Internet Communication : Voice Over Internet Protocol (IP)

Vaibhav Panwar*

ABSTRACT

Communication via packet and data networks such as IP, ATM; Frame Relay has become a preferred strategy for both corporate and public networks. Experts predict that data traffic will soon exceed telephone traffic if it already hasn't. As a result of this there has been considerable interest in transmitting voice over data networks, as opposed to the more traditional data over voice networks. This process is called "Voice over IP". This paper seeks to identify the challenges involved in integrating voice over data networks such as IP. It also proposes solutions to common problems involved in transmitting voice over data networks such as delay, echo, jitter, interoperability, density, scalability, reliability, lost packet compensation as well as bandwidth reservation

1. INTRODUCTION

Companies and Organizations around the world want to reduce rising communication costs. The consolidation of separate voice and data networks offers an opportunity for significant reduction in communication costs. Accordingly, the challenge of integrating voice and data networks is becoming a rising priority for network managers. Since data traffic is growing much faster than telephone traffic, a need has been identified to transfer voice over data networks, as opposed to the transmission of data over voice networks. This give rise to Voice over IP (VoIP). VoIP has become especially attractive given the low-cost, flat rate pricing of the public Internet. Many components will have to be designed to accommodate voice over data networks, such as the access gateways that link the data and the telephony networks among others. Applications that offer Voice over IP services will have to include a sending voice over networks.

Comprehensive technology set that reduces the impairments caused by data that were not designed to handle it. An important factor for the network designers is the problem of Quality of Service (QoS). This is because IP is a best effort service and therefore provides no guarantees on delivery and data integrity. Voice processing will need to handle greater and variable delays, jitter, and cancel echoes that will be introduced from the telephony side. It will also have to include an appropriate algorithm to mask the gaps caused by dropped packets due to congestion on the network. A protocol needs to be implemented which guarantees bandwidth for the duration of a session and also better compression technologies need to be put in place. An understanding of how to handle call set up translation for different types of networks, connections, and internetworking is essential for competent handling of every call. A Voice over IP application meets the challenges of combining legacy voice networks by allowing both voice and signaling information to be transported over IP.

*Lecturer School of IT, Institute of Management Studies Dehradun

Typically there are three styles of Voice over IP calls. It can be either from Computer to Computer, or it could be from a Computer to phone or from phone to phone. The key components here are the gateways, which adapt traditional telephony to the Internet. A call goes over the local public switched telephone network (PSTN) to the nearest gateway server, which digitizes the analog voice signal using pulse code modulation, compresses it into IP packets, and moves it onto the Internet for transport to a gateway at the receiving end.

VoIP is the technology for transmission of voice traffic over packet switched networks or the Internet. It is hence also known as Internet telephony, IP telephony, and packetized voice. From a business perspective the trends show the integration of voice, video and data in multiple applications. The amount of bandwidth consumption is 8:1 in favor of packet-based networks. Packet-based networks such as IP help reduce the rising communication costs. Also, a single network is much easier to manage. Most of all from a consumer perspective, making calls through packet networks is much lower than circuit switched networks.

The technology is a real shift from using the traditional circuit switched network, which has dedicated 64 Kbps of bandwidth reserved end to end. The current circuit-switched system is connection oriented in which the user picks up the phone. The dial tone indicates resources are available in central office. After dialing the number, the signaling system (SS7) allocates resources through network. One usually has to wait for their party to answer. A fast busy signal indicates congestion in the network. VoIP is an emerging technology defined under the H.323 set of protocols. It specifies components, protocols and procedures that provide multimedia communication services such as real-time audio, video and data over packet networks. Packet based networks include IP-based LANs, enterprise networks, MANs, WANs. The protocols specified by H.323 are - audio CODECs; video CODECs; H.225 registration, admission, and status (RAS) H.225 call signaling, H.245 control signaling, real-time transfer protocol (RTP), real-time control protocol (RTCP).

2. VOICE OVER IP DESIGN AND DEVELOPMENT CHALLENGES

The goal of VoIP developers is to add telephone-calling capabilities (both voice transfer and signaling) to IP based networks and interconnect these to the public telephone network and to private voice networks, in such a way as to maintain current voice standards and preserve the features everyone expects from the telephone. VoIP development needs to take place in five specific areas:

1. Voice quality should be comparable to what is available using the PSTN, even over networks having variable levels of QoS.
2. The underlying IP network must meet strict performance requirements and criteria including minimizing call refusals, network latency, packet loss, and disconnects. This is required even when there is heavy congestion in the network or when resources have to be shared among multiple users.
3. Call control (the actual signaling) should be done transparent to the user in such a way that they should be unaware of what technology is actually implementing the service.

PSTN/ VoIP service internetworking (and equipment interoperability) involves gateways between the voice and data network environments.

