VoIP Bandwidth Considerations - design decisions

When calculating the bandwidth requirements for a VoIP implementation the two main protocols are: a signalling protocol such as SIP, H.323, SCCP, IAX or MGCP for call set up, and control; and a protocol to carry the voice audio stream. The Real-Time Transport Protocol (RTP) is normally used to carry the voice audio stream. An IETF standard RTP will work with any signalling protocol.

Factors effecting the quality of a VoIP call

- Reliability (needs to be less than 1% loss)
- Delay (<250ms to prevent 'speech collisions', <150ms ideal) ITU recommendation calls for <150ms domestic, <200ms internationally and <400ms via satellite for the one way delay
- Delay Variation = Jitter (<10ms)
- Bandwidth (prioritized queuing & call admission control) LFI (link fragmentation and interleave) required on slow data/voice links

A VoIP source can be configured to sample audio and produce a voice packet every 10, 20, 30 or 40ms, depending on the vendor's implementation. The packet of digitized speech can be uncompressed, compressed and/or encrypted, changing the overhead and the final size of the packet carried over the network.

A user will detect voice break-up and poor quality if there is too much delay, loss or distortion. The choice of packet size is made based upon delay through the network, bandwidth available and packet loss within the network. A short packet creation delay, enables more network delay to be tolerated. Shorter packets cause less of a problem if the packet is lost, but shorter packets require more bandwidth to to the increased overhead. Long packets that contain more voice audio reduce the bandwidth requirements but produce a longer delay and are therefore more noticeable if they are lost somewhere in the network. Most vendors opt for a middle of the road packet size of 20 or 30ms.

The CODEC (COder/DECoder) is the element used to digitise the voice audio and converts it back into an analogue audio stream. The CODEC is the analogue-to-digital-to-analogue converter in the IP Phone, IP PBX, softphone, gateway etc...The different CODECs perform at differing bit sample rates and also perform compression and decompression.

Most vendors support one or more of the following ITU standards :

- **G.711** often the default standard for IP PBX, IP phones and ATA vendors, as well as for the digital PSTN and trunks. This standard digitizes voice into 64 Kbps before encapsulation with no compression
- **G.729** is supported by many vendors for compressed voice operating at 8 Kbps, 8 to 1 compression. With quality just below that of G.711, although hard to detect the difference. There is a licensing fee associated with every G.729 active call stream in use. Ideal CODEC for WAN and low bandwidth circuits. To be avoided on links going through tandem encoding such as to GSM networks.
- **G.723.1** was once the recommended compression standard. It operates at 6.3 Kbps and 5.3 Kbps. Although this standard further reduces bandwidth consumption, voice is noticeably poorer than with G.729, so it is not very popular for VoIP.
- **G.722** operates at 64 Kbps, but offers high-fidelity speech. Whereas the three previously described standards deliver an analogue sound range of 3.4 kHz, G.722 delivers 7 kHz. This version of digitized speech has been announced by several vendors and will become common in the future.

The quality of a voice call is defined by the Mean Opinion Score (MOS) 0 to 5. A score of 4.4 to 4.5 is considered to be toll quality. Voice compression affects the MOS as the sampling rate drops so does the MOS. An MOS below 3.5 will usually produce complaints from the users. Cell phone calls average about 3.8 to 4.0 for the MOS.

Standard	Speed	MOS	Sampling delay per phone
G.711	64 Kbps	4.4	0.75 ms
G.711/PCM	64Kbps	4.1	0.125ms
G.729	8 Kbps	4.2	10 ms
G.729A	8Kbps	3.7	10ms
GSM	14Kbps	3.5	20ms
G.728	16Kbps	3.8	0.625
G.723.1	6.3 Kbps 5.3 Kbps	4.0 3.5	30 ms

Packet Format for the digitised voice audio stream It is useful to remember the OSI layered construction of the data through the network. Each layer in the OSI adding the headers, addressing and port information to

The RTP header field contains version, identifiers, M bit (marker for playout burst), time stamp and sequence numbers (for packet loss detection) for the content of the following voice sample. A payload type field within RTP defines the compression technique used in the sample and is used by SDP to identify the CODEC. In practice the 'standard payload type codes' are not used because of variations in how the vendors have implemented their CODECs and a mapping of RTP payload type to configured CODEC is found in most VoIP IP PBX systems.

