VoIP Bandwidth Calculation

A digital telephone system converts an analog voice signal into a stream of bits expressed in K bits per second (where K is used to mean one thousand). For standard PCM digital encoding this stream of bits is 64 K bits per second. This is 64 K bps in each direction (transmit and receive).

To improve transmission efficiency, this bit stream is compressed using a standard compression algorithm such as G.729. The result is still a bit stream but with fewer bits per second. For example G.729 will reduce the 64 K bits per second to a bit stream of 8 K bits per second.

This bit stream is then divided into chunks that can be placed in packets for transmission over a data network. Some textbooks refer to these chunks of voice data as Voice Samples but I prefer to call them Voice Frames.

This reduced bit stream is examined repeatedly in fixed time intervals. This examination time is called the Voice Frame Interval. This is the time used to collect the bits for one Voice Frame. The Voice Frame Interval is expressed in milliseconds (ms). A millisecond is one thousandth of a second.

If X is the number of bits per second in the reduced bit stream and Y is the time of the Voice Frame Interval in milliseconds, then the number of bits in a Voice Frame is:

\[ X \text{ Kbps} \times Y \text{ ms} \]

The dimensions “thousand” and “second” cancel out and the resulting dimension is bits.

\[ XY \text{ bits of Voice data} \]

These bits need to be converted to Bytes so they can be combined with the data added by the Internet protocols handling the data transmission. (It needs to be converted so we can add apples to apples so to speak). Divide the equation above by 8 to convert bits to Bytes (because there are 8 bits per Byte).

\[ XY / 8 \text{ Bytes of Voice data} \]

As the voice data is processed by the various Internet protocols, each adds more data to the Voice data. The data added by each is expressed in Bytes. The User Datagram Protocol (UDP) adds 8 Bytes. The Real Time Protocol (RTP) adds 12 Bytes. The Internet Protocol (IP) adds 20 bytes of data. This is a total of 40 Bytes of Overhead data added by these transport protocols.

The size of the packet is the sum of the Voice data and the Overhead data.

\[ \text{Packet size} = \text{Overhead Data} + \text{Voice Data (Bytes per Packet)} \]
\[ = (40 + XY / 8) \text{ (Bytes per Packet)} \]

Bandwidth is expressed in bits per second, so this needs to be converted back to bits.

\[ \text{Packet size} = 8 \times (40 + XY / 8) \text{ (bits per Packet)} \]
\[ = (320 + XY) \text{ (bits per Packet)} \]

To calculate the Bandwidth (bits sent per second) we multiply the bits per packet by the number of packets that can be sent in one second. The number of packets that can be sent in a second is found by
dividing one second by the time it takes to collect the Voice data (the Voice Frame Interval). It is possible to place more than one Voice Frame in each packet. If there are “n” Voice Frames in each packet, then the time to collect the data for a packet is,

\[
\text{Time to collect Voice data for each packet} = nY
\]

and the number of packets sent in one second is,

\[
\text{Packets sent per Second} = \frac{1}{nY} \text{ (K packets per second)}
\]

The Bandwidth is found by multiplying the (bits per packet) by (packets sent per second)

\[
\text{Voice Bandwidth} = \frac{(320 + XY)}{nY} \text{ (K bits per second)}
\]

This is the Bandwidth formula for a VoIP Voice signal before it is placed on the network. This formula does not take into consideration the overhead added by Layer 2 protocols such as Ethernet and Frame Relay or the overhead added by the VPN protocol used for security. Using H as the Overhead of all protocols, a more general formula for voice bandwidth is,

\[
\text{Voice Bandwidth} = \frac{(8H + XY)}{nY} \text{ (K bits per second)}
\]

This is the formula for the voice part of a standard VoIP connection. This does not include the call control portion of RTP, which is called RTCP. This is usually calculated by multiplying the Voice Bandwidth by 0.0526 to get the VoIP Bandwidth. This is applied to Voice Bandwidth before adding the Layer 2 and security overhead.

\[
\text{RTCP Bandwidth} = 0.0526 \frac{(320 + XY)}{nY} \text{ (K bits per second)}
\]

The Bandwidth per Voice Channel is the sum of the Voice Bandwidth plus the RTCP Bandwidth

\[
\text{Bandwidth per Voice Channel} = \left[ \frac{0.0526 (320 + XY)}{nY} \right] + \left[ \frac{(8H + XY)}{nY} \right]
\]

\[
= \left( 16.832 + 8H + 1.0526 XY \right) / nY \text{ (K bits per second)}
\]

The telephone control signals will add more bandwidth. Each Aspire telephone uses 16 K bps of control signals. Since there is one telephone associated with each voice channel, the 16 K bps control signal is added to the Voice Channel Bandwidth. The Total bandwidth is then,

\[
\text{Total Bandwidth per channel} = 16 + \left( 16.832 + 8H + 1.0526 XY \right) / nY \text{ (K bits per second)}
\]

Each telephone call has 2 voice channels, one to send and one to receive. In reality, the protocols (and the resulting bandwidth) used by endpoint are not required to be the same. For calculation purposes, it is easier to assume that the same protocol is used by both endpoints in a telephone call. So, we multiply the Bandwidth per channel by 2 to calculate Bandwidth per telephone call.

\[
\text{Bandwidth per telephone call} = 2 \left( 16 + \left( 16.832 + 8H + 1.0526 XY \right) / nY \right)
\]

Multiply this by the total number of VoIP ports per system to calculate the bandwidth requirements for the system.