Proceedings of the National Conference; INDIACom-2007
Computing for Nation Development, February 23 – 24, 2007
Bharati Vidyapeeth's Institute of Computer Applications and Management, New Delhi

Optimization Of Voice Over IP

Mrs. Parul Arora
Lecturer, BVICAM, New Dehli-110063
Email: parul@micromaxinfo.com

ABSTRACT
The transfer of voice traffic over packet networks, and especially Voice Over IP(VoIP), is rapidly gaining acceptance. Many industry analysts estimate that the overall VoIP market will become a multi-billion dollar business within three years. While many corporations have been using voice over Frame Relay to save money by utilizing excess Frame Relay capacity, the dominance of IP has shifted most attention from VoFR to VoIP. Initialy, the prime motivation for developing VoIP was reducing the cost of long-distance phone calls. But today many organizations are looking to deploy VoIP because it also enables them to provide a wide range of capabilities. These high-speed backbones take advantage of the convergence of Internet and voice traffic to form a single managed network.
This network convergence also opens the door to novel applications. Interactive shopping are just one example, while streaming audio, electronic white-boarding and CD-quality conference calls in stereo are other exciting applications. But along with the initial excitement, customers are worried over possible degradation in voice quality when voice is carried over these packet networks. Whether these concerns are based on experience with the early Internet telephony applications, or whether they are based on understanding the nature of packet networks, voice quality is a critical parameter in acceptance of VoIP services. As such, it is crucial to understand the factors affecting voice over packet transmission, as well as obtain the tools to measure and optimize them. This paper covers the basic elements of voice over packet networks, the factors affecting voice quality and discusses techniques of optimizing voice quality as well as solving common problems in VoIP networks.

KEYWORDS
VoIP, Gateways, Gatekeepers, Soft Phone, Hard Phone, Latency, Jitter, Silence Suppression, CODECS.

INTRODUCTION
VoIP devices – IP endpoints, call managers, media gateways, gatekeepers,- communicate using a common "language," or protocol. These protocols serve to setup and tear down calls, locate and negotiate resources on the network, register endpoints and transmit connection-related data during the voice session. Call setup protocols allow for control and signaling of IP voice traffic so that a pre-determined path across the network is negotiated in advance of the voice conversation taking place. They also ensure interoperability so that the device of one vendor, say an IP phone, interoperates with the media gateway of another vendor[1]-[5]
Typically, VoIP devices are built to conform to one or more of the following standards:

H.323 The international Telecommunications Union, or ITU, designed H.323 to define how multimedia, such as video, audio and other multicast applications travel over a packet-switched network.

SIP The Internet Engineering Task Force (IETF) defines Session Initiation Protocol in RFC 3261. SIP is a low overhead protocol that works along side other protocols on the network. SIP performs very basic call setup functions such as establishment of user location (i.e. translating from a user's name to their current network address), feature negotiation, call management, and changing features of a session while it is in progress.

MGCP Media Gateway Control Protocol runs in conjunction with other IP protocols such as H.323 or SIP to bridge circuit switched and packet networks. MGCP, running on a media gateway, can lower the total cost of ownership by enabling "dumb" IP endpoints, such as analog phones, to connect to an IP backbone and function with the same feature set as its IP phone counterpart.

RTP/RTCP The Real-time Transport Protocol, RT Control Protocol, provides end-to-end delivery services for data with real-time characteristics, such as interactive audio and voice or simulation data, over multicast or unicast network services. RTP supports data transfer to multiple destinations using multicast distribution if provided by the underlying network.

