is an Application Layer Gateway (ALG), called an Edgemarc. The ALG is able to perform call admission control as a way to help ensure adequate resources exist end-to-end to sustain calls. If enough resources are available to sustain the voice call end-to-end, then the call goes through. If not, the user gets a fast-busy signal. The ALG does not implement this exactly, as the end-to-end Resource Reservation Protocol (RSVP) needed as the basis for IntServ was never widely adopted.

The ALG does, however, know how many active calls are in progress. Remember that all of the phones register themselves with the ALG, and the ALG is in their forwarding path. The ALG is also configured with a specific number of calls the provisioned network capacity should be able to sustain. With these two pieces of information (number of active calls in progress and number of calls configured), Speakeasy can determine if a call has a chance of going through. In the event that a new call pushes the call capacity over the limit, then the ALG has the ability to disrupt the signaling. In this situation, the caller gets a busy signal of sorts, but the quality of the existing calls that are being forwarded through the ALG is not degraded.

It would be ideal if the network never dropped a packet and the customer only wanted to use their circuit for voice traffic. However, customers usually want to converge their infrastructure to run both voice and data over the same links. The ALG does not have the ability to control non-voice based traffic inbound to the customer's location: It cannot fully prioritize traffic entering the customer's LAN in a fully converged voice and data environment. This is where the Speakeasy edge router comes into play.

In the table below, there are three possible sources/destinations from which the voice traffic, or any traffic for that matter, can flow. As an aid, here are the acronym expansion and definitions used in the table:

- PSTN: Public Switched Telephone Network (voice)
- OnNet: An IP address that is within Speakeasy's address blocks
- OffNet: An IP address that is NOT within Speakeasy's address blocks

Sender	Receiver		
	PSTN	OnNet	OffNet
PSTN	N/A	Voice	N/A
OnNet	OffNet	OnNet	OffNet
OffNet	N/A	OffNet	N/A

Under the receiver section of the table, we see what the actual scheduling behavior of the traffic is. Note: there are three receiver options.

Below is a table which displays Speakeasy's relative scheduler weights:

	Non-Voice User	Voice User
Voice	-	strict
OnNet	-	.66/1*
OffNet	1*	.33/1*

As before, the numbers represent the relative frequency that a particular queue is serviced. For the OnNet and OffNet division that occurs for the Voice user, in the absence of one classification of data traffic, its resources can be borrowed by the other non-voice traffic class.

In the business voice scenarios, the network of classification is:

- Speakeasy Core Routers, which classify traffic coming from the Voice Gateways and mark the header in such a way so that the voice per-hop behavior can be acted upon through Speakeasy's network. [PSTN -> OnNet]
- The ALG classifies voice traffic based upon registration [Voice -> Speakeasy across the backhaul]
- The Speakeasy Edge Router rewrites the IP precedence bits based upon the source and destination of the packet header. If both are within Speakeasy's Classless InterDomain Routing (CIDR) space, then the packet is given the behavioral aggregate for OnNet. Otherwise, it is considered OffNet.

Scheduling looks like:

- The ALG sends voice traffic ahead of data traffic [Voice -> Speakeasy across the backhaul]
- The Speakeasy Edge Router maps all packets based upon their behavior aggregate into the appropriate traffic class and, by extension, queue.

SUMMARY

Built for dynamic businesses, Speakeasy's business voice offers an integrated voice and data solution that lowers operating costs, increases business efficiency, and guarantees world-class reliability while delivering crystal clear call quality. Speakeasy's hosted system eliminates the capital expenditures and long-term leasing requirements of traditional PBX systems, resulting in significant cost savings. And with Speakeasy, voice quality and service reliability are never compromised. This solution truly is the technology of the future.

Voice is a lot more complicated service to maintain than casual web browsing. Before deciding on a managed voice solution for your business, be certain that your provider takes your voice service as seriously as you do, as Speakeasy does.

Author: Brad Volz, Senior Network Engineer, Speakeasy, Inc.