Bandwidth Optimization.
By Sergio J. Castro, CETyS University

From the user perspective In developed countries bandwidth now a days is a commodity, it really is; users hardly worry about bandwidth limitations anymore, bandwidth is an afterthought. In the rest of the world it is not the same case, especially in places where the digital gap is enormous; bandwidth is still a precious resource. Service providers everywhere still manage their bandwidth carefully, regardless of their location as new interactive multimedia services demand more and more bandwidth to run properly. This article focuses on describing the bandwidth concept and uses one media type, voice, to illustrate how bandwidth is utilized to run in a cost effective optima way and still delivering a good quality service.

For practical purposes bandwidth is defined as the capacity in Hertz (Hz) necessary for offering different multimedia services such as telephony, video and data transfer. However, it is interesting to define bandwidth scientifically as a measurement unit for transmitting our telecommunication services.

The definition of bandwidth originated from analog waves that transmit signals; frequency is the variation in the amplitude of a signal wave. Bandwidth is the width of a frequency range and it is expressed mathematically by:

$BW = f^1 - f^2$

Where BW = bandwidth; f^1 = the highest frequency; f^2 = the lowest frequency, so it clearly is the difference between the highest and lowest frequencies[1].

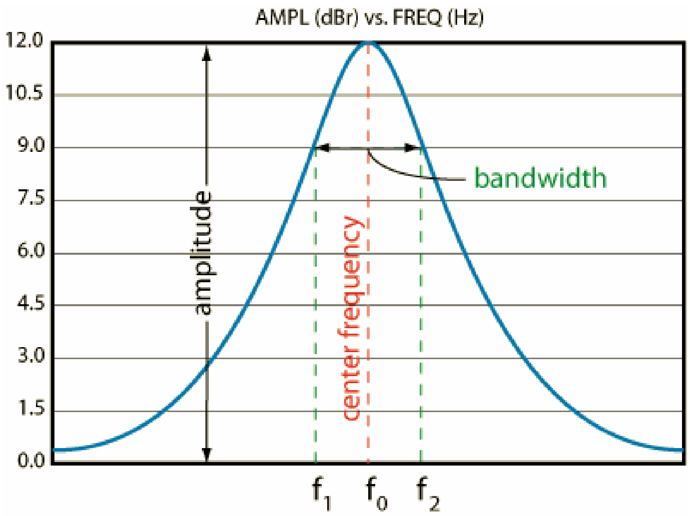

Figure 1. Bandwidth.

Vowels and consonants provide a perfect example of bandwidth. Vowels occupy higher frequencies and consonants occupy lower frequencies. The human ear can detect frequencies in the range of 40 to 18,000 Hz. Early telecommunications transmission devices did not have the capacity to transmit in the higher frequencies, fortunately this was not a major problem as the human brain has the capacity to reconstruct sound characteristics and the intention of a message[2].

Initially, for voice transmission, a 4000Hz channel was defined as the optimal one. This meant that the channel could transmit frequencies ranging from 0 to 4KHz. The first telephone companies connected only a handful of subscribers as each subscriber received a pair of copper cables at 4 KHz. It was a great disadvantage having to provision a cable to each of the subscriber's home[1].

There's a direct relationship between bandwidth and transmission speed (bits per second), the more bandwidth allocated to a channel the more transmission capacity there will be. Transmission capacity (R) is expressed by $R = 2 f^1$. In copper cables it is possible to transmit beyond 100,000 Hz, so a method for transmitting beyond the 4,000 Hz per user was needed and here's where FDM plays the important role of Multiplexing. FDM stands for Frequency Division Multiplexing. With FDM up to 30 voice channels can be transmitted simultaneously; the first channel is transmitted in that range from 0 to 4 KHz as previously explained, the second channel is transformed to transmit in the range between 4 KHz to 8 KHz, the third channel transmits from 8 KHz to 12 KHz and so on, until completing the 30 channels. Figure 2 explains this concept in greater detail. The advantage with FDM is that it reduced considerably the cable requirements to connect subscribers.

