

VoIP Bandwidth Calculation

VoIP Bandwidth Calculation

Executive Summary

Calculating how much bandwidth a Voice over IP call occupies can feel a bit like trying to answer the question; "How elastic is a piece of string?" However, armed with a basic understanding of the parts that make up the whole, the question becomes easier to understand. This White Paper examines the process that turns voice into Voice over IP.

Overview

The amount of bandwidth required to carry voice over an IP network is dependent upon a number of factors. Among the most important are:

- Codec (coder/decoder) and sample period
- IP header
- Transmission medium
- Silence suppression

The codec determines the actual amount of bandwidth that the voice data will occupy. It also determines the rate at which the voice is sampled. The IP/UDP/RTP header can generally be thought of as a fixed overhead of 40 octets per packet, though on point-to-point links RTP header compression can reduce this to 2 to 4 octets (RFC 2508). The transmission medium, such as Ethernet, will add its own headers, checksums and spacers to the packet. Finally, some codecs employ silence suppression, which can reduce the required bandwidth by as much as 50 percent.

The Codec

The conversion of the analogue waveform to a digital form is carried out by a codec. The codec samples the waveform at regular intervals and generates a value for each sample. These samples are typically taken 8,000 times a second. These individual values are accumulated for a fixed period to create a frame of data. A sample period of 20 ms is common. Some codecs use longer sample periods, such as 30 ms employed by G.723.1. Others use shorter periods, such as 10 ms employed by G.729a.

The important characteristics of the codec are:

- The number of bits produced per second
- The sample period - this defines how often the samples are transmitted

Together, these give us the size of the frame. For example, take a G.711 codec sampling at 20 ms. This generates 50 frames of data per second. G.711 transmits 64,000 bits per second so each frame will contain 64,000 / 50 = 1,280 bits or 160 octets.

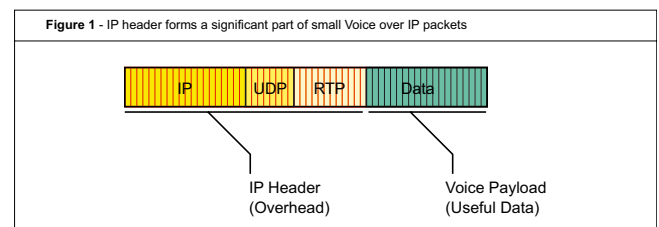
Frames and Packets

Many IP phones simply place one frame of data in each packet. However, some place more than one frame in each packet. For example, the G.729a codec works with a 10 ms sample period and produces a very small frame (10 bytes). It is more efficient to place two frames in each packet. This decreases the packet transmission overhead without increasing the latency excessively.

Latency and Packet Overhead

The choice of the number of frames per packet is a trade-off between two characteristics: latency and packet overhead.

Long sample periods produce high latency, which can affect the perceived quality of the call. Long delays make interactive conversations awkward, with the two parties often talking over each other. Based on this fact alone, the shorter the sample period, the better the perceived quality of the call. However, there is a price to pay. The shorter the sample period, the smaller the frames and the more significant the packet headers become. For the smallest packets, well over half of the bandwidth used is taken up by the packet headers; clearly an undesirable case. See figure 1 below:



The IP Header

The term 'IP header' is used to refer to the combined IP, UDP and RTP information placed in the packet. The payload generated by the codec is wrapped in successive layers of information in order to deliver it to its destination. These layers are:

- IP – Internet Protocol
- UDP – User Datagram Protocol
- RTP – Real-time Transport Protocol

RTP is the first, or innermost, layer added. This is 12 octets. RTP allows the samples to be reconstructed in the correct order and provides a mechanism for measuring delay and jitter.

UDP adds 8 octets, and routes the data to the correct destination port. It is a connectionless protocol and does not provide any sequence information or guarantee of delivery.

IP adds 20 octets, and is responsible for delivering the data to the destination host. It is connectionless and does not guarantee delivery or that packets will arrive in the same order they were sent.

In total, the IP/UDP/RTP headers add a fixed 40 octets to the payload. With a sample period of 20 ms, the IP headers will generate an additional fixed 16 kbps to whatever codec is being used.

The payload for the G.711 codec and 20 ms sample period calculated above is 160 octets, the IP header adds 40 octets. This means 200 octets, or 1,600 bits sent 50 times a second – result 80,000 bits per second. This is the bandwidth needed to transport the Voice over IP only, it does not take into account the physical transmission medium.

There are other factors, which can reduce the overhead incurred by the IP headers, such as compressed RTP (cRTP). This can be implemented on point-to-point links and reduces the IP header from 40 to just 2 or 4 octets. Though this is not that common today, its use will become more widespread with it being implemented with 3G mobile networks.

The Transmission Medium

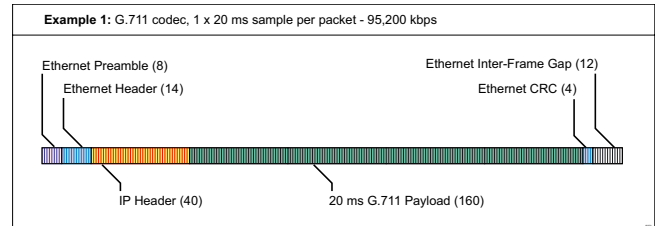
In order to travel through the IP network, the IP packet is wrapped in another layer by the physical transmission medium. Most Voice over IP transmissions will probably start their journey over Ethernet, and parts of the core transmission network are also likely to be Ethernet.

