SmartBits

Performance Analysis System

Voice over IP (VoIP)

SPIRENT Communications

P/N 340-1158-001 REV A, 8/01

Spirent Communications, Inc.
(800) 886-8842 Toll Free
(818) 676-2300 Phone
(818) 881-9154 FAX

Disclaimer

The information contained in this document is subject to change without notice and does not represent a commitment on the part of Spirent Communications. The information in this document is believed to be accurate and reliable, however, Spirent Communications assumes no responsibility or liability for any errors or inaccuracies that may appear in the document.

Trademarks

AST II™, ScriptCenter™, SmartApplications™, SmartBits®, SmartCableModem™, SmartFabric™, SmartFlow™, SmartLib™, SmartMetrics™, SmartMulticastIP™, SmartSignaling™, SmartTCP™, SmartVoIPQoS™, SmartWindow™, SmartxDSL™, TeraMetrics™, TeraMobileIP™, TeraRouting Tester™, TeraVPN™, VAST™, and WebSuite™ are trademarks or registered trademarks of Spirent Communications, Inc.

All other trademarks and registered trademarks are the property of their respective owners.

Warranty

Spirent Communications warrants to recipient that hardware which it supplies with this document (“Product”) will be free from significant defects in materials and workmanship for a period of twelve (12) months from the date of delivery (the “Warranty Period”), under normal use and conditions.

Defective Product under warranty shall be, at Spirent Communications’ discretion, repaired or replaced or a credit issued to recipient’s account for an amount equal to the price paid for such Product provided that: (a) such Product is returned to Spirent Communications after first obtaining a return authorization number and shipping instructions, freight prepaid, to Spirent Communications’ location in the United States; (b) recipient provide a written explanation of the defect claimed; and (c) the claimed defect actually exists and was not caused by neglect, accident, misuse, improper installation, improper repair, fire, flood, lightning, power surges, earthquake or alteration. Spirent Communications will ship repaired Product to recipient, freight prepaid, within ten (10) working days after receipt of defective Product. Except as otherwise stated, any claim on account of defective materials or for any other cause whatsoever will conclusively be deemed waived by recipient unless written notice thereof is given to Spirent Communications within the Warranty Period. Product will be subject to Spirent Communications’ standard tolerances for variations.

TO THE EXTENT PERMITTED BY APPLICABLE LAW, ALL IMPLIED WARRANTIES, INCLUDING BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, NONINFRINGEMENT AND FITNESS FOR A PARTICULAR PURPOSE, ARE HEREBY EXCLUDED, AND THE LIABILITY OF SPIRENT COMMUNICATIONS INC., IF ANY, FOR DAMAGES RELATING TO ANY ALLEGEDLY DEFECTIVE PRODUCT SHALL BE LIMITED TO THE ACTUAL PRICE PAID BY THE PURCHASER FOR SUCH PRODUCT. IN NO EVENT WILL SPIRENT COMMUNICATIONS INC. BE LIABLE FOR COSTS OF PROCUREMENT OF SUBSTITUTE PRODUCTS OR SERVICES, LOST PROFITS, OR ANY SPECIAL, DIRECT, INDIRECT, CONSEQUENTIAL, OR INCIDENTAL DAMAGES, HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, ARISING IN ANY WAY OUT OF THE SALE AND/OR LICENSE OF PRODUCTS OR SERVICES TO RECIPIENT EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGES AND NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY.

Contents

Voice over IP (VoIP)

Introduction

Interest in Voice over IP (VoIP) has increased steadily over the past few years. Enterprises, ISPs, ITSPs (Internet Telephony Service Providers), and carriers view VoIP as a viable way to implement packet voice. Reasons for implementing VoIP typically include toll-bypass, network consolidation, and service convergence. Toll-bypass allows long-distance calls to be placed without incurring the usual toll charges. Through network consolidation, voice, video, and data can be carried over a single network infrastructure, thereby simplifying network management and reducing cost through the use of common equipment. With service convergence, enhanced functionality can be implemented through the coupling of multimedia services. This full integration permits new applications, such as unified messaging and web call center.

However, designing a VoIP network requires careful planning to ensure that voice quality can be properly maintained. This document examines the factors that affect voice quality and the test and analysis strategy for a VoIP network.

VoIP Components

Figure 1 on page 2 shows the major components of a VoIP network. The gateway converts signals from the traditional telephony interfaces (POTS, T1/E1, ISDN, E&M trunks) to VoIP. An IP phone is a *terminal* that has native VoIP support and can connect directly to an IP network. In this paper, the term terminal will be used to refer to either a gateway, an IP phone, or a PC with a VoIP interface.[1] The server provides management and administrative functions to support the routing of calls across the network. In a system based on H.323, the server is known as a *gatekeeper.* In SIP/SDP, the server is a *SIP server.* In a system based on MGCP or MEGACO, the server is a *call agent*. Finally, the IP network provides connectivity between all the terminals. The IP network can be a private network, an Intranet, or the Internet.

1. This is consistent with the terminology used in H.323. However, this paper does not assume that the VoIP system is necessarily based on H.323.

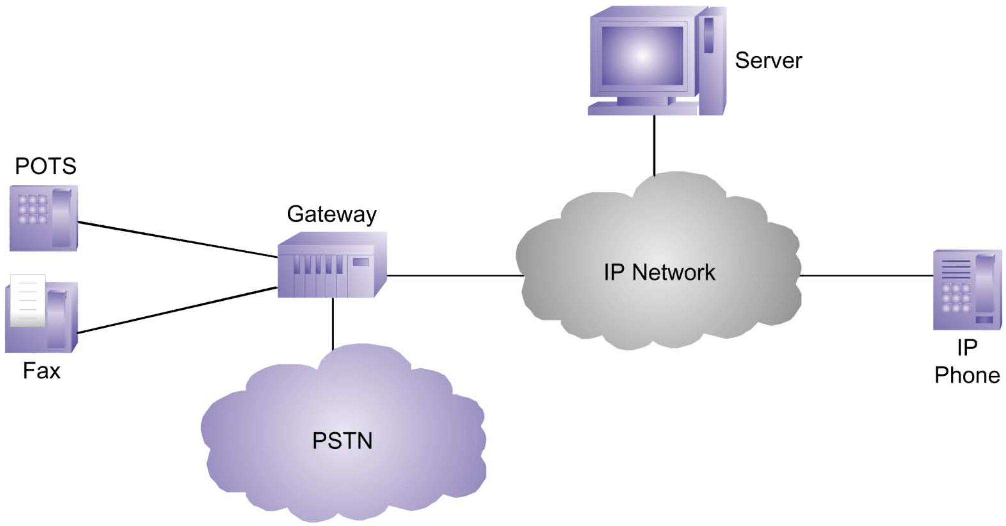


Figure 1. VoIP Components

Once a call has been set up, speech will be digitized and then transmitted across the network as IP frames. Voice samples are first encapsulated in RTP (Real-time Transport Protocol) and UDP (User Datagram Protocol) before being transmitted in an IP frame. *Figure 2* shows an example of a VoIP frame in both LAN and WAN.

PPP	IP	UDP	RTP	Voice samples	FCS	
4	20	8	12	Depends on CODEC	2	octets

ETH	IP	UDP	RTP	Voice samples	FCS	
14	20	8	12	Depends on CODEC	4	octets

Figure 2. Encapsulation of VoIP Frame

For example, if the CODEC used is G.711 and the packetization period is 20 ms, the payload will be 160 bytes. This will result in a total frame length of 206 bytes in WAN and 218 bytes in LAN.

Voice Quality

In designing a VoIP network, it is important to consider all the factors that will affect voice quality. A summary of the major factors follows.