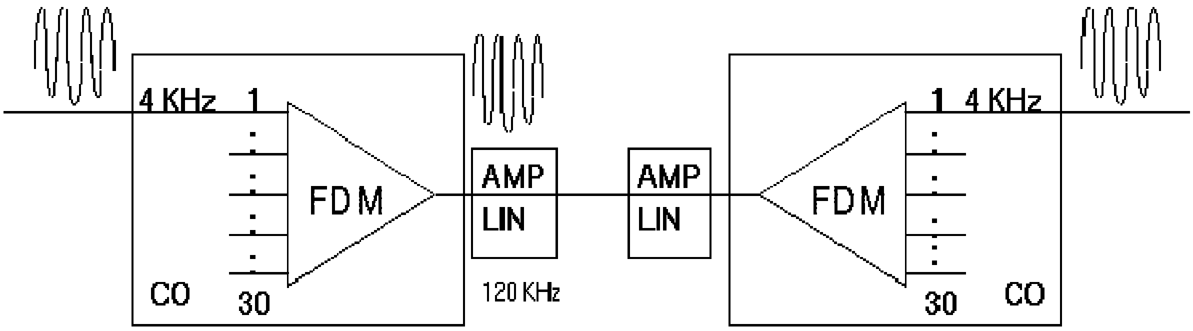

Figure 2. Frequency Division Multiplexing.

<u>Noise</u>.

Cables, as antennas, receive airwave signals, but in cables these signals interfere with the original information being sent, this interference is called noise, unwelcomed noise. Noise is generated mostly form sources of electrical radiation such as electric engines, fluorescent lights and even from other telephone lines [1]. Unfortunately in analog transmission systems noise cannot be separated from the original signal. As a transmitted voice signal progresses throughout a cable it starts to weaken and it could fade away before it reaches its destination; to avoid this the signal needs to be amplified; along the signal path noise has been incorporating to the original signal so when the signal is amplified the incorporated noise is amplified as well resulting in a corrupted signal at the receiving end. Figure 3 highlights this phenomenon.

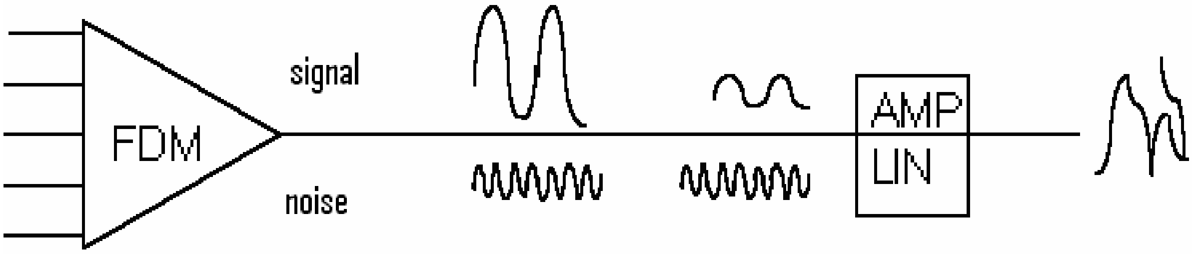

Figure 3. Signal and Noise are amplified together resulting in a corrupted signal.

Digital Systems.
There have been two major developments that changed telephony networks: Digital transmission and Common Channel Signaling [3]. Digital telephony transmit voice signals as bit strings, these strings maintain very low levels of noise facilitating signal switching and transmitting different signals in the same telephone line. Common Channel Signaling (CCS) allows control information to optimize the deployment of digital services. Figure 4 highlights a digital signal.

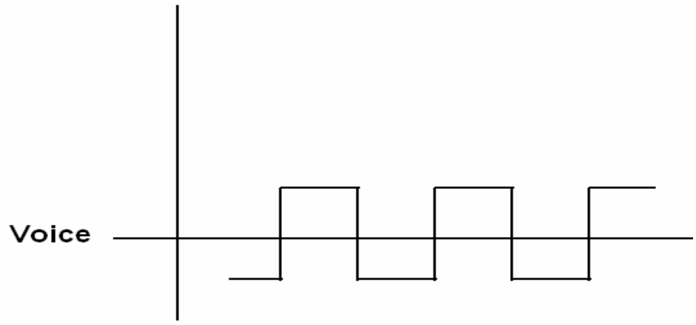

Figure 4. Digital transmission.

Voice digitalization tackles the noise problem very well. The main difference between analog and digital systems is that digital systems already “know” what they will be receiving from the transmitter. An analog signal could be any value in the signal curve, therefore any small variation caused by noise is difficult to detect and impossible to remove. A digital signal on the other hand can only have few values (-1, 0, +1), so any variation represents noise; during the transmission the digital signal is regenerated periodically restoring the original signal transmitted (Figure 5).