The RTP packet format for VoIP is shown below with example CODEC sample

IP header	UDP header	RTP header	CODEC sample Xms	CODEC sample Xms
20 Bytes	8 Bytes	12 Bytes	10 Bytes	10 Bytes

The RTP can be encapsulated using Ethernet, frame relay, ATM, PPP etc… with the digitized CODEC sample, RTP, UDP and IP headers remaining the same through the network. Each of the above packets will contain a snap-shot of the voice audio with a packet rate of 50 packets per second for 20ms and 33.3 packets per second for 30ms voice samples.

Depending upon the layer-2 protocol, the header and trailer overheads can make or break the VoIP service consuming 20% to 80% of the bandwidth. As shown in the RTP packet layout, there can be at least 40 bytes of overhead before encapsulation into yet another 18 bytes of overhead for Ethernet frames or 6 to 8 bytes for Frame Relay and/or PPP. As the size of the voice field gets larger with longer packets, the percentage of overhead decreases -- therefore the needed bandwidth decreases. In other words, bigger packets are more efficient than smaller packets.

Header compression Cisco's header compression technique is now the 'industry standard' called RTP header compression. This technique actually compresses the RTP, UDP and IP headers and significantly reduces the RTP, UDP and IP overhead from 40 bytes to between 4 and 6 bytes. The bandwidth consumption for compressed voice packets can be reduced by nearly 60%. This technique has less value for large uncompressed voice packets.

Calculating bandwidth consumption for VoIP As seen above, the bandwidth required for VoIP will depend on several factors: the compression technology CODEC, packet overhead, network protocol used and whether silence suppression is used. Also several factors effect the bandwidth, delay, jitter and losses within a network: bandwidth available, queuing strategy, QoS, reliability and resilience.

Not all of these factors are under the control of the enterprise. In the real world, a compromise often has to be made between cost and bandwidth. It is not always possible or cost effective to just add bandwidth to implement VoIP, QoS is not always available on the lowest cost Teleco circuits and all the Telco options of DDI, PRI, BRI, SIP trunks, leased circuits and DSL with the required supporting hardware have to be evaluated. The enterprise will often have to accept that during times of busy voice activity the data network response will be slightly slower, if there are not enough funds to provide the luxury of ample bandwidth.

For example The choice of CODEC G.711 @ 64kbps (8000 Bytes per sec) or G.729 @ 8kbps (1000 Bytes per sec) offers MOS of 4.1 or 3.94 which most people can not differentiate, but G.729 will suffer if tandem encoding has to be used to reach a GSM network and drops the MOS to 3.15. In noisy environments a higher bit rate is also desirable as the noise levels are higher with the more complex CODECs. In short;

- Choose a higher bit rate CODEC for noisy environments
- Do not use low bit rate CODECS for tandem encoding routes
- Use single CODEC frame per packet when using low bit rate CODECs on poor networks with suboptimal QoS
- Use low bit rate CODECs were possible on WAN routes and on conferencing scenarios

There are cost implications also…. a G.729 CODEC will often cost an additional $10 licensing fee per concurrent one-way audio stream. Some of the lowest cost hardware does not support higher bit rate CODECs.

That said, the bandwidth requirements for a one-way voice stream using G.729, with RTP compression and a 20ms sample is approximately 32kbps. Using G.729, with RTP compression and a 40ms sample is approximately 22kbps. G.711 with RTP compression and 20ms sample is 88kbps. G.711 with RTP compression and 40ms sample is 76kbps……… The choice of packet size, voice compression and header compression make it difficult to determine the required VoIP bandwidth.

- Bandwidth requirements reduce with compression, G.711 vs. G.729.
- Bandwidth requirements reduce when longer packets are used, thereby reducing overhead, but packets more prone to audio drop during loss
- Even though the voice compression is an 8 to 1 ratio, the bandwidth reduction is about 3 or 4 to 1. The overhead negates some of the voice compression bandwidth savings.
- Compressing the RTP, UDP and IP headers (cRTP) is most valuable when the packet also carries compressed voice.
- If silence suppression/voice activity detection is used, the bandwidth consumption may drop 50% -- to 8 Kbps total per VoIP call.

Bandwidth allocation for call setup and RTCP is shown in RFC3556 to require about 5% of the RTP session bandwidth. For G.711 call using 88kbps the RTCP would require 4kbps, for G729 call using 22kbps the RTCP would require about 1kbps.

Comparison of circuits available form the Telco …. There are a bewildering amount of circuits and managed options available from the Telcos, here is just a few. Most enterprises allocate about 33% of their connected telephones the facilities for concurrent calling…. i.e. a company with 99 extensions would be expected to have the capacity to have 33 concurrent voice calls at any time.