CODEC

Before analog voice can be transmitted over an IP network, it must first be digitized. The common coding standards are listed in the following table:

Coding Standard	Algorithm	Data Rate
G.711	PCM (Pulse Code Modulation)	64 kbps
G.726	ADPCM (Adaptive Differential Pulse Code Modulation)	16, 24, 32, 40 kbps
G.728	LD-CELP (Low Delay Code Excited Linear Prediction)	16 kbps
G.729	CS-ACELP (Conjugate Structure Algebraic CELP)	8 kbps
G.723.1	MP-MLQ (Multi-Pulse Maximum Likelihood Quantization)	6.3 kbps 5.3 kbps
	ACELP (Algebraic Code Excited Linear Prediction)	6.3 kbps 5.3 kbps

There is a general correlation between the voice quality and the data rate: the higher the data rate, the higher the voice quality. The relationship between the two will be examined in greater detail in *"Mean Opinion Score (MOS)" on page 10*.

Frame Loss

VoIP frames have to traverse an IP network, which is unreliable. Frames may be dropped as a result of network congestion or data corruption. Furthermore, for real-time traffic like voice, retransmission of lost frames at the transport layer is not practical because of the additional delays. Hence, voice terminals have to deal with missing voice samples, also referred to as *frame erasures*. The effect of frame loss on voice quality depends on how the terminals handle frame erasures.

In the simplest case, the terminal leaves a gap in the voice stream if a voice sample is missing. If too many frames are lost, the speech will sound choppy with syllables or words missing. One possible recovery strategy is to replay the previous voice sample. This works well if only a few samples are missing. To better cope with burst errors, interpolation is usually used. Based on the previous voice samples, the decoder will predict what the missing frames should be. This technique is known as *Packet Loss Concealment* (PLC).

For example, ITU-T G.711 Appendix I describes a PLC algorithm for PCM. A circular history buffer consisting of 48.75 ms of the previous voice samples is kept. Once frame erasure is detected, the contents of the history buffer will be used to estimate the current pitch period. This information will then be used to generate a synthesized signal to fill in the gap. With PLC in G.711, the audio output is delayed by an additional 3.75 ms to provide a smooth transition between the real and synthesized signals. CELP-based speech coders such as G.723.1, G.728, and G.729 also have PLC algorithms built into their standards. In general, if the erasures are not too long, and the signal is not changing very rapidly, the erasures may be inaudible after concealment.

ITU-T G.113 Appendix I provides some provisional planning guidelines on the effect of frame loss on voice quality. The impact is measured in terms of Ie, the equipment impairment factor.[2] This is a dimensionless number in which 0 means no impairment. The larger the Ie factor, the more severe the impairment. The following table is derived from G.113 Appendix I and it shows the impact of frame loss on the Ie factor.

When the frame loss rate is 2%, the equipment impairment is 35 for standard G.711. However, with PLC, the equipment impairment is reduced to 7. Note that with low bit-rate coders such as G.729A and G.723.1, there is equipment impairment of 11 and 15 respectively even with no frame loss. A 2% frame loss will increase the impairment to 19 and 24 respectively.

Codec	Ie (0% loss)	Ie (2% random frame loss)	Ie (5% random frame loss)
G.711 without PLC	0	35	55
G.711 with PLC	0	7	15
G.729A	11	19	26*
G.723.1 (6.3 kbps)	15	24	32†

* The values were for 4% random frame loss. The values for 5% were not provided in the Appendix.
† The values were for 4% random frame loss. The values for 5% were not provided in the Appendix.

2. The Ie factor is derived subjectively using a Mean Opinion Score (MOS). The methodology is described in G.113 Annex E. The Ie factor can also be used in the E-Model, which is discussed later in *"Transmission Characteristics and the E-Model" on page 12*.

Delay

Another important consideration in designing a VoIP network is the effect of delay. Impairments caused by delays include echo and talker overlap. The effect of delay on voice transmission is discussed in ITU-T G.114.

Sources of Delays

Before assessing the impact of delay, it is useful to first identify the sources of delays.

Algorithmic Delay. This is the delay introduced by the CODEC and is inherent in the coding algorithm. The following table summarizes the algorithmic delay of common coding standards.

Coding Standards	Algorithmic Delay (ms)
G.711	0.125*
G.726	1
G.728	3-5
G.729	15†
G.723.1	37.5‡

* The algorithmic delay can be 3.75ms if PLC is implemented.
† Includes lookahead buffer.
‡ Includes lookahead buffer.

Packetization Delay. In RTP, voice samples are often accumulated before putting into a frame for transmission to reduce the amount of overhead. RFC 1890 specifies that the default packetization period should be 20 ms. For G.711, this means that 160 samples will be accumulated and then transmitted in a single frame. On the other hand, G.723.1 generates a voice frame every 30 ms and each voice frame is usually transmitted as a single RTP packet.

Serialization Delay. This is the time required to transmit the IP packet. For example, if G.711 is used and the packetization period is 20 ms (i.e., there are 160 bytes in the RTP payload), then the entire frame will be 206 bytes assuming PPP encapsulation. To transmit the frame, it will require 1.1 ms on a T1 line, 3.2 ms at 512 kbps, and 25.8 ms at 64 kbps. Furthermore, the serialization delay is incurred whenever it passes through another store-and-forward device such as a router or a switch. Thus, a frame that traverses 10 routers will incur this delay 10 times.

Propagation Delay. This is the time required for the electrical or optical signal to travel along a transmission medium and is a function of the geographic distance. The propagation speed in a cable is approximately 4 to 6 microseconds per kilometer. For satellite transmission, the delay is 110 ms for a 14000-km altitude satellite and 260 ms for a 36000-km altitude satellite.

Component Delay. These are delays caused by the various components within the transmission system. For example, a frame passing through a router has to move from the input port to the output port through the backplane. There is some minimum delay due to the speed of the backplane and some variable delays due to queuing and router processing.

Echo Cancellation

The first impairment caused by delay is the effect of echo. Echo can arise in a voice network due to poor coupling between the earpiece and the mouthpiece in the handset. This is known as *acoustic echo*. It can also arise when part of the electrical energy is reflected back to the speaker by the hybrid circuit[3] in the PSTN (Public Switched Telephone Network). This is known as *hybrid echo*.

When the one-way end-to-end delay is short, whatever echo that is generated by the voice circuit will come back to the speaker very quickly and will not be noticeable. In fact, the guideline is that echo cancellation is not necessary if the one-way delay is less than 25 ms. In other words, if the echo comes back within 50 ms, it will not be noticeable. However, the one-way delay in a VoIP network will almost always exceed 25 ms. Therefore, echo cancellation is always required.

Talker Overlap

Even with perfect echo cancellation, carrying on a two-way conversation becomes difficult when the delay is too long because of talker overlap. This is the problem that occurs when one party cuts off the other party's speech because of the long delay. G.114 provides the following guidelines regarding the one-way delay limit:

0 to 150 ms	Acceptable for most user application
150 to 400 ms	Acceptable provided that Administrations are aware of the transmission time impact on the transmission quality
Above 400ms	Unacceptable for general network planning purposes

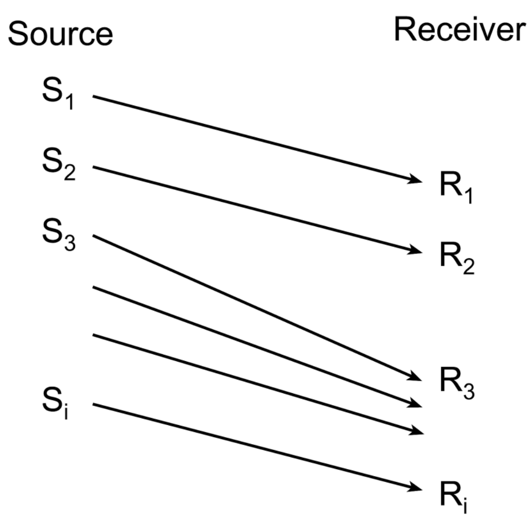
Delay Variation (Jitter)

When frames are transmitted through an IP network, the amount of delay experienced by each frame may differ. This is because the amount of queuing delay and processing time can vary depending on the overall load in the network. Even though the source gateway generates voice frames at regular intervals (say, every 20 ms), the destination gateway will typically not receive voice frames at regular intervals because of jitter. This is illustrated in *Figure 3 on page 7*.