VOIP Architecture
VoIP services need to be able to connect to traditional circuit-switched voice networks. The basic elements of the H.323 network are shown in the network diagram below where H.323 terminals such as PC-based phones (left side of drawing) connect to existing ISDN, PSTN and wireless devices (right side):

Figure 1: VOIP Architecture

There are two forms of a VOIP call. If you have Microsoft's NetMeeting, you can set up a PC-to-PC call without working with a call server. This is typically how the early users of VoIP made calls. However, the prevalent enterprise VoIP solution requires a call server (the standards community calls this a "gatekeeper") to be part of the network configuration. Although it is called a server, the server does not operate like a traditional server. An e-mail server and a PC are in constant contact for the e-mail operations. In VoIP, the call server (see Figure 1) controls all the services offered, provides control over the call, supports the telephone features, authenticates and authorizes the caller and implements security. The call server is NOT the telephone switch. Once the call server sets up a phone (peer-to-peer) call, the server becomes dormant during the speech transmission unless the phones contact the server to indicate a change in status or the call server wants to change the call configuration, such as indicating there is a call waiting. The server is there to process the signaling, but does not switch the speech. The speech packets are passed directly from phone to phone.

There are two major categories of IP phone implementations: hard phone and soft phone. The hard phone contains all the hardware and software to implement VoIP. It is not a PC, but is specifically designed as a phone. Hard phones can be simple in their functions, but can also have color displays with touch sensitive screens and may even support web browsing. There is no typical hard phone on the market. The soft phone, is a headset connected to a PC with all the telephone features implemented by the sound card and software resident in the PC.

Another piece of hardware, the gateway, is usually part of the VoIP network. Most organizations will have legacy phones, fax-machines, modems, connections to the PSTN, and other devices that originally connected to the organization's telephone switch, called a PBX. When migrating to VoIP, these devices and interfaces will have to be connected to a conversion system that supports the legacy devices and interfaces on one side and connects to the IP network on the other. The legacy devices will be connected to an access/gateway and the PSTN interface connection will be terminated on a trunk gateway[13][14][15].

Understanding the factors affecting voice quality

In the traditional circuit-switched network, each voice channel occupied a unique T1 timeslot with fixed 64 Kbps bandwidth. When traveling over the packet network, voice packets must contend with new phenomena that may affect the overall voice quality as perceived by the end-customer. The premier factors that determine voice quality are choice of **CODEC, latency, jitter and packet loss.**

Audio CODECS

Voice channels occupy 64 Kbps using PCM (pulse code modulation) coding when carried over T1 links. Over the years, compression techniques were developed allowing a reduction in the required bandwidth while preserving voice quality. Such techniques are implemented as CODECS.

Different compression schemes can be compared using four parameters:

Compressed voice rate – the CODEC compresses voice from 64 Kbps down to a certain bit rate. Some network designs have a big preference for low-bit-rate CODECS. Most CODECS can accommodate different target compression rates such as 8, 6.4 and even 5.3 Kbps.

Complexity – the higher the complexity of implementing the CODEC, the more CPU resources are required.

Voice quality – compressing voice in some CODECS results in very good voice quality, while others cause a significant degradation.

Digitizing delay – Each algorithm requires that different amounts of speech be buffered prior to the compression. This delay adds to the overall end-to-end delay network with excessive end-to-end delay, often causes people to revert to a half-duplex conversation ("How are you today? over…") instead of the normal full-duplex phone call.

Understanding latency

In contrast to broadcast-type media transmission (e.g., RealAudio), a two-way phone conversation is quite sensitive to latency, Most callers notice round-trip delays when they exceed 250mSec, so the one-way latency budget would typically be 150mSec. G.114 recommendation as the maximum desired one way latency to achieve high-quality voice. Beyond that round-trip latency, callers start feeling uneasy holding a two-way conversation and usually end up talking over each other. At 500mSec round-trip delays and beyond, phone calls are impractical, where you can almost tell a joke and have the other guy laugh after you've left the room. For reference, the typical delay when speaking through a geo-stationary satellite is 150-500mSec.