Ethernet has a minimum payload size of 46 octets. Carrying IP packets with a fixed IP header of 40 means that the codec data must be at least 6 octets – typically not a problem. The Ethernet packet starts with an 8 octet preamble followed by a header made up of 14 octets defining the source and destination MAC addresses and the length. The payload is followed by a 4 octet CRC. Finally, the packets must be separated by a minimum 12 octet gap. The result is an additional Ethernet overhead of 38 octets.

Ethernet adds a further 38 octets to our 200 octets of G.711 codec frame and IP header. Sent 50 times a second – result 95,200 bits per second, see example 1 below. This is the bandwidth needed to transmit Voice over IP over Ethernet.

Transmission of IP over other mediums will result in different overhead calculations.

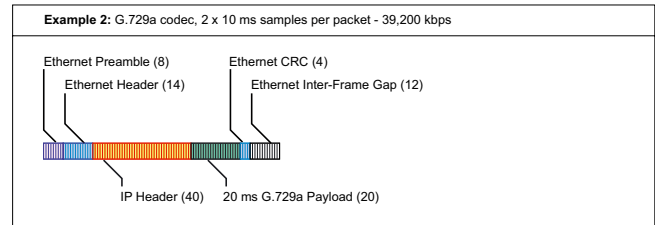
Voice over IP over Ethernet, Example 1: G.711



- Codec G.711 – 64 kbps, 20 ms sample period
- 1 frames per packet (20 ms)
- Standard IP headers
- Ethernet transmission medium

One packet is sent every 20 ms, 50 packets per second. Payload is $64,000 \div 50 = 1,280$ bits (160 octets). Fixed IP overhead 40 octets, fixed Ethernet overhead 38 octets. Total size 238 octets. Bandwidth required is $(160 + 40 + 38) \times 50 \times 8 = 95,200$ kbps.

Voice over IP over Ethernet, Example 2: G.729a



- Codec G.729a – 8 kbps, 10 ms sample period
- 2 frames per packet (20 ms)
- Standard IP headers
- Ethernet transmission medium

One packet is sent every 20 ms, 50 packets per second. Payload is $8,000 \div 50 = 160$ bits (20 octets). Fixed IP overhead 40 octets, fixed Ethernet overhead 38 octets. Total size 98 octets. Bandwidth required is $(20 + 40 + 38) \times 50 \times 8 = 39,200$ kbps.

Silence Suppression

Certain codecs support silence suppression. Voice Activity Detection (VAD) suppresses the transmission of data during silence periods. As only one person normally speaks at a time, this can reduce the demand for bandwidth by as much as 50 percent. The receiving codec will normally generate comfort noise during the silence periods.



Additional Codec Data

Codec	Bandwidth	Sample period	Frame size	Frames/packet	Ethernet Bandwidth
G.711 (PCM)	64 kbps	20 ms	160	1	95.2 kbps
G.723.1A (ACELP)	5.3 kbps	30 ms	20	1	26.1 kbps
G.723.1A (MP-MLQ)	6.4 kbps	30 ms	24	1	27.2 kbps
G.726 (ADPCM)	32 kbps	20 ms	80	1	63.2 kbps
G.728 (LD-CELP)	16 kbps	2.5 ms	5	4	78.4 kbps
G.729A (CS-CELP)	8 kbps	10 ms	10	2	39.2 kbps
AMR (ACELP)	4.75 kbps	20 ms	12	1	36.0 kbps
AMR (ACELP)	7.4 kbps	20 ms	19	1	38.8 kbps
AMR (ACELP)	12.2 kbps	20 ms	31	1	43.6 kbps
AMR-WB/G.722.2 (ACELP)	6.6 kbps	20 ms	17	1	38.0 kbps

Summary

Although there are many factors that influence the amount of bandwidth required to transmit a voice call over an IP network, by approaching the problem one element at a time the final calculation becomes relatively simple. Other factors may influence the actual bandwidth used, such as RTP header compression, silence suppression and other techniques still under development.

Glossary

ACELP	Algebraic CELP (codec)
ADPCM	Adaptive Differential Pulse Code Modulation (codec)
AMR	Adaptive Multi-Rate (GSM2 codec)
AMR-WB	Adaptive Multi-Rate Wideband (3GPP codec G.722.2)
CELP	Code Excited Linear Predictive (codec)
Codec	Coder/Decoder
CRC	Cyclic Redundancy Code
cRTP	Compressed RTP (RFC 2508)
CS-CELP	Conjugate-Structure CELP
IP	Internet Protocol
LD-CELP	Low Delay CELP (codec)
MP-MLQ	Multi-Pulse Maximum Likelihood Quantization
Octet	A sequence of 8 bits
PCM	Pulse Code Modulation (codec)
RFC	Request For Comment
RTP	Real-time Transport Protocol
SB-ADPCM	Sub-Band ADPCM
UDP	User Datagram Protocol
VAD	Voice Activity Detection