3. The hybrid circuit converts the two-wire PSTN circuit to a four-wire circuit.

In general, jitter will result in clumping and gaps in the incoming data stream. The general strategy in dealing with jitter is to hold the incoming frames in a playout buffer long enough to allow the slowest frames to arrive in time to be played in the correct sequence. The larger the amount of jitter, the longer some of the frames will be held in the buffer, which introduces additional delay.

To minimize the delay due to buffering, most implementations use an adaptive jitter buffer. In other words, if the amount of jitter in the network is small, the buffer size will be small. If the jitter increases due to increased network load, the buffer size will increase automatically to compensate for it. Therefore, jitter in the network will impair voice quality to the extent that it increases the end-to-end delay due to the playout buffer. Sometimes when the jitter is too large, the playout buffer may choose to allow some frame loss to keep the additional delay from getting too long.



$$D_i = (R_i - S_i) - (R_{i-1} - S_{i-1})$$
$$= (R_i - R_{i-1}) - (S_i - S_{i-1})$$
$$Jitter = \frac{\sum_i^n |D_i|}{n}$$

Figure 3. Jitter

Delay Budget

Figure 4 shows an example of a VoIP network and the sources of delay. The following delay budget can be constructed. Assume an end-to-end delay target of 150 ms.

G.723.1 (algorithmic delay)	37.5
G.723.1 (processing delay)	30.0
Serialization delay (two T1s)	2.0
Propagation delay (5000km of fiber)	25.0
Other component delays	2.0
Total fixed delay	96.5

Variable delay limit = 150 – 96.5 = 53.5 ms

In this example, the fixed (minimum) delay is calculated to be 96.5 ms. The presence of jitter will add to the end-to-end delay. How much jitter can the system tolerate? If the end-to-end delay target is 150 ms, then the maximum jitter that can be tolerated is 53.5 ms. The assumption is that the jitter will be removed by a playout buffer which can delay frames by up to 53.5 ms to remove the jitter.

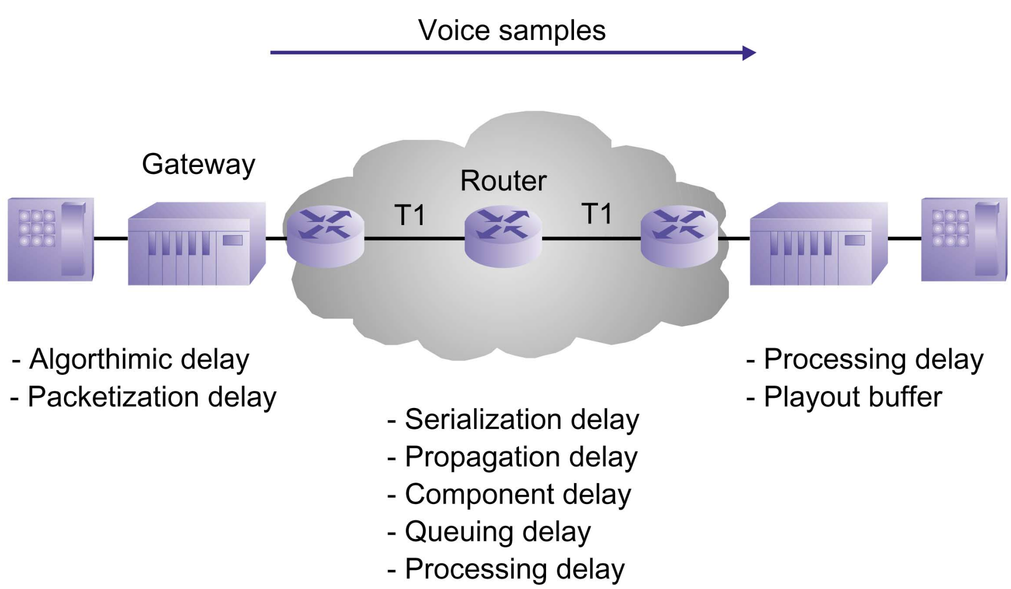


Figure 4. Delay Budget Example 1

However, this example assumes that you knew the exact topology of the network, and thus were able to calculate all the delay components. In the next example (*Figure 5 on page 9*), we assume that the voice gateways are connected via a VPN service offered by an ISP.

Assume an end-to-end delay target of 150 ms:

G.723.1 (algorithmic delay)	37.5
.723.1 (processing delay)	30.0
Total gateway delay	67.5

Internet delay limit = 150 – 67.5 = 82.5 ms

In this example, we can only identify the delays due to the two gateways. To stay within the delay target of 150 ms, the delay introduced by the ISP must not exceed 82.5 ms. Note that this represents both the fixed and variable delays. In other words, the minimum delay along the VPN path might be 50 ms. The maximum jitter that the system can tolerate will be 32.5 ms, which will be compensated by the playout buffer. Today, many ISPs offer VPN service with a Service Level Agreement (SLA). An SLA will typically guarantee a certain round-trip delay between sites.

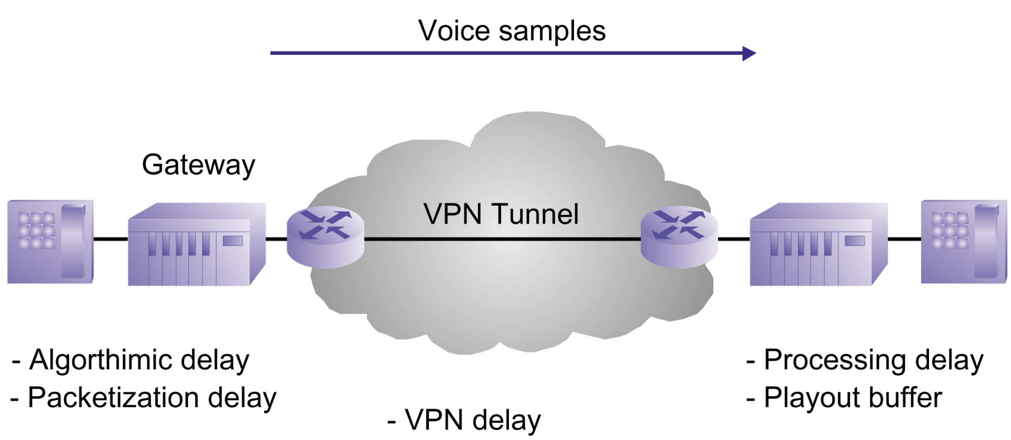


Figure 5. Delay Budget Example 2

Measuring Voice Quality

In the previous section, guidelines were provided on how to design a network to ensure good voice quality. However, we still need the ability to measure and compare voice quality. This can be done using a number of different techniques.

Mean Opinion Score (MOS)

Described in ITU-T P.800, MOS is the most well-known measure of voice quality. It is a subjective method of quality assessment. There are two test methods: conversation-opinion test and listening-opinion test. Test subjects judge the quality of the voice transmission system either by carrying on a conversation or by listening to speech samples. They then rank the voice quality using the following scale:

5 – Excellent, 4 – Good, 3 – Fair, 2 – Poor, 1 – Bad

MOS is then computed by averaging the scores of the test subjects. Using this scale, an average score of 4 and above is considered as toll-quality. MOS was originally designed to assess the quality of different coding standards. The following is a summary of the MOS for different coding algorithms.

Coding Standard	MOS
G.711	4.3 – 4.4 (64 kbps)
G.726	4.0 – 4.2 (32 kbps)
G.728	4.0 – 4.2 (16 kbps)
G.729	4.0 – 4.2 (8 kbps)
G.723.1	3.8 – 4.0 (6.3 kbps)
	3.5 (5.3 kbps)

MOS is the most relevant test, because it is humans who use the voice network and it is humans whose opinions count. However, a subjective test that involves human subjects can be time-consuming to administer. Hence, there is a lot of interest in devising objective tests that can be used to approximate human perception of voice quality.

Perceptual Speech Quality Measure (PSQM)

Described in ITU-T P.861, PSQM (*Figure 6*) uses a psychoacoustic model to mathematically compute the differences between the input and output signals.