Data networks were not affected by delay. An additional delay of 200mSec on an email or web page goes mostly unnoticed. Yet when sharing the same network, voice callers will notice this delay. When considering the one-way delay of voice traffic, one must take into account the delay added by the different segments and processes in the network, as shown in the following diagram:

Some components in the delay budget need to be broken into fixed and variable delay. For example, for the backbone transmission there is a fixed transmission delay which is dictated by the distance, plus a variable delay which is the result of changing network conditions.
The most important components of this latency are:

- **Backbone (network) latency**. This is the delay incurred when traversing the VoIP backbone. In general, to minimize this delay, try to minimize the router hops that are traversed between end-points.
- **CODEC latency**. Each compression algorithm has certain built-in delay. For example, G.723 adds a fixed 30mSec delay. When this additional gateway overhead is added in, it is possible to end up paying 32-35mSec for passing through the gateway. Choosing different CODECS may reduce the latency, but reduce quality or result in more bandwidth being used.
- **Jitter buffer depth**. To compensate for the fluctuating network conditions, many vendors implement a jitter buffer in their voice gateways. This is a packet buffer that holds incoming packets for a specified amount of time before forwarding them to decompression. This has the effect of smoothing the packet flow, increasing the resiliency of the CODEC to packet loss, delayed packets and other transmission effects. However, the downside of the jitter buffer is that it can add significant delay.

Understanding jitter

While network latency effects how much time a voice packet spends in the network, jitter controls the regularity in which voice packets arrive. Typical voice sources generate voice packets at a constant rate. The matching voice decompression algorithm also expects incoming voice packets to arrive at a constant rate. However, the packet-by-packet delay inflicted by the network may be different for each packet.
The result packets that are sent in equal spacing from the left gateway arrive with irregular spacing at the right gateway, as shown in the following diagram:

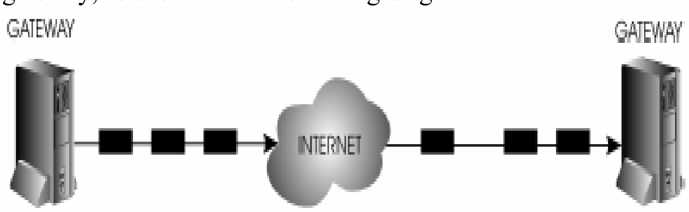

Figure 2: VOIP Architecture

Since the receiving decompression algorithm requires fixed spacing between the packets, the typical solution is to implement a jitter buffer within the gateway. The jitter buffer deliberately delays incoming packets in order to present them to the decompression algorithm at fixed spacing. The jitter buffer will also fix any out-of-order errors by looking at the sequence number in the RTP frames. The operation of the jitter buffer is analogous to a doctor's office where patients that have appointments at fixed intervals do not arrive exactly on time and are deliberately delayed in the waiting room so they can be presented to the doctor at fixed intervals. This makes the doctor happy because as soon as he is done with a patient, another one comes in, but this is at the expense of keeping patients waiting. Similarly, while the voice decompression engine receives packets directly on time, the individual packets are delayed further in transit, increasing the overall latency.

Packet loss

Packet loss is a normal phenomenon on packet networks. Loss can be caused by many different reasons: overloaded links, excessive collisions on a LAN, physical media errors and others. Transport layers such as TCP account for loss and allow packet recovery under reasonable loss conditions. Audio CODEC s also take into account the possibility of packet loss, especially since RTP data is transferred over the unreliable UDP layer. The typical CODEC performs one of several functions that make an occasional packet loss unnoticeable to the user. For example, a CODEC may choose to use the packet received just before the lost packet instead of the lost one, or perform more sophisticated interpolation to eliminate any clicks or interruptions in the audio stream.
However, packet loss starts to be a real problem when the percentage of the lost packets exceeds a certain threshold (roughly 5% of the packets), or when packet losses are grouped together in large packet bursts. In those situations, even the best CODECS will be unable to hide the packet loss from the user, resulting in degraded voice quality. Thus, it is important to know both the percentage of lost packets, as well as whether these losses are grouped into packet bursts.

Tunable factors in VoIP Network

Having discussed the parameters that affect voice quality especially jitter and loss, it is a good time to elaborate on some of network conditions affect these parameters[6][7].