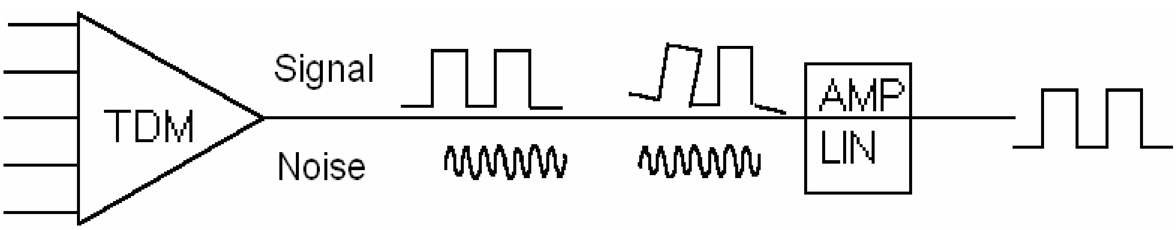

Figure 5. The digital signal is regenerated eliminating Noise.

The transmitter and digital logic lead to another technology: Time Division Multiplexing (TDM). TDM assigns time intervals to transmission channels (Generally 30 or 24 channels, depending of European or North American standards), and it rotates those channels. The most common voice digitalization method is Pulse Code Modulation, PCM. To convert analog voice into a digital signal it goes through a CODEC process (Codification-Decodification). The conversion takes place in two steps:

1.- Modulation by pulse amplitude. Based on the mathematical theorem by Harry Nyquist, a signal representing the voice variations entering the modulator is sampled 8,000 times per second, the modulator utilizes the resulting sample by transmitting a narrow wave pulse per each individual sample, which voltage (height) is the same as in the analog signal [3].

2.- Digital Codification. The pulse height is converted into a digital value; the output is an octet (byte) that represents the pulse voltage, i.e. the sampled voice. The two steps process converts the analog signal into a 64,000 bit string. (8,000 samples x 8 bits), see figure 6. The 64 KBPS (Kilo Bits Per Second) is commonly known as a DS-0 (Digital Signal level 0).

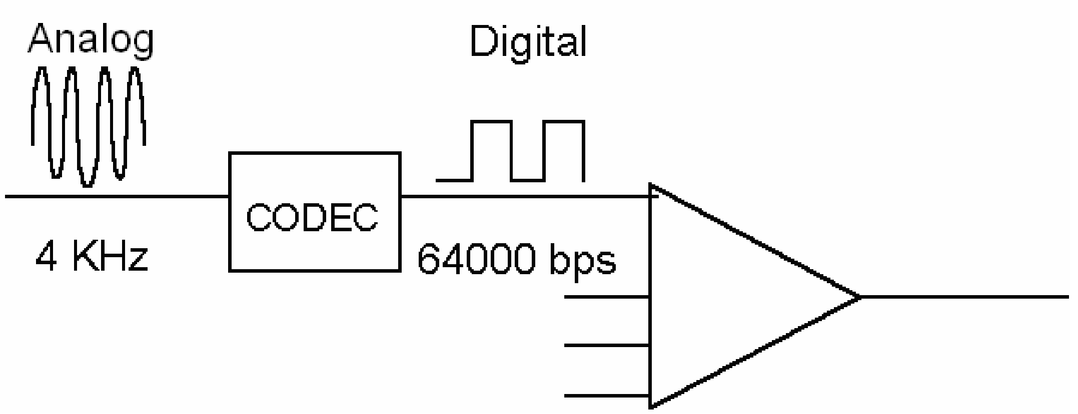

Figure 6. Codification-Decodification.

<u>Bandwidth Requirements</u>.
With the above description we can define now bandwidth in the most trivial way and it is the amount of information that it can be transmitted in any form of connection in a given time. Usually it is measured in bits per second.