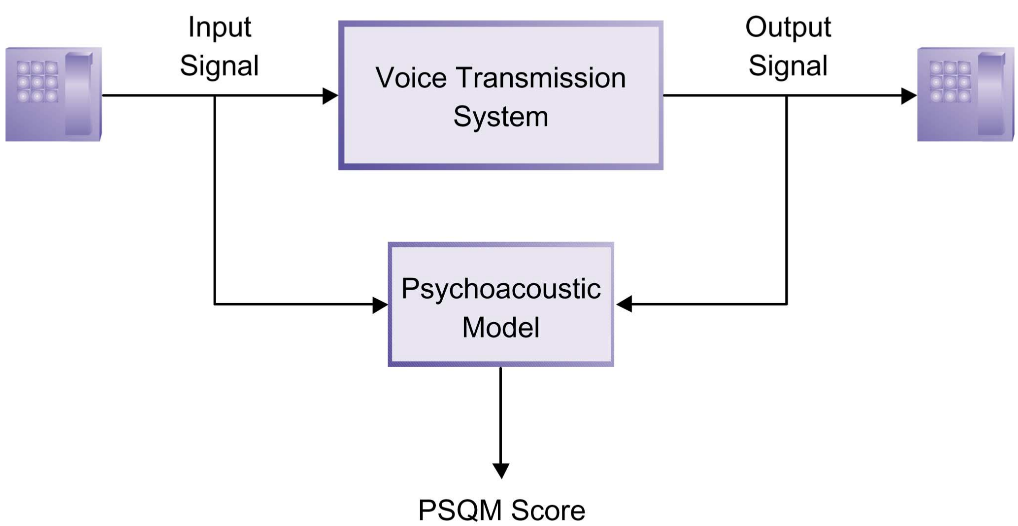


Figure 6. PSQM

Using this method, if the input and output signals are identical, the PSQM score will be zero. The bigger the differences, the higher the score will be up to a maximum of 6.5. However, unlike more traditional measurements such as signal-to-noise ratio (SNR), the emphasis of PSQM is on differences that will affect human perception of speech quality.

One of the PSQM's criticisms is that it was originally designed to measure the quality of coding standards. Therefore, it does not fully take into account the effect of various transmission impairments. PSQM+ was proposed in December 1997 and accounts for:

- Different perceptions due to volume or loud distortions
- Speech that has dropouts

With PSQM+, the correlation between the objective score and MOS is improved.

Other Speech Quality Measures

There are a number of other objective measures that either have been proposed or are in use, including:

- Measuring Normalizing Blocks (MNB) – described in P.861 Appendix II
- Perceptual Analysis Measurement System (PAMS) – a proprietary system developed by British Telecom

- Perceptual Evaluation of Speech Quality (PESQ) – a proposed standard being considered by the ITU-T

Transmission Characteristics and the E-Model

In a VoIP network, transmission impairments play a very important role in determining voice quality. As discussed in *"Voice Quality" on page 3*, these transmission impairments include frame loss, delay, and jitter. Another approach in voice quality testing is to measure directly those transmission impairments and then predict what the voice quality will be given those impairments.

The E-Model, as described in ITU-T G.107, provides a useful computational model for predictive analysis. The basic equation of the model is as follows:

$$R = Ro - Is - Id - Ie + A$$

R	Transmission rating factor
Ro	Basic signal-to-noise ratio. This is computed from all circuit noise powers.
Is	Simultaneous impairment factor. This accounts for impairments caused by non-optimum sidetone and quantizing distortion.
Id	Delay impairment factor. This accounts for impairments caused by a delay in the network.
Ie	Equipment impairment factor. This accounts for impairments caused by low bit rate coders as well as the effect of frame loss on the coder. This was discussed earlier in detail in *"Frame Loss" on page 3*.
A	Expectation factor. This is a correction factor that adjusts perceived quality based on user expectation. For example, if users are aware that they are communicating with a hard-to-reach location via multi-hop satellite connections, they may be more willing to tolerate impairments due to long delays.

For example, once the transmission impairments in an IP network have been measured, the E-Model can be used to calculate the transmission rating factor.[4] The transmission rating factor can then be transformed into MOS using the following equations:

- For $R < 0$:
  $MOS = 1$
- For $0 < R < 100$
  $MOS = 1 + 0.035R + 7R(R-60)(100-R) \times 10-6$

4. Based on the transmission impairments, different transmission rating factors can be computed depending on the coder used. Each coder will result in a different Ie factor and thus a different R factor.

- For R > 100
  MOS = 4.5

Which Voice Quality Measure Should be Used

Given the plethora of measurement methods, which one should be used? In practice, a number of methods can be used in combination. As mentioned previously, MOS is the most relevant measure because it is the human opinion that counts the most. So it should always be used as a reality check. Instead of conducting a formal MOS test, you may choose to run a pilot of a VoIP network, let a select group of users try out the system, and then provide you with feedback. However, when you are configuring the system, you may make many adjustments, and getting human subjects to assess the effect may be impractical. In these cases, an objective test system such as PSQM, PAMS, or PESQ may be more convenient. In VoIP, QoS is a very important component. In measuring the effectiveness of QoS, measuring the transmission impairments (frame loss, delay, and jitter) is more useful because it helps to directly answer questions such as:

- If the network is congested, are voice frames given higher priority compared with data frames?
- What is the average delay experienced by voice frames?
- What is the average jitter seen by voice frames?

In testing a VoIP network, it is necessary to create a realistic test environment. Typically, this means that there are a number of concurrent voice sessions and that both voice and data traffic are present and competing for bandwidth.

Testing VoIP

The previous section examined the various voice quality measures. In this section, we will examine the different test configurations and the objectives of the tests.

IP Network Analysis

The main objective of this test is to measure the transmission characteristics of an IP network to determine if it can support VoIP applications. It is also an important test to measure the effectiveness of QoS mechanisms. *Figure 7* shows the typical test configuration.

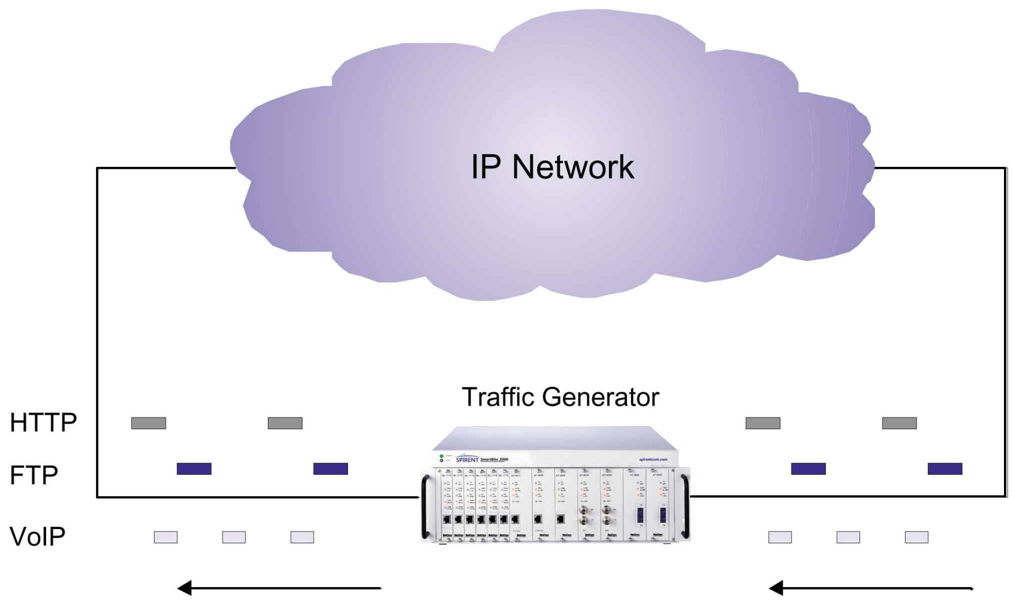


Figure 7. IP Network Analysis

To test the effectiveness of QoS, you must be able to simulate a mixture of voice and data traffic. By measuring the transmission characteristics of each flow, you can test whether the IP network provides different treatment to voice and data traffic. Note that this test does not involve the use of any VoIP equipment.