Network load

A very important factor affecting voice quality is the total network load. When the network load is high and especially for networks with statistical access such as jitter and frame loss typically increase. For example, when using Ethernet, higher load leads to more collisions. Even if the collided frames are eventually sent over the network, they were not sent when intended to, resulting in excess jitter. Beyond a certain level of collisions, significant frame loss occurs.

While good network design takes into account the network load, it is not always under your control. However, even in congested networks it is sometimes possible to employ packet prioritization schemes, based on port numbers or on the IP precedence field. These methods, typically built into routers and switches, allow giving timing-sensitive frames such as voice priority over data frames. There is often no perceived degradation in the quality of data service, but voice quality significantly improves. Another alternative is to use bandwidth reservation protocols such as RSVP (resource reservation protocol) to ensure that the desired class of service is available to the specific stream.

Jitter buffer settings

The jitter buffer can be configured in most VoIP gear. The jitter buffer size must strike a delicate balance between delay and quality. If the jitter buffer is too small, network perturbations such as loss and jitter will cause audible effects in the received voice. If the jitter buffer is too large, voice quality will be fine, but the two-way conversation might turn into a half-duplex one.

One can decide on a jitter buffer policy that specifies that a certain percentage of packets should fit in the jitter buffer, say 95%. Since the utilization of the jitter buffer depends on the arrival times of the packets, it is useful to look at the jitter buffer problem in view of following terms

- Sequence number. This designates the RTP sequence number of the incoming packet.
- Absolute time – the absolute arrival time of the packet.
- Delta time – the inter-arrival time (absolute time of each packet – absolute time of previous packet).
- Delay-Expected Inter-Arrival time – suppose the expected inter-arrival time is 20 mSec (the inter-emission time), this term shows how much the inter-arrival time deviated from the expected inter-arrival time. If all packets arrive exactly on schedule, this term will be always 0.

Packet size

Packet size selection is also about balance. Larger packet sizes significantly reduce the overall bandwidth but add to the packetizing delay as the sender needs to wait more time to fill up the payload. Overhead in VOIP communications is quite high. Consider a scenario where you are compressing down to 8 Kbps and sending frames every 20mSec. This results is voice payloads of 20 bytes for each packet. However, to transfer these voice payloads over RTP, the following must be added: an Ethernet header of 14 bytes, IP header of 20 bytes, UDP header of 8 bytes and an additional 12 bytes for RTP. This is a whopping total of 54 bytes overhead to transmit a 20-byte payload.

In some cases, such an overhead is fine. In others, there are two solutions to the problem:

- Increase packet size. By deciding to send packets every 40mSec, it is possible to increase the payload efficiency. Before the inter-arrival time is increased, it should be verified that the delay budget can support this.
- Employ header compression. Header compression is popular with some vendor's equipment, especially on slow links such as PPP, FR or ISDN. This is commonly called CRTP or Compressed RTP. It compresses the header down to a few bytes on a hop-by-hop basis. This can be done because the "logial channel" is determined by the FRDLCI and thus some header information is redundant.

Silence suppression[5]

Silence suppression takes advantage of prolonged periods of silence in conversations to reduce the number of packets. In a normal interactive conversation, each speaker typically listens for about half the time, so it is not necessary to transmit packets carrying the speaker's silence. Many vendors take advantage of this to reduce the bandwidth and number of packets on a link.

Proposed Complementary Solutions

Organizations are challenged today to optimize the voice services all through the world. We need comprehensive solutions to manage VOIP implementations and support IT professionals to enhance their networked applications in a variety of ways

1. One way to optimize VoIP is to track the overall link traffic volume , as well as the volume consumed by individual data applications, voice media traffic (RTP, RTCP) and voice signaling traffic (H.323, SIP, IPSI, SCCP, and MGCP). Application that are generating this traffic should reveal whole link utilization. This information aids IT Managers in troubleshooting and traffic engineering and makes it easier to understand which applications are impacting the network.