System administration or System management, security, addressing, and accounting must be provided, preferably consolidated with the PSTN operation support systems.

3. QUALITY OF SERVICE ISSUES IN IP NETWORKS

The advantages of reduced cost and bandwidth savings of carrying voice over data networks are associated with some Quality of Service (QoS) issues unique to packet networks. Delivering quality voice signals from one point to another cannot be considered successful unless the quality of the delivered signal satisfies the recipient. Providing a level of quality that at least equals the PSTN (this is usually referred to as "toll quality voice") is viewed as a basic requirement. Although QoS usually refers to the fidelity of the transmitted voice and facsimile document it can also be applied to network availability, telephone feature availability, and scalability. Many factors have been identified that play a big role in determining the quality of service. They are as follows:

1. **Delay** Two probems that result from high end-to-end delay in a voice network is echo and talk over lap. Echo is caused by signal reflections of the speaker's voice from the far end telephone equipment back into the speaker's ear. Echo becomes a problem when the round trip delay exceeds 50 milliseconds. Since echo is perceived as a significant quality problem, Voice over IP systems have to address the need for echo control and implement means for echo cancellation. Talker's overlap is the problem of one caller stepping on the other talker's speech. This becomes significant if the one-way delay becomes greater than 250 milliseconds. Delay can be subdivided into two sub-components. They can be fixed delay components as well as variable delay components. Fixed delay components include propagation, serialization, and processing. The variable delay components include the queuing delay, dejitter buffers as well as variable packet sizes. Figure 1 shows the different components of delay and where it takes place in the network.

The following are sources of delay in an end-to-end Voice over IP call.

1. **Accumulation delay**, which is also called algorithmic delay at times, is caused by the need to collect a frame of voice samples to be processed by the voice coder. This depends on the type of voice coder used and varies from a single sample time, which is .125 microseconds, to many milliseconds.

2. **Processing delay** is caused by the actual process of encoding and collecting the encoded samples into a packet for transmission over the IP network. The encoding delay is a function of both the processor execution time and the type of algorithm used. To reduce this delay what is often done is to collect multiple voice coder frames in a single packet.

2. **Network delay**, is caused by the physical medium and protocols used to transmit the voice and data, and by the buffers used to remove packet jitter on the receiver side. This delay is a function of the capacity of the links in the network as well as the processing that happens as the packets pass through the network. The jitter buffers add delay used to remove the packet delay variation that each packet experiences as it transits the packet network. This delay can

be a significant part of the overall delay as packet delay variations can be as high as 70msec to 100msec in IP networks.

3. **Jitter (Delay Variability)** Jitter is the variation in inter-packet arrival time as introduced by the variable transmission delay over the network. Removing jitter requires collecting packets and holding them long enough to allow the slowest packets to arrive in time to be played in the correct sequence, which in turn causes additional delay. The conflicting goals of minimizing delay and removing jitter has led to the development of various schemes to adapt the jitter buffer size to match the time varying requirements of network jitter removal.

4. **Packet Loss** IP networks cannot provide a guarantee that packets will be delivered at all, much less in order. Packets will be dropped under peak loads and during periods of network congestion. But due to the time sensitivity of voice transmissions, however the normal TCP- based retransmission schemes are not suitable. Approaches that compensate for packet loss have to be developed to overcome this problem.

5. **Bandwidth Availability** Bandwidth is the portion of the network that is available to an application to transfer information on the network. The level of reliability and sound quality that is acceptable among users has not yet been reached and this is primarily because of bandwidth limitations and this also leads to packet loss. In voice communications, packet loss shows up in the form of gaps or periods of silence in the conversation, thus leading to a "clipped speech" effect that is unsatisfactory for most users and unacceptable in business communications.

6. **Proposed solutions for problems associated in sending Voice over IP** Maintenance of acceptable voice quality levels despite inevitable variations in network performance is achieved using a variety of techniques. These techniques and solutions to problems that have been detailed above with regard to transmission of voice over IP are as follows.

One of the main problems of a very big end-to-end delay is the problem of echoes. This happens whenever the round-trip delay exceeds 50 milliseconds. Echo in a telephone network is caused by signal reflections generated by the hybrid circuit that converts between a 4-wire circuit (a separate transmit and receive pair) and a 2-wire circuit (a single transmit and receive pair). Echo in a telephone network is acceptable because the round-trip delays through the network are smaller than 50 milliseconds. Echoes are a problem in Voice over IP as the round-trip delays are almost always greater than 50 milliseconds. A way to improve speech quality is to implement some kind of echo cancellation mechanism. The ITU standard G.165 defines performance requirements for echo cancellers. The way the echo cancellers work is that when the echoes are generated from the telephone network toward the packet network, the echo canceller compares the voice data received from the IP network to the voice data that is being transmitted to the IP network. The echo from the telephone network is removed by a digital filter on the transmit path to the IP network.