End-to-End Voice Analysis

The objective is to test the ability of the VoIP network to transmit voice and other related signals from end to end. The most important part of this test is to assess speech quality. *Figure 8* shows the typical test configuration. This configuration allows a number of different tests, as described below.

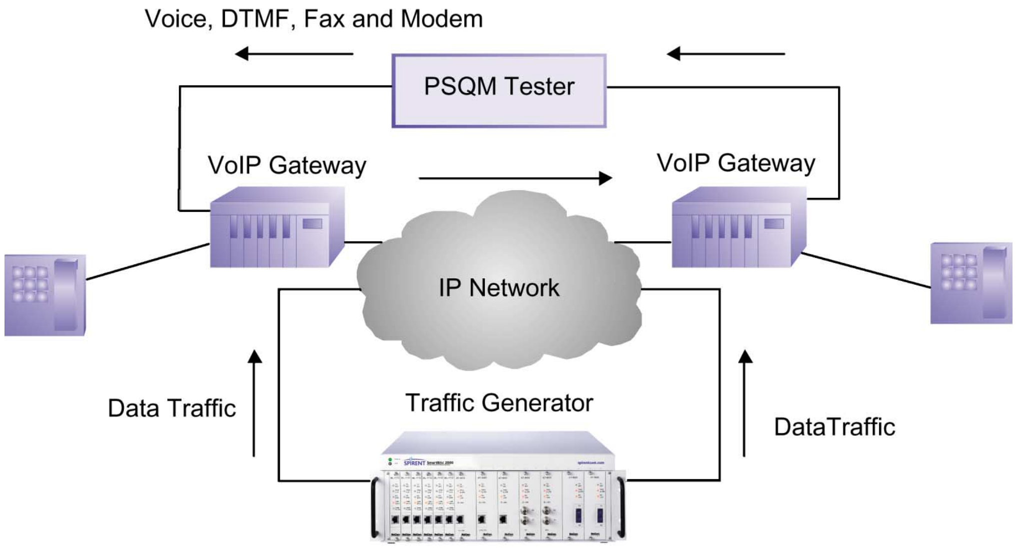


Figure 8. End-to-End Voice Analysis

Using the handsets, a voice call can be placed and human subjects can be used to assess the quality of the voice transmission.

Using a PSQM test system, the same test can be done objectively. This allows different gateway configurations (changing the CODEC or turning voice activity detection on and off) to be rapidly tested. The traffic generator can also be used in this configuration to increase the load on the IP network, thereby testing the QoS of the network.

In addition to testing voice transmission, other types of information may also be transmitted over the VoIP network. These include DTMF tones, fax, and modem. For example, if you configure the gateway to use a low bit-rate coder such as G.729 with VAD, a fax will not transmit properly. However, some gateways can detect the presence of the fax tone and will switch over to G.711 and turn off VAD automatically. Other gateways implement fax-relay that extracts the fax data and only transmits the data across the IP network. These mechanisms need to be tested.

Another common test configuration involves the use of an impairment simulator. When testing voice terminals, it is often desirable to test their performance under degraded operating conditions. For example, determining what happens to voice quality if the frame loss rate is 1%, 2%, 3%, and so on. To a certain extent, the insertion of data traffic into the network to cause congestion can achieve that. However, it is difficult to adjust the data traffic level to cause a frame loss rate of, say, exactly 3%. An impairment simulator can be used to precisely create a set of degraded operating conditions. The resulting voice quality can again be measured either subjectively by human listeners or objectively using a PSQM test system. This test configuration is illustrated in *Figure 9*.

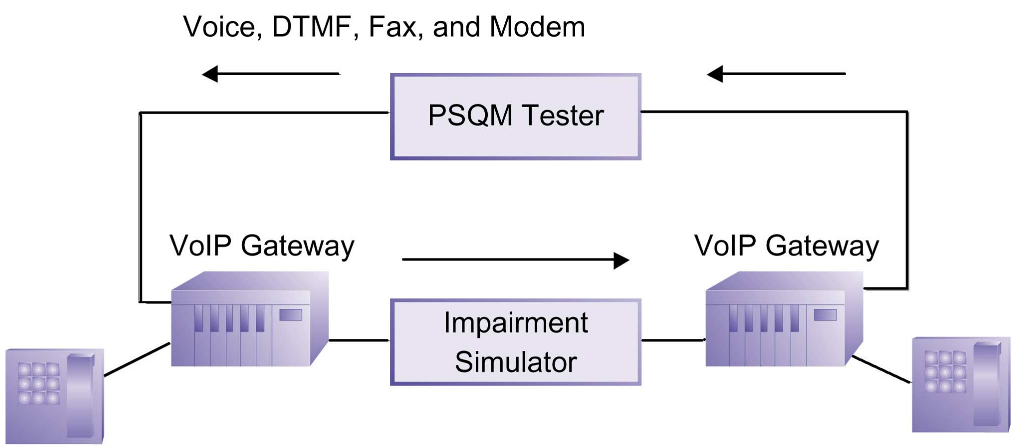


Figure 9. Impairment Simulation

Signaling Stress Test

This test focuses on the scalability of a VoIP network. In the PSTN, traditional telephone switches or PBXs have been tested extensively to ensure that they can handle a large number of calls. In a VoIP network, call routing is handled by devices such as gatekeepers (in H.323), call agents (in MGCP and MEGACO) and SIP servers (in SIP). These devices must also be tested in a similar way. This test can be performed with the help of a bulk call generator as shown in *Figure 10 on page 17*.

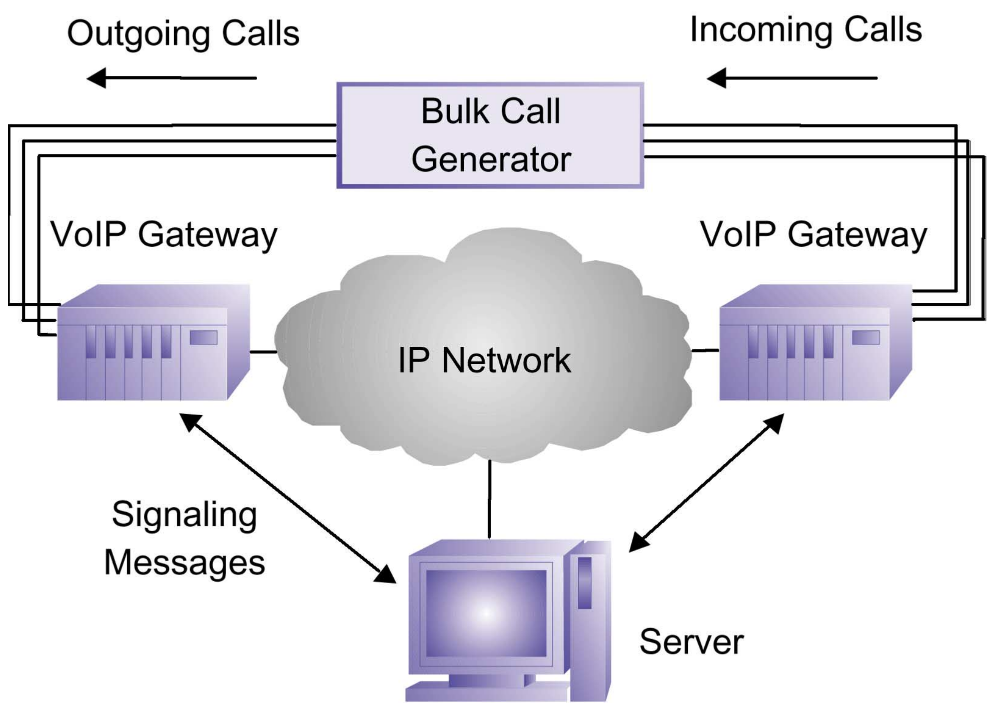


Figure 10. Signaling Stress Test

Typical measurements include:

- What is the highest call rate the servers can sustain?
- What is the highest number of calls the servers can maintain simultaneously?
- What is the call setup time in relation to the load?