Thanks to the Internet and all the services being offered over this technology, the per-capita bandwidth has been increasing worldwide as new services such as streaming video, Voice over IP (VoIP), conferencing, etc. are demanding more and more bandwidth, and bandwidth has been becoming more accessible to the end user, mostly in developed countries. Bandwidth availability has not been equal throughout the world. In underdeveloped countries prices are still prohibited. Table 1 highlights the current scenario in 30 countries:

Ranking	Nation	Household Penetration	Speed (Average Mbps)	Price (monthly)$	Composite Score
1	South Korea	0.93	49.5	0.37	15.92
2	Japan	0.55	63.6	0.13	15.05
3	Finland	0.61	21.7	0.42	12.20
4	Netherlands	0.77	8.8	1.90	11.77
5	France	0.54	17.6	0.33	11.59
6	Sweden	0.54	16.8	0.35	11.53
7	Denmark	0.76	4.6	1.65	11.44
8	Iceland	0.83	6.1	4.93	11.20
9	Norway	0.68	7.7	2.74	11.05
10	Switzerland	0.74	2.3	3.40	10.78
11	Canada	0.65	7.6	3.81	10.61
12	Australia	0.59	1.7	0.94	10.53
13	United Kingdom	0.55	2.6	1.24	10.30
14	Luxembourg	0.56	3.1	1.85	10.25
15	United States	0.57	4.9	2.83	10.25
16	Germany	0.47	6.0	1.10	10.17
17	Belgium	0.57	6.3	3.58	10.17
18	Portugal	0.44	8.1	1.24	10.15
19	New Zealand	0.42	2.5	1.05	9.68
20	Spain	0.49	1.2	2.27	9.68
21	Italy	0.41	4.2	1.97	9.54
22	Austria	0.45	7.2	4.48	9.37
23	Ireland	0.46	2.1	4.72	9.01
24	Greece	0.18	1.0	1.41	8.26
25	Hungary	0.29	3.3	4.67	8.22
26	Poland	0.23	7.9	6.47	7.83
27	Czech Republic	0.30	2.0	9.70	7.03
28	Slovak Republic	0.22	3.5	9.38	6.77
29	Turkey	0.23	2.0	15.75	5.25
30	Mexico	0.20	1.1	18.41	4.41
	Average	0.51	9.2	3.77	10.00

Table 1. Bandwidth prices[4].

As you can appreciate from the table above the bandwidth availability and price difference between South Korea and Mexico is huge.
Even though bandwidth has become more affordable, providers are still continuing to economize the bandwidth, that's why we have different rates for uploading and downloading and that's the reason why voice and video compression algorithms exist, to economize on bandwidth resources.

VoIP telephony providers require great lengths of bandwidth for servicing their end customers, for example let's estimate the bandwidth requirements for market standard Media Gateways. A typical Media Gateway that supports up to 16 E1s, which is the equivalent of 480 concurrent voice ports (calls), are normally 4 rack units (RU) tall therefore a conventional 19 inch rack could host 7 Media Gateways. This means that a full rack would serve 3,360 users simultaneously. If we acknowledge that G7.11 is the best CODEC then we need to reserve at least 64 KBPS per voice port for this CODEC, that's 215 MBPS of bandwidth to transmit voice packets over a data network, and this requirement is for voice media only, it is not considering the packet's that transmit the informational headers and trailers data.

And this requirement is minimum in comparison to physically smaller equipments with higher capacity, there are 2RU Media Gateways with DS-3 (Digital Signal, Level 3) interfaces out there. This means that a single one of these equipments can support up to 4,042 simultaneous voice calls. Being 2 RU tall means that 14 of these units would fit a typical 19 inch rack. Using a higher compression CODEC, this time, G.729, the amount of bandwidth needed for a fully loaded rack is 452 MBPS. That's a lot of bandwidth!!! Normally tier 1 carriers have more than one fully loaded rack in their location; they have several. So regardless of how inexpensive the Bandwidth is every provider will still try to economize on their Bandwidth resources.

There are several data network types out there, worldwide, that can be used for transmitting packet voice such as Frame Relay or MPLS (Multi Protocol Label Switching); to save bandwidth, voice compression algorithms are still needed.

<u>Compression Algorithms</u>.
We have loosely calculated bandwidth requirements and we spoken vaguely about compression algorithms. I mentioned the G.711and G.729 compression methods, in reality these "G" methods are ITU (International Telecommunications Union, a United Nations telecommunications body) recommendations.

The G.711 recommendation allows voice to be sent on 64 KBPS channels, the G.726 recommendation allows sending the same voice over half the bandwidth at 32 KBPS while the G.729 recommendation requires only 8 KBPS of bandwidth to send the same voice. The G.711 algorithm is the voice compression that we use in regular telephony, is the one that offers the best quality and it is the one that we're used to in plain old wireline telephony; and it is the one that uses the most bandwidth too. To offer this high quality voice service, telephone companies have deployed very expensive networked equipment near our homes, repeaters are needed every x amount of miles to ensure the level of service that we are used to with plain old telephony systems (POTS).