E1 European lease line circuit of 2.048Mbps used to carry 32 x 64kbps

T1 US leased line circuit of 1.544 Mbps used to carry 24 x 64kbps channels

Frame Relay provides access circuits to Telcos frame and cell networks or can be used as private enterprise networks

ATM as above. Various Telco offerings provide a leased access circuit and many options of bandwidth and QoS

ISDN (integrated services digital network) sold as ISDN2 BRI (basic rate interface) (2 x 64kbps + 16kbps control) or ISDN30 PRI (primary rate interface) (30 x 64kbps + 1 64kbps control + 1 64kbps alarm channel) E1 circuit switched digital technology or US PRI sold as 23 channels of 64kbps + 1 64kbps control) T1 circuit switched

aDSL (asymmetric digital subscriber line) sold over standard POTS lines using a 276kbps split in the full duplex uplink/downlink providing upto 8Mbps downstream and upto 1Mbps upstream. Limited by distance from nearest DSLAM exchange where the speed falls dramatically as the distance exceeds 1.8 miles. Sold in differing levels of contention ratio.

aDSL2 as above with improved standards providing upto 12Mbps downstream and 1Mbps upstream within 1.8 miles of the DSLAM.

aDSL2+ as above with improvement to provide upto 24Mbps downstream and 1Mbps upstream within 1.8 miles of the DSLAM

aDSL2+M as above with improvement to provide 3.5Mbps upstream

sDSL (symmetrical digital subscriber line) makes use of the entire broadband on the POTS cable (i.e. can not co-exist with analogue telephony) to provide upto 2.3Mbps in both directions within 1.8 miles of the DSLAM aDSL CONTENTION RATIO can be seen as 50:1, 20:1, 5:1 and 1:1 offering from various Telcos. Simply translates to the maximum amount of connections competing for your allocated bandwidth at the busiest time. The following figures show worst case downstream during busy periods IF full contention is reached ;

1Mbps aDSL 50:1 20.48kbps

4Mbps aDSL 50:1 81.8kbps

2Mbps aDSL 20:1 102.4kbps

4Mbps aDSL 20:1 204.9kbps

2Mbps aDSL 5:1 409.8kbps

4Mbps aDSL 5:1 819.6kbps

4Mbps aDSL 1:1 4Mbps

With some aDSL circuits only achieving 256kbps in the upstream direction and on a contention ratio of 50:1 it can be seen that the upstream minimum bit rate could be as low as 5kbps in theory and therefore unsuitable for VoIP. 20:1 may just support a single voice stream during the worst contention period. In practice a Telco will not allow the bandwidth to drop as low as the maximum contention and will actively monitor and upgrade links to keep the bandwidth at a 'reasonable level'. However, due to this unknown bandwidth and lack of QoS, an aDSL link is normally limited to only support one or two simultaneous voice calls.

When using xDSL for VoIP it is important to calculate the worst case as the minimum upstream bandwidth that will be available for the contention ratio supplied. QoS is available from some Telcos but only using their 1:1 contention offerings. Generally QoS is not supported across the Telco via aDSL.

QoS is a requirement for any bottle-neck within the network. The bottle-necks are usually connections to the WAN. Remember it is important to preserve the VoIP requirements of little or no packet loss, <200ms delay, <10ms jitter and enough bandwidth for the call. That will not be achieved if you WAN is fed with a router that has no QoS and allows an FTP transfer to send several 1500 Byte packets onto the WAN during a VoIP call causing the VoIP audio stream to wait more than 10ms before it is sent….. the result would be broken audio. Therefore it is import to configure the switches and routers within the LAN to mark and classify the packets to enable the router to queue and prioritize the VoIP packets and ensure minimum delays to the voice packets.

Recommended steps are to classify the enterprise network traffic into real-time, mission-critical and best-effort for the video/voice, email and import applications followed by the web and data transfer applications etc… In this way all traffic can be classified and queues configured on the routers and switches to ensure the optimum bandwidth is used by the services selected. In many cases, an enterprise will have to conduct a baseline and audit exercise to determine the traffic levels and nature of their bandwidth utilization before the designs for VoIP can be begun.

For more information see www.kccvoip.com

Keith Campbell MIET

KCcommunications
VoIP # 8829900284400
UK # +44 1207 668118
US # +1 425 906 4810
www.kccommunications.com
www.kccvoip.com www.kccvoip.net