Case Studies

Quality-of-Service (QoS)

Whether VoIP is implemented using dedicated or shared facilities, QoS is an important consideration. A QoS-enabled network will differentiate between different types of traffic and offer different treatments. Using standard-based methods, this can be achieved using either the TOS (Type Of Service) bits or the DiffServ (Differentiated Services) field in the IP header, or through the use of signaling protocols such as RSVP (Resource reSerVation Protocol) and MPLS (Multi-Protocol Label Switching). Routers and switches also support prioritization based on physical port, protocol, IP addresses, transport addresses, or even frame length.

In analyzing whether an IP network can support VoIP, the effectiveness of its QoS must be evaluated. In particular, questions that are of interest include the following:

- How can the network differentiate VoIP traffic from other types of traffic?[5]
- In the event of network congestion, what is the frame loss rate of voice vis-à-vis data?
- What is the average delay experienced by voice frames?
- What is the average jitter?
- Can the network scale if there is a large number of flows?

The following series of tests were designed to investigate the behavior of common prioritization schemes in routers. *Figure 11 on page 19* shows the test configuration.

Using the SmartBits, a number of traffic flows were defined and injected into the IP network. These traffic flows represented different types of traffic. They differed from one another in terms of IP addresses, IP precedence, port numbers, size of packets, or a combination of these factors. By observing the output of the IP network, we can determine how the network treated the traffic differently. To see the effect of QoS, there must be some resource contention. Therefore, the bandwidth of the serial link was configured to be 500 kbps. Because traffic arrived from a 10 Mbps Ethernet port, congestion started to occur when the input load exceeded about 6%.[6] This number was quite arbitrary. If the WAN bandwidth was changed, the congestion point would also change.

5. The issue is that both data and voice appear as IP frames. Furthermore, RTP does not use well-known port numbers.
6. The congestion point varies depending on the frame size. For example, using 200 byte frames (including header and FCS), the maximum frame rate on 10 Mbps Ethernet is 5682 frames per second. If the load is 6%, the frame rate will be 341 frames per second. Using PPP encapsulation on the WAN link, the bit rate will be 513 Kbps, which is just beyond the capacity of the link. The congestion point will occur sooner if longer frames are used because the overhead will be less.

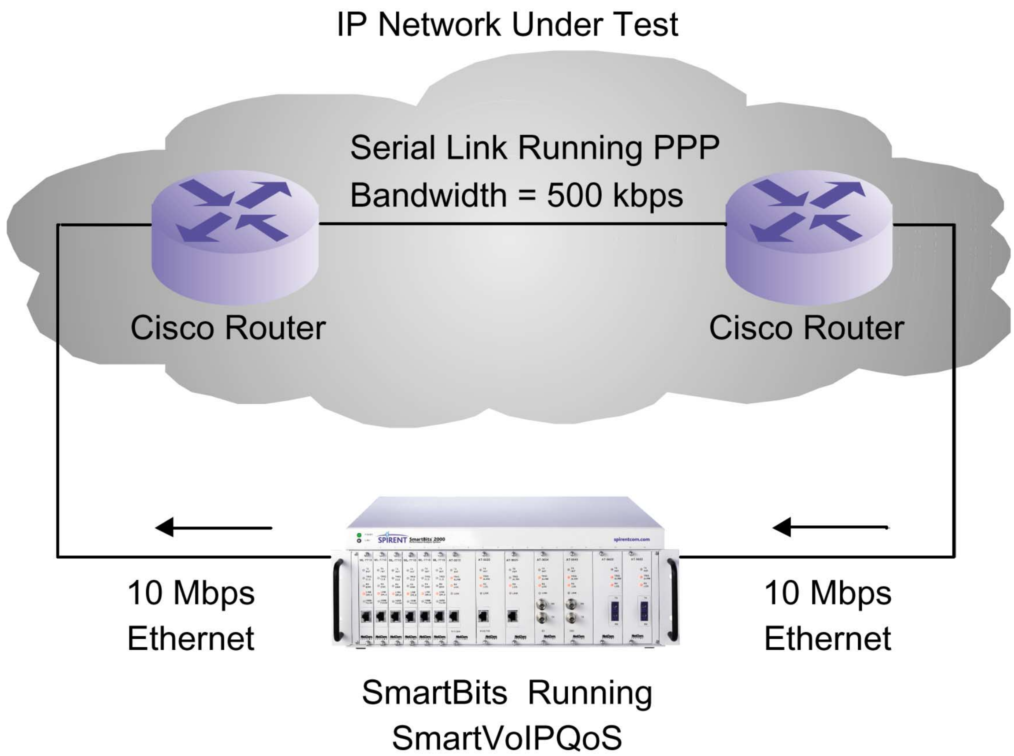


Figure 11. QoS Test Configuration

Priority Queuing

In priority queuing, traffic is classified into separate queues – for example, high, medium, normal, and low. The queues are serviced in the strict order of priority. In other words, the high priority queue must be empty before the medium priority queue is serviced, and both the high and medium priority queues must be empty before the normal queue is serviced and so on. *Figure 12 on page 20* shows the frame loss behavior of such a queuing strategy as the input load is increased.

In this example, when the input load exceeded the WAN bandwidth (at around 6%), frames were discarded. The low priority traffic was discarded until there was none left. Then the normal priority traffic was discarded followed by medium priority traffic. The results illustrate a potential problem with this strategy. If there is a high volume of high priority traffic, then it will squeeze out all traffic of lesser priority.

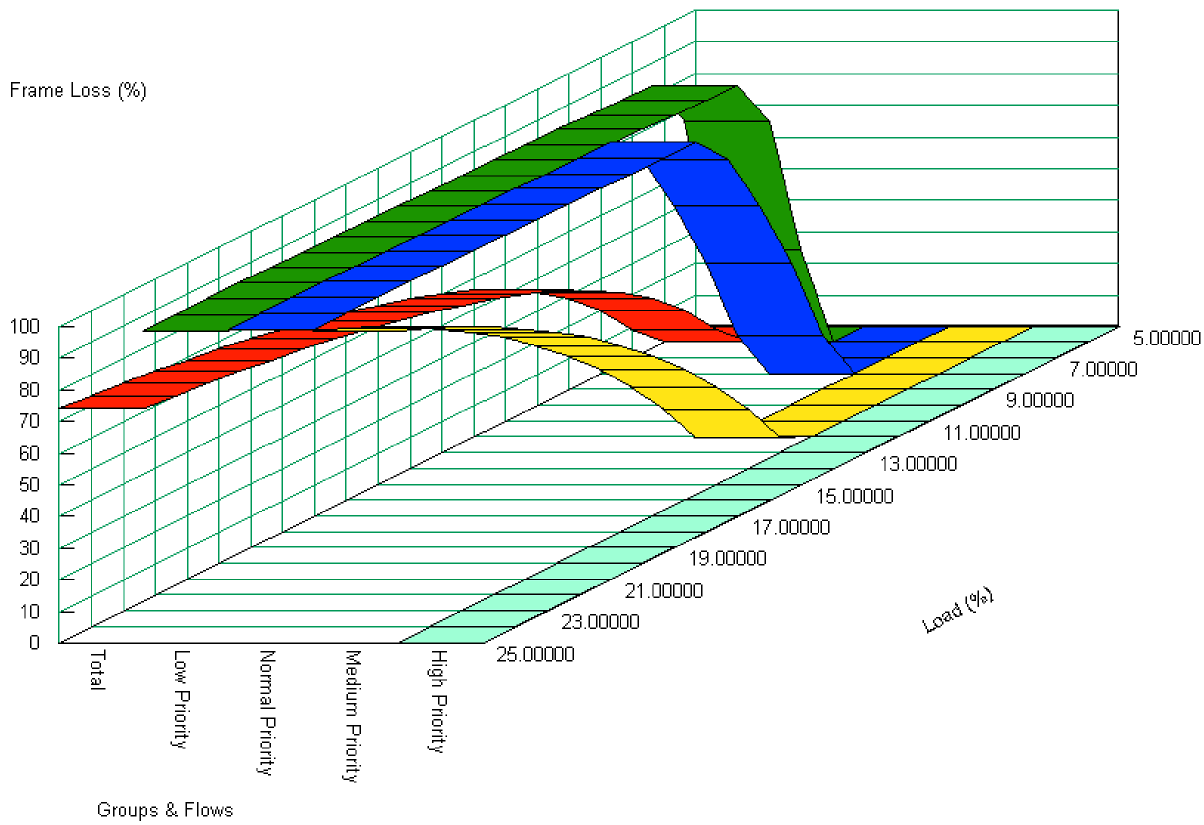


Figure 12. Classical Priority Queuing

The test was repeated again. This time, three traffic flows were configured – one for FTP, one for HTTP, and one for VoIP. Prioritization was done based on frame length. The VoIP traffic was given high priority whereas FTP and HTTP were given low priority. The following table shows the results of the test.