As mentioned above the task of solving the problem of jitter in transmitting voice over IP has two conflicting goals. They are of minimizing delay as well as removing jitter. This has led to the development of various schemes to adapt the jitter buffer size to match the time varying requirements of network jitter removal. These schemes have the explicit goal of minimizing the size and delay of the jitter buffer while at the same time preventing buffer underflow caused by jitter. One approach, which is used in IP networks to adapt the jitter buffer size, is as follows. This approach involves counting the number of packets that arrive late and create a ratio of these packets to the number of packets that are successfully processed. This ration is then in turn used to adjust the jitter buffer to target a predetermined allowable late packet ratio. This approach works best with networks with highly variable packet and inter-arrival intervals such as IP.

Lost packets are a big problem in networks. This is an even more important problem in IP networks as IP does not guarantee service, and as a result of this they will usually exhibit a much higher incidence of lost packets. In current IP networks all voice frames are treated like data. Under peak loads and congestion, voice frames will be dropped equally with data frames. Data frames are not time sensitive like voice frames and there is no point in retransmission of lost frames as in voice transmission, if a packet is late, it is as good as not reaching the receiver at all. If the lost packets are left untreated, the listener hears annoying pops and clicks. Some schemes called lost packet compensation schemes used by voice over IP to overcome the problem of lost packets are as under.

1. Interpolate for lost speech packets by replaying the last packet received during the interval when the last packet was supposed to be played out. This works well when the incidence of lost frames is infrequent. It does not work very well for bursty loss of packets.
2. Another way is to send redundant information at the expense of bandwidth utilization. The basic approach replicates and sends the n'th packet of voice information along with the (n+1) th packet. This method has the advantage of being able to exactly correct for the lost packet. However, this approach uses more bandwidth and creates greater delay.
3. An alternative approach is to develop an algorithm in the digital signal processor that detects missing packets, and then replays the last successfully received packet at a decreased volume in order to fill the gaps.
4. Another problem is that of Out of Order Packets. When an out of order condition is detected in the network, the missing packet is replaced by its predecessor, as was lost. When the late packet finally arrives, it is discarded.

Quality levels that are tolerated by the users are also achieved using such techniques such as compression, silence compression. The way silence compression works is that whenever it detects a gap in speech, it suppresses the transfer of things like pauses, breaths, and other periods of silence. This can amount to 50%-60% of the time of a call, resulting in considerable bandwidth conservation. Since the lack of packets is interpreted as complete silence at the output, another function is needed at the receiving end to add "comfort noise" to the output.

5. SOFTWARE SUPPORT TO ENABLE VOICE OVER IP

Transmitting voice over an IP network is just one of the many applications that use IP, with software to support the application and the interface to the network. Two major types of information must be handled in order to interface telephony equipment to a packet network- voice and signaling.

The software functionality required for voice-to-packet conversion in a VoIP terminal or gateways are:

Voice packet software – typically runs on a Digital Signal Processor (DSP), prepares voice samples for transmission over the packet network. Its components perform tone detection and generation, echo cancellation, voice compression, voice activity detection jitter removal, resampling and voice packetization.

Telephony signaling gateway software – this software interacts with the telephony equipment, translating signaling into state changes used by the network protocol module to set up connections. These state changes are on-hook, off-hook, trunk seizure, etc. This software supports E&M (wink, delay, and immediate), loop or ground start foreign exchange station and foreign exchange office, integrated services digital network, basic and primary rate interface and QSIG.

Packet processing software – this software processes voice and signaling packets, adding the appropriate transport headers prior to submitting the packets to the IP network. Signaling information is converted from telephony protocols to the packet signaling protocols.

6. TELEPHONY SIGNALING GATEWAY SOFTWARE MODULES

The components of the telephony signaling software are as follows.

- Telephony interface unit software: periodically monitors the signaling interfaces of the module and provides basic debouncing and rotary digit collection for the interface.
- Signaling protocol unit: contains state machines implementing the various telephony signaling protocols such as E&M
- Network control unit: maps telephony signaling information into a format compatible with the packet voice session establishment signaling protocol.
- Address translation unit: maps the E.164 dial address to an address that can be used by the packet network.

DSP interface driver: relays control information between the host microprocessor and DSP's.

DSP downline loader: responsible for downline load of the DSP's at start-up, configuration update, or mode changes

7. NETWORK SUPPORT FOR QUALITY OF SERVICE (QOS)

A key requirement for successful VoIP deployment is the availability of an underlying IP based network that is capable of supporting real-time telephone and facsimile. As was noted above, voice quality is affected by delay, jitter, unreliable packet delivery, and bandwidth limitations – all of which are characteristics of the basic IP network service.

