Voice over IP
(IP Telephony)

March 28, 2001

Done by: Daniel Rodriguez
To: Dr. Farmer

What is VoIP?

VoIP stands for **V**oice **o**ver **I**nternet **P**rotocol. VoIP is a new technology that allows users to send human voice over IP networks. Human voice gets transformed into IP packets and these IP packets travel over many distant computer networks like the Internet. VoIP can use accelerating hardware to achieve this purpose and it is also used in the PC environment. VoIP is also known as *IP telephony* because that is how this technology is mostly used; it resembles the typical telephone call, in the fact that people can use it to make a call to different users.

How does it work?

Just like in Data networks, where we are able to route packets of data as long as they are digitized, so does VoIP uses this technique to send voice over the Internet. All we have to do is digitialized our voice into a packet and send it across the network just like a data packet. This architecture requires the need for a sender component that will be in charge of digitizing the voice into a digital signal and a receiver that will convert it back to its analog form. Then it uses the same techniques as data networks to transmit the packets across the Internet. VoIP works like that, digitizing voice in data packets, sending them and reconverting them in voice at destination.

Digital format can be better controlled, because we can compress it, route it, convert it to a new better format. Another advantage of the digital signal is that it is more tolerant to noise than the analog. (see GSM vs TACS). TCP/IP networks are made of IP packets containing a header (to control communication) and a payload to transport the data: VoIP use it to go across the network and reach its destination.

Converts Analog to Digital

Converts Digital to Analog

Voice (source) - ----- ADC ------ - - - Internet - - ---- DAC ---- - - Voice (destination)

What are the advantages using VoIP rather than PSTN?

When you are using PSTN (Public Switch Telephone Network), you typically pay for time spent connected to this network, the more time you stay at the phone the more you will pay. In addition you could not talk with more that one person at a time, which is easily done with the VoIP mechanism, where you can talk all the time with every person you want (all that is needed is that the other people also connect to Internet at the same time). If you are still not persuaded you can consider that at the same time that you are talking to the other person you can also exchange data, images and videos.

Then, why isn't everybody using it?

Unfortunately we have to report that there are still some problems with the integration of VoIP architecture and the Internet. The exchange of voice needs to be done on real-time, meaning that the information traveling back and forth between the source and destination has to be done very quickly and in a reliable way, so users won't be interrupted by delays on the signal and be forced to wait for a response.

A fast and reliable method of communication is not one of the Internet's major characteristics. The Internet is a heterogeneous architecture that is composed of many routers (machines that route packets), about 20-30 of these routers can have a very high round trip time (RTT), so we need to modify something to get it properly working. There are suggested solutions to this dilemma, for example using totally different protocol that sends packets in order, or the most promising solution would be to increase the bandwidth of the network.

Technical info about VoIP

To setup a VoIP communication we need:

1. First the ADC to convert analog voice to digital signals (bits).

2. Now the bits have to be compressed in a good format for transmission: there are a number of protocols we will review after.
3. Here we have to insert our voice packets in data packets using a real-time protocol (typically RTP over UDP over IP)
4. We need a signaling protocol to call users: ITU-T H323 does that.
5. At the receiving end we have to disassemble packets, extract the data, then convert them to analog voice signals and send them trough a sound card (or phone)
6. All of that must be done in a real time fashion, because we cannot wait for too long for an answer.

Analog to Digital Conversion

The analog to digital conversion is mostly taken care by the hardware, typically by a card integrated with ADC. Today every sound card allows you convert with 16 bit a band of 22050 Hz (for sampling it you need a frequency of 44100 Hz for Nyquist Principle) and also to obtain a throughput of 2 bytes * 44100 (samples per second) = 88200 Bytes/s, 176.4 kBytes/s for stereo stream.

Compression Algorithms

Now that we have digital data we may convert it to a standard format that could be quickly transmitted, such as PCM (Pulse Code Modulation) Standard ITU-T G.711. Voice bandwidth is 4 kHz, so sampling bandwidth has to be 8 kHz (for Nyquist). We represent each sample with an 8 bit (having 256 possible values), so throughput is 8000 Hz *8 bit = 64 kbit/s, as a typical digital phone line. In real application mu-law (North America) and a-law (Europe) variants are used which code analog signal with a logarithmic scale using 12 or 13 bits instead of 8 bits (see Standard ITU-T G.711).