Load	Traffic	Frame Loss (%)	Latency (ms)	Max Latency (ms)	Average Jitter (ms)
2%	FTP	0	25.4	30.5	N/A
	HTTP	0	12.8	15.4	N/A
	VoIP	0	5.8	27.1	4.0
4%	FTP	0	25.1	27.9	N/A
	HTTP	0	12.8	15.2	N/A
	VoIP	0	10.1	27.1	8.2
6%	FTP	2.55	825.0	3097.2	N/A
	HTTP	2.55	683.8	3028.7	N/A
	VoIP	0	43.7	214.3	20.2
8%	FTP	36.97	1680.9	2934.2	N/A
	HTTP	13.82	610.6	2830.3	N/A
	VoIP	0	41.0	200.3	19.2

Congestion started at 6% and 8%, at which point data traffic was discarded while voice traffic was preserved. At the same time, voice traffic also experienced much lower delay compared with data.

Weighted Fair Queuing (WFQ)

Another common queuing strategy in use in the Cisco router is Weighted Fair Queuing. Flow-based WFQ allocates bandwidth based on the IP precedence bits with each traffic flow. For example, if there are eight flows with eight different precedences (0 to 7), then the bandwidth allocation will be 1/36, 2/36, 3/36, …, 8/36 respectively. In other words, the traffic flow with IP precedence 0 will receive 1/36 of the total bandwidth and the traffic flow with IP precedence 7 will receive 8/36 of the bandwidth. *Figure 13 on page 23* shows a graph for the frame loss rate with 8 traffic flows, each with a different IP precedence.

Unlike priority queuing, traffic with a lower priority (lower IP precedence) still gets some bandwidth. This allows more equitable sharing of bandwidth. Furthermore, if there are multiple flows with the same IP precedence, bandwidth is also shared fairly among the flows. Higher priority will typically be given to traffic flows with lower volume (for example, interactive traffic), while traffic flows with higher volume (for example, bulk data transfer) will be given lower priority.

Load	Traffic	Frame Loss (%)	Latency (ms)	Max Latency (ms)	Average Jitter (ms)
2%	FTP	0	25.2	27.1	N/A
	HTTP	0	12.8	15.2	N/A
	VoIP	0	5.9	27.1	4.0
4%	FTP	0	25.1	27.9	N/A
	HTTP	0	12.8	15.2	N/A
	VoIP	0	10.1	27.1	8.2
6%	FTP	0.7	999.5	3187	N/A
	HTTP	1.5	400	3113	N/A
	VoIP	0	26.8	212.5	15.6
8%	FTP	40.7	1759	3152	N/A
	HTTP	7.0	586	2837	N/A
	VoIP	0	33.9	215.5	17.2

The results show that WFQ is equally effective in providing better service to voice over data while providing the additional benefits of fairer sharing of bandwidth.

IP RTP Priority

One of the problems with flow-based WFQ is scalability. The queuing strategy tends to break down when there are a large number of flows. This is evident from the following test. *Figure 13 on page 23* shows the results of WFQ with a total of 8 flows. Instead of one flow per IP precedence, 50 flows were configured per IP precedence. This resulted in a total of 400 flows. In theory, the bandwidth allocation should stay the same, i.e., 1/36 for IP precedence 0, 2/36 for IP precedence 1, and so on. *Figure 14 on page 24* shows the test results. Clearly the scheduling algorithm broke down when there was a large number of flows.