The Internet infrastructure is not as expensive as the POTS infrastructure and the Internet is being used to transport every imaginable, and even beyond imagination, multimedia type of service, including telephony. Many telephone companies have begun transporting telephony traffic over the Internet. Of course, these major carriers have to offer the best service possible to their subscribers and they have to guarantee the same quality of voice that we are so used to in POTS. Most of the carriers are still using G.711 as the voice compression algorithm. When G.711 is transported over the Internet it requires lots of bandwidth, more than when it is sent via the POTS because data packets, by the nature of their architectural design, attaches informational header and trailer packets that contains all of the specific information of that particular packet, like where is it going to, where is it coming from, the length of the packet, priority, etc. This header and trailer packets sometimes consume more bandwidth than the actual information.

Some companies have bandwidth limitations, especially smaller companies in developing countries, some of these companies look for alternative ways to G.711 while still maintaining high quality voice services. The CODEC that arguably offers a similar voice quality to G.711 is G.729 which compresses the voice to 1/8 of its original “size”. Compressing to 1/8 means in theory that a carrier can send eight voice channels using G.729 where only one G.711 channel would fit. That’s a huge saving!

The G.729 CODEC utilizes the CS-ACELP (Conjugate Structure Algebraic-Code-Excited Linear Prediction) methodology to compress the voice wave. CS-ACELP first obtains a wave length filtration of the analog signal from the telephone line, then it samples the signal at 8,000 Hz; it then converts it to a 16 bit linear PCM (Pulse Code Modulation) which is the equivalent to 128 KBPS. 128 KBPS is then compressed to 8 KBPS by assigning only 1 bit per each 8 KHz wave sample, to be more precise 80 bits per each 10 millisecond frame i.e. (80 bits x 1000 milliseconds = 8 KBPS). The output needs to be converted back to an analog signal by opposite-similar methods[5].

Over Head.

In the examples provided above for the equipments we mentioned that overhead (header and trailer packets) were not included in out math.

IP and Frame Relay are data networks that offer many variables for transmitting voice packets; where there is capacity restrains (and even where there are not) bandwidth calculation is necessary to design a network for transmitting voice services. Let’s examine all the variables for these two network types. In the samples below each field represents the quantity in BYTES required for a complete set of packets (Headers, media and trailers) for a given number of channels or calculating the number of channels that can be accommodated over a given WAN rate. But before we do that, here’s a small glossary:

FRF.11: Frame Relay Forum. Implementation agreement 11. This agreement specifies how to transport voice application data over a Frame Relay network.
FCS: Frame Check Sequence.
DLCI: Data Link Connection Identifier.
IP: Internet Protocol.
UDP: User Datagram Protocol
RTP: Real Time Protocol.
VPO: Voice Packet Optimization. VPO makes it possible to send voice packets from different channels under one overhead.

Packet Formats:

FRF.11 SubDLCI < 64

Delimiter (1)	FRF11 (3)	Payload	FCS(2)

OH = 6 bytes

FRF.11 SubDLCI > 63

Delimiter (1)	FRF11 (4)	Payload	FCS(2)

OH = 7 bytes

IP with FRF.11

Ethernet (26)	IP (20)	UDP (8)	FRF11(2)	Payload

OH = 56 bytes

IP with RTP

Ethernet (26)	IP (20)	UDP (8)	RTP(12)	Payload

OH = 66 bytes

IP with FRF.11and VPO

Ethernet (26)	IP (20)	UDP (8)	VPO (2)	FRF11(2)	Payload(1)

OH = 54 bytes + N*4bytes

VPO (2)	FRF11(2)	Payload(N)

Table 2. Packet Formats

Bandwidth Formulas:

The required “bandwidth” (bits per second) of each packet is given by:

PBW = ((OH + PAYLOAD) * 8)/PP where OH and PAYLOAD are in bytes and PP is the packetization period.