RTTP Real Time Transport Protocol

Now that we have the raw data, we can encapsulate it into the TCP/IP stack. Following this structure:

VoIP data packets
RTP
UDP
IP
I, II layers

VoIP data packets live in RTP (Real-Time Transport Protocol) packets, which are inside UDP-IP packets. First, VoIP does not use TCP because it is too slow for real time applications, so instead UDP (datagram) is used. In UDP we cannot order packets to arrive in time (which is a must in VoIP) because UDP does not use a virtual circuit between the source3 and the destination, each packet is independent from the others (datagram concept), so we have to introduce a new protocol, such as RTP.

RSVP

There are also other protocols used in VoIP, like RSVP, that can manage Quality of Service (QoS). RSVP is a signaling protocol that requests a certain amount of bandwidth and latency in every network hop that supports it.

Quality of Service (QoS)

We said many times that VoIP applications require a real-time data streaming because we expect an interactive data voice exchange. Unfortunately, TCP/IP cannot guarantee this kind of exchange, it just makes a "best effort" to deliver the packets. So we need to introduce tricks and policies that could manage the packet flow in every router we cross. The following are what is used to increase QoS:

1. A TOS field in the IP protocol to describe type of service: high values indicate low urgency while more and more low values bring us more and more real-time urgency.
2. Queuing packets methods:
    1. FIFO (First in First Out), it is an easier method that allows packets to arrive in order.
    2. WFQ (Weighted Fair Queuing), consisting in a fair passing of packets (for example, FTP cannot consume all available bandwidth), depending on kind of data flow, typically one packet for UDP and one for TCP in a fair fashion.
    3. CQ (Custom Queuing), users can decide priority.

    4. PQ (Priority Queuing), there is a number (typically 4) of queues with a priority level each one: first, packets in the first queue are sent, then (when first queue is empty) starts sending from the second one and so on.
3. Shaping capability, that allows limiting the source to a fixed bandwidth in:
    1. Download.
    2. Upload.
4. Congestion Avoidance, like RED (Random Early Detection).

H323 Signaling Protocol

H323 protocol is used, for example, by Microsoft NetMeeting to make VoIP calls. The h323 protocol allows not only VoIP but also video and data communications.

Hardware requirement

To create a VoIP system you need the following hardware:

1. PC 386 or more.
2. Sound card, full duplex capable.
3. A network card or connection to the Internet or other kind of interface to allow communication between 2 PCs

All of these have to be present on both ends to simulate standard communication. The tools above are the minimal requirement for a VoIP connection: next we will see that we should (and in Internet we must) use more hardware to do the same in a real situation. Sound card has to be full duplex unless we could not hear anything while speaking.

Software requirement

We can choose what O.S. to use:

1.Win9x
2.Linux

Under Win9x we have Microsoft Netmeeting, Internet Phone, DialPad or others or Internet Switchboard (from Quicknet web site) for Quicknet cards. Also you can use free software you download from OpenH323.

Under Linux we have free software from OpenH323 web site: simph323 or ohphone that can also work with Quicknet accelerating hardware.

Bandwidth consideration

From all we said before we noticed that we still have not solved problems about bandwidth or how to create a real time streaming of data. We know we could not find a solution unless we enable a right real-time manager protocol in each router we cross, so what do we can do? First we use a very high rate compression algorithm, such as: LPC10 (which only consumes a 2.5 kbps bandwidth, about 313 bytes/s). Then we start to classify our packets, in TOS field, with the most high priority level, so every router will transfer the packet urgently. Important: all of that is not sufficient to guarantee our conversation would always be ok, but without a great infrastructure managing bandwidth reservation and so on, it is not possible to do it, TCP/IP is not a real time protocol.

Summary

VoIP has a very bright future in the field of telecommunication; it has a lot room to expand and a lot of applications to contribute to the world. As network connections become faster and more cost effective the use of VoIP will be more than welcome in our daily lives.

Useful links

Voxilla	http://www.voxilla.org/
Linux Telephony	http://www.linuxtelephony.org/
International Communication Union	http://www.itu.int/home/index.html
Quicknet Web site	http://www.quicknet.net/
Open H323 web site	http://www.openh323.org/
Speak Freely	http://www.speakfreely.org/
Cisco Systems	http://www.cisco.com/