2. Create power alerts to notify system managers when a threshold has been exceeded, and forward the evidence of the event that triggered the alert. Some applications generate large amount of traffic and over occupy the link. Such applications, hosts and conversations should send an alert when the threshold crosses. Along with these alerts system manager also needs information about the application, client, and server that were experiencing the condition when the threshold was crossed

3. Configure Alerts to report Time over Threshold thus avoiding frivolous or nuisance alarms, and to provide the most relevant information, the alarms can be configure to report only after the threshold has been exceeded continuously within the specified interval.

4. Identify specific individual phone users on a per link basis to determine how it is being utilized, who is

using it and how much they are using for highlighting protocol usage details and valid/failed calls.

5. Analyse details related to the most recent conversations for individual phone users' activity including: call setup (i.e. sender and receiver codec), quality (i.e. sender and receiver packet loss - by volume and percentage, sender / receiver jitter, and sender / receiver DSCP for resolving mis configuration problems

CONCLUSIONS

Voice over IP services offer lucrative advantages to customers and service providers alike. However, as with any new technology, it brings its own sets of network design and optimization issues. By understanding the important parameters, and acquiring the proper methodology, you can reap the benefits of voice over packet services. The good news is that after experiencing these problems, system administrators (or their replacements) will know which areas need to be fixed/adjusted/expanded to accommodate this voice traffic and these infrastructure changes may carry the enterprise forward until another change causes a new set of problems to surface

REFRENCES

[1] IEEE Std. 802.11, IEEE Standard for Wireless LAN Medium Access Control (MAC) and Physical Layer Specification, November 1997.

[2] G. Anastasi and L. Lenzini, "QoS provided by the IEEE 802.11wireless LAN to advanced data applications: a simulation analysis," *Wireless Networks*, (6):99—108, 2000.

[3] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications*, 18(3):535—547, March 2000.

[4] B. P. Crow, I. Widjaja, J. G. Kim and P. Sakai, "Investigation of the IEEE 802.11 medium access control (MAC) sublayer functions," *Proeedings of INFOCOM'97*, volume 34, pages 126—133, April 1997.

[5] M. Veeraraghavan, N. Cocker and T. Moors, "Support of voice services in IEEE 802.11 wireless LANs," *Proceedings of INFOCOM'01,* 2001.

[6] J.-Y. Yeh and C. Chen, "Support of multimedia services with the IEEE 802.11 MAC protocol," *Proceedings of ICC'02*, New York, April 2002.

[7] P. Brady, "A model for generating on-off speech patterns in two-way conversation," *Bell Syst. Tech. Journal,* vol. 48, no. 7, pp.2245—2272, Sept. 1969.

[8] C. E. May and T. J. Zebo, "A summary of speech statistics measured during the TASI-E Rego Park-Ojus field trial", submitted for publication.

[9] S. Garg and M. Kappes. "On the throughput of 802.11b networks for VoIP", *submitted for publication*.

[10] D. Chen, S. Garg, M. Kappes and K. S. Trivedi, "Supporting VoIP traffic in IEEE 802.11 WLAN with enhanced medium access control (MAC) for quality of service," *submitted for publication*

[11] Application Notes for NetScout *nGenius* with Avaya Communication Manager and Extreme Networks Summit48 - Issue 1.0, Solution & Interoperability Test Lab Application Notes. http://www.avaya.com/masterusa/enus/resource/assets/applicationnotes/ns-acm-xtrm.pdf

[12] Application Notes for NetScout *nGenius* with Avaya Communication Manager Running on the Avaya S8300 Media Server and Avaya G700 Media Gateway - Issue 1.0, Solution & Interoperability Test Lab

[13] Application Notes. http://www.avaya.com/gcm/master-usa/enus/resource/assets/applicationnotes/nss8300p333.pdf

[14] Network Protocols Handbook, Javvin Technologies Inc,handbook@javvin.com

[15]http://www.bitpipe.com/detail/RES/ 1166456162_688.html&asrc=EM_KAR_20061220&src=KA_RES_20061220Sachs