For Voice or FAX, PAYLOAD = (Rate * PP)/8 where rate is in bits per second. This result in:

PBW = ((OH + (Rate* PP)/8)*8)/PP
= ((OH*8)/PP) + Rate

For a FAX rate of 9600bps and packetization period of 40 ms, this equates to:

FBW = ((OH*8)/PP) + Rate = ((OH*8)/.040) +9600 (half-duplex)

For a worst case data upspeed rate of 64000bps and packetization period of 10 ms, this equates to:

DBW = ((OH*8)/PP) + Rate = ((OH*8)/.010) +64000 (full-duplex)

For a voice rate of V, packetization period of PP ms and voice activity A, this equates to:

VBW = ((OH*8)/PP) + Rate = (((OH*8)/PP) + V)*A

Average channel bandwidth for N channels with PF percentage FAX and PD percentage data is:

ABW = ((FBW * N*PF) + (DBW * N * PD) + (VBW * N * (1-PF-PD))/N

Since VPO is a dynamic structure in which the frame size and corresponding overhead may vary, a simplified approach is taken. The assumption is that an average 5 packets will be present in any frame. This result in a total overhead of 54 + 20 bytes distributed across five channels. The resultant OH per channel is approximately 14.8 bytes or 118 bits.

Delay Formulas:

Average Delay = (4.5 * FP) + (PP-FP) where FP is the basic frame period and PP is the packetization period.

Packets Per Second:

Packets per second = 1/PP where PP is the packetization period.

Average packets per second = weighted average of voice (including VAD) + FAX + Data pps requirements.

Let's use the above formulas for an actual bandwidth calculation example:

For this example we will be compressing a voice wave signal using the G.729 CODEC over the Internet Protocol and using RTP as the payload. We will use a packetization period of 20 milliseconds.

The formula is as follow:

a)PP=20 means that the packetization period is 20 milliseconds, i.e. each 20 milliseconds a G.729 packet will be processed and sent out to the network, therefore in a given second we will have 50 packets (each second has 1,000 milliseconds 1,000/20=50).
b)The RTP packet is 12 bytes long
c)The UDP packet is 08 bytes long
d)The IP packet is 20 bytes long
e)The G.729 packet in this case is 20 bytes long (G.729 is 8,000 bits long, dividing it by 8, which is the amount of bits in a byte, will give us 1,000 bytes. Since we have 50 instances in a second this result in 1,000/50=20bytes.

Therefore the total size of a G.729 packet over IP using RTP as the media payload is: 12+8+20+20 = 60 bytes. Each 20 ms 60 bytes are sent meaning that each second 3,000 bytes are sent. Multiplying the 3,000 bytes per 8 (amount of bits in a byte) then the bandwidth requirement is 24,000 bits of bandwidth per channel for an 8,000 bit compression.

Other elements play a factor in bandwidth savings like a concept called VAD (Voice Activity Detection). VAD helps to save bandwidth by not sending packets during silence periods during a regular conversation. You would be surprised that in alphabet based languages up to 60% of a conversation is composed by silence pauses.

Conclusion.
Bandwidth calculation is still necessary in our days, it still need to be economized, one day it will be as hard disk space is today: very affordable and available to anybody, in the mean time we still have to do some math !!!

References and further reading.

Stufflebeam, Jim; The Bandwidth Calculator. Nuera 2001.
Chhatwal, Harprit; McNeil, Scott, Wong, Daniel; Wang Dave; Zuercher Fredy; Creamer, Mike.
The Bandwidth Calculator, Internal discussions. Nuera 2001 and 2008.
[1] Flanagan William A. The Guide to T-1 Networking. Telecom Library, 1988.
[2] Black Uyless. ATM: Foundation For Broadband Networks. Prentice Hall, 1995.
[3] Walrand Jean. Communication Networks. Aksen Associates, 2007.
[4] Meiselmann, Arnoldo. Cablemás.
[5] ITU-T Recommendation G.729. 2001
www.eveliux.com
www.audiocodes.com
www.develcon.com/kb/a014.htm

SERGIO J. CASTRO is a Cisco Certified Network Associate; he has an extensive career as a Senior Systems Engineer for such companies as Movius Interactive, IP Unity, Nuera Communications, IBM and Red Uno.

Working for these companies he has planned, sized, designed, deployed and documented Voice over Data Networks all over the world. He is currently pursuing a MSc in Networks and Telecommunications degree at CETyS University in Mexico. He holds a BS in Computer Science from The University of Baja California, Ensenada. He calls home both San Diego, California and Ensenada, Baja California. He can be reached at: scastro@rocketmail.com