A basic requirement is that most of today's data network equipment, such as routers, LAN switches, ATM switches, network interface cards, PBX's, will have to be configured to support voice transmissions and traffic. Furthermore, equipment specific to Voice over IP will have to either be integrated into these devices or will have to work compatibly with them. Many different techniques and protocols are used to improve the quality of service in networks. Some of these techniques and protocols that provide better network support in terms of reliability, availability, etc. are discussed below and in he next section.

Providing a controlled networking environment in which capacity can be preplanned and adequate performance can be assumed (at least for a majority of the time). This would generally be the case with a private IP network like an Intranet that is owned and operated by a single owner.

Using management tools to configure the network nodes, monitor performance, and manage capacity and flow on a dynamic basis. Most internetworking devices such as routers and switches include a variety of mechanisms that can be used in supporting voice. Traffic can be prioritized by protocol, location, and application type, thereby allowing real-time traffic to be given precedence over non-critical traffic. Queuing mechanisms may also be manipulated to minimize delays for real-time data flows. Recent research has shown that tag switching and flow switching can also improve overall performance and reduce delays.

Adding control protocols and mechanisms that help avoid or alleviate the problems inherent in IP networks. Protocols such as Real-time Transport Protocol (RTP), Real-time Transport Control Protocol (RTCP), and Resource Reservation Protocol (RSVP) are being implemented to provide greater assurances of controlled Quality of Service within the network. These protocols and the way they provide solutions to the problem of QoS in IP networks are discussed in the next section (Section 8). Other mechanisms such as admission controls and traffic shaping may also be used to avoid performance deterioration and overloading the network.

Some other Networking tools to provide Quality of Service include

1. Congestion Management (Weighted Fair Queuing)
2. QoS signaling (IP Precedence and RSVP)
3. Packe Residency (MLPPP/MTU Size Reduction)
4. RTP header compression
5. Generic Traffic Shaping
6. Weighted Random Early Detection (WRED)

8. CONCLUSION

Data has traditionally been transmitted over the voice networks. The Internet has created the opportunity to reverse this trend and now send voice over data networks such as IP, with the integration of video and other multimedia applications close behind. The Internet and its underlying TCP/IP protocol suite have become the driving force for new technologies, with the unique challenges of real-time voice being the latest in the series of developments. Telephony over the Internet cannot make compromises in voice quality, reliability, scalability, and manageability. It must also work seamlessly with telephone systems all over the world. Therefore all of today's network devices need to be voice enabled. As mentioned above, future extensions will have to include conference bridging, voice/data synchronization, and text to speech conversion and voice responsive systems. In conclusion, the market for VoIP is established and is beginning its rapid growth phase and corporations must look to capitalize on this new wave of voice over packet networks, like IP, in order to cut communication costs and improve efficiency due to the unification characteristic of IP, and to keep pace with a fast growing global economy.

REFERENCES

[1] RFC 2205, RSVP: Resource Reservation Protocol. R. Braden, L. Zhang, S. Berson, S. Herzog, S. Jamin September 1997. http://www.ietf.org/rfc/rfc2205.txt

[2] RFC 1889, RTP: Real-time Transport Protocol. H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson. January 1996. http://www.ietf.org/rfc/rfc1889.txt

[3] Technology Guide Series. Voice over IP http://www.techguide.com/

[4] Voice over IP. J.Q. Maguire Jr. http://www.it.kth.se/edu/gru/Internet/Lectures/VOIP.html

[5] Voice over IP: Issues and Challenges. Raj Jain http://www.cis.ohio-state.edu/~jain/cis788-99/h_8voip.htm

[6] Voice over IP: Network Design. Jonathon Davidson, Cisco Systems http://www.cisco.com/warp/public/625/ccie/presentation/921_jonathan_rev4/index.htm

[7] Telogy Networks. Voice over IP White Paper http://www.telogy.com/our_products/golden_gateway/VOPwhite.html

[8] Trillium Digital Systems H.323 Tutorial http://www.webproforum.com/h323/tp01.html

[9] Internet Telephony Tutorial. Siemens Telecom Networks. http://www.webproforum.com/siemens2/full.html http://www.webproforum.com/acrobat/int_tele.pdf

[10] Voice over Packet. Telogy Networks. http://www.webproforum.com/acrobat/voice_packet.pdf

[11] Voice Transmission over Asynchronous Transfer Mode. TERAbridge http://www.webproforum.com/acrobat/vtoa.pdf

[12] Convergence of Voice, Fax, and Video over IP. Network World Seminars and Diginet. http://www.nwfusion.com/whitepapers/voip_wp.html

[13] RFC 2543, SIP: Session Initiation Protocol. M. Handley, H. Schulzrinne, E. Schooler, J. Rosenberg. March 1999. http://www.ietf.org/rfc/rfc2543.txt

[14] H. Schulzrinne, J. Rosenberg, "A comparison of SIP and H.323 for Internet Telephony".