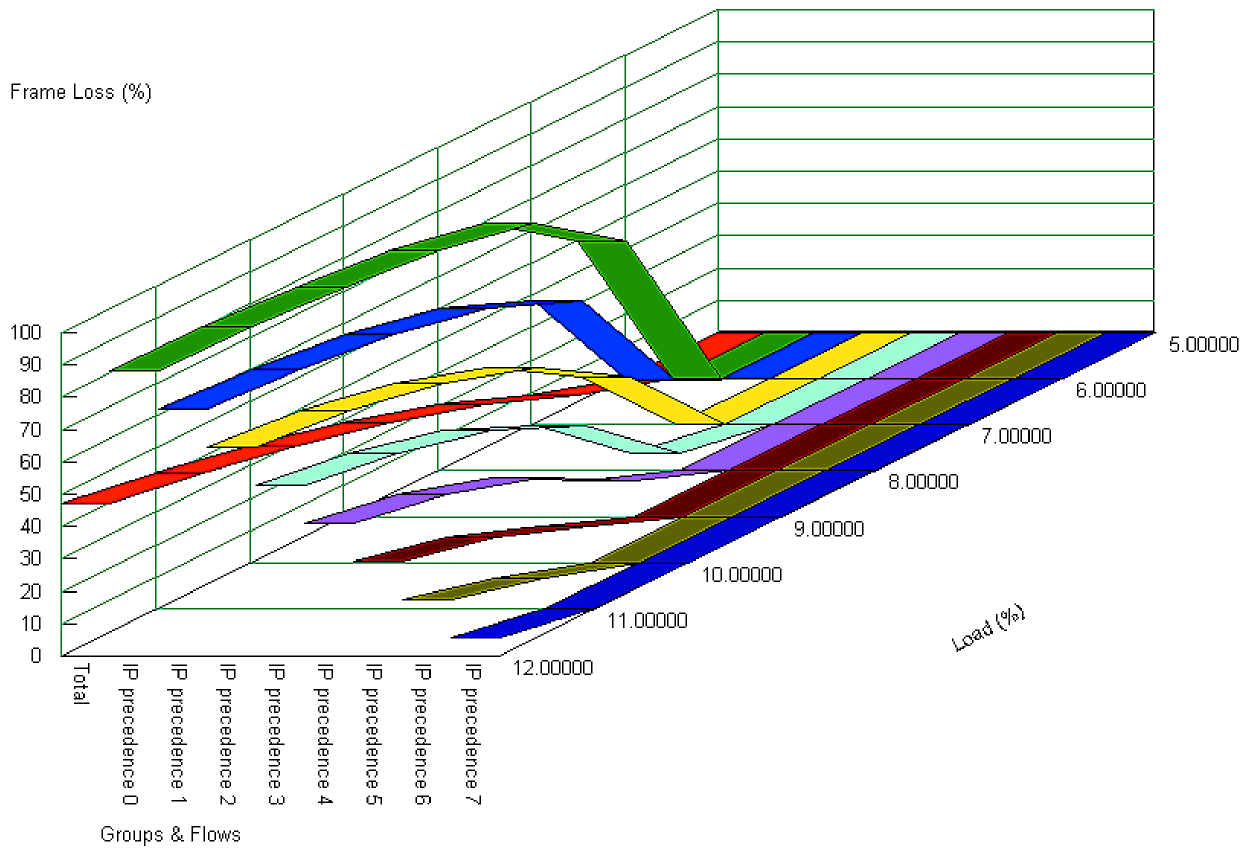


Figure 13. Weighted Fair Queuing

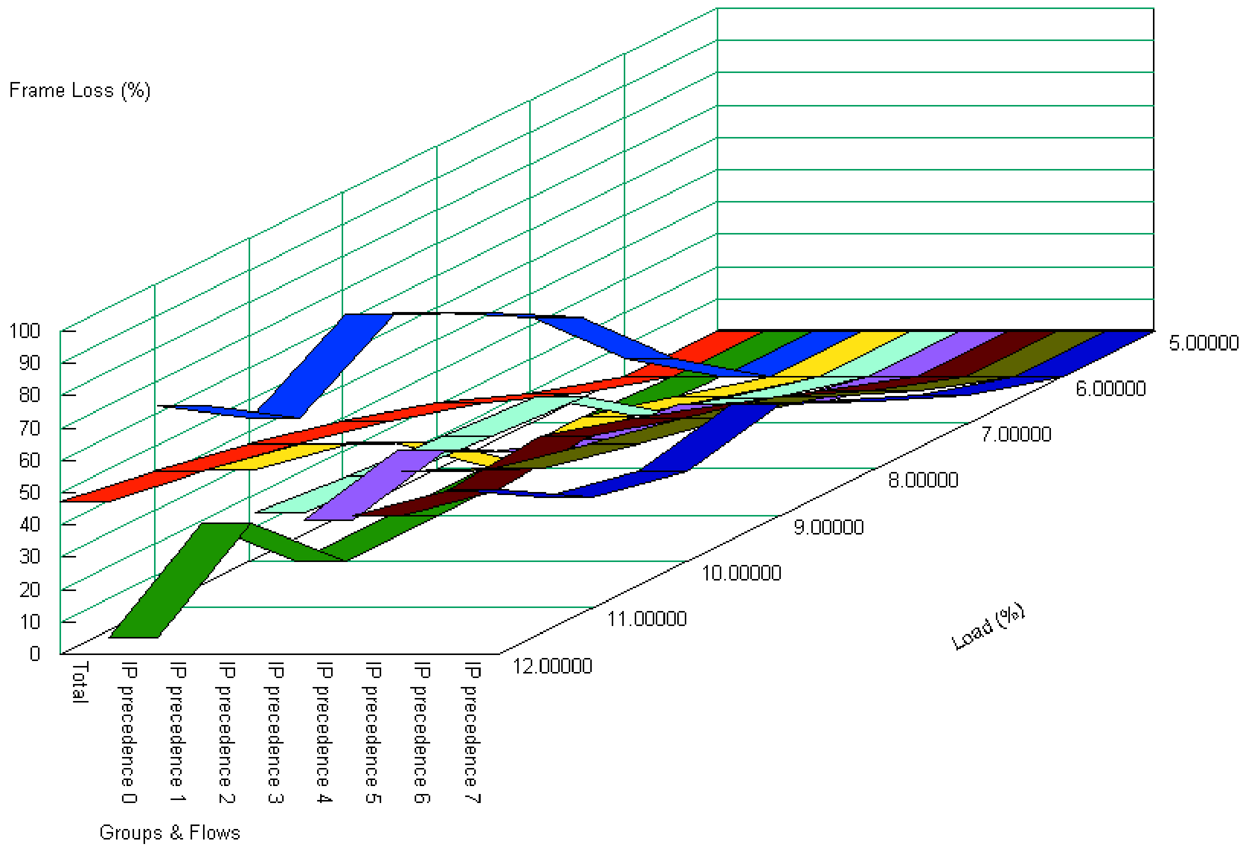


Figure 14. WFQ With 400 Flows

If there is a large number of VoIP flows, IP RTP priority can be used instead. It combines priority queuing with WFQ. VoIP traffic (as a group as opposed to per flow) can be treated with higher priority. This can be done by specifying a UDP port range. All other traffic will be scheduled using WFQ. *Figure 15 on page 25* shows the test results.

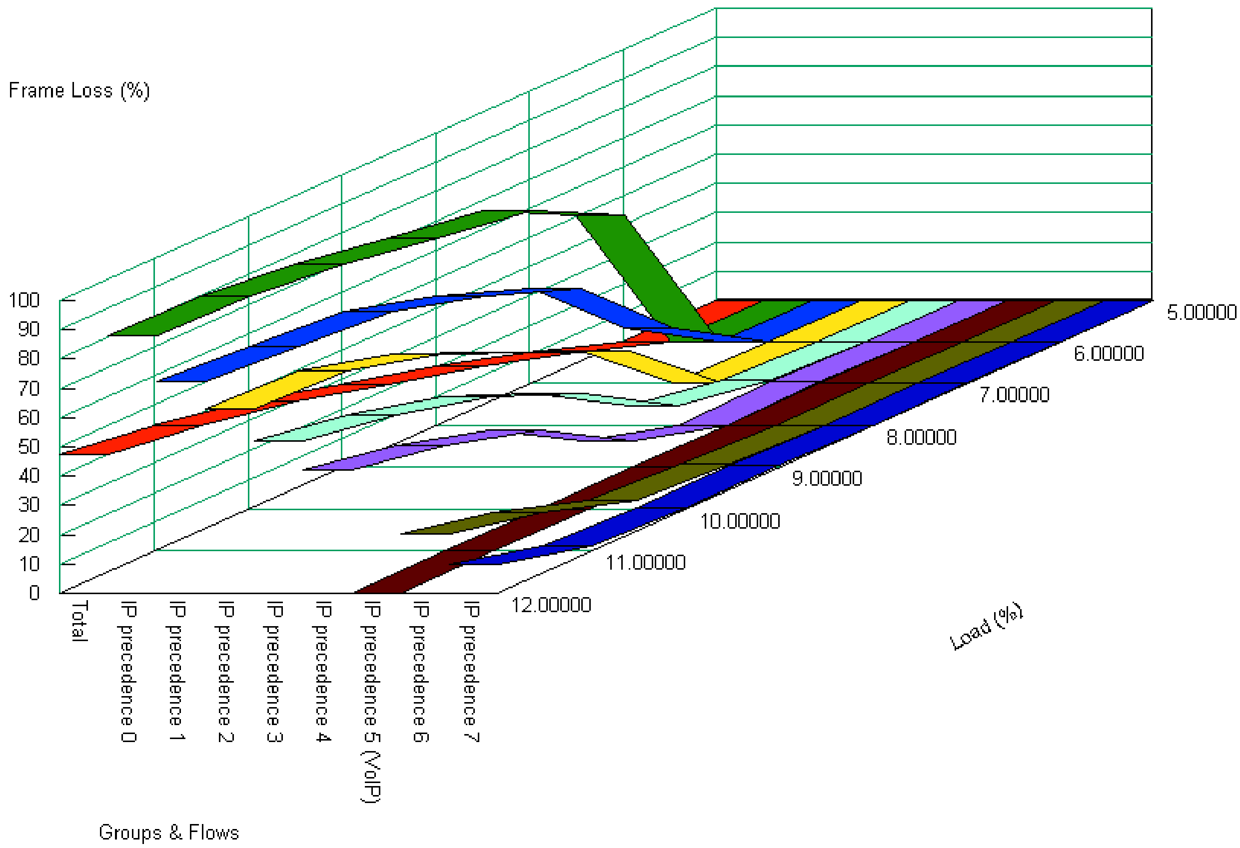


Figure 15. IP RTP Priority

In this case, absolute priority was given to VoIP. However, to prevent VoIP traffic from monopolizing the bandwidth, an upper limit was placed on the bandwidth consumption (although the bandwidth cap was not reached in the test).

Effect of Transport Impairments on Voice Quality

In *"Voice Quality" on page 3*, the transport impairments and their impact on voice quality were discussed. In evaluating voice products, it is highly desirable to measure how well each product can deal with these impairments. This requires creating a test environment that has a controlled set of impairments.

Figure 16 on page 26 shows the test configuration for the following tests. A PC running IP Wave with two Ethernet NICs acted as an "imperfect" router. It passed traffic from one segment to the other with frame loss, latency, and latency variation that could be controlled. The Abacus was used to initiate calls, generate voice input, and measure the quality of the voice output at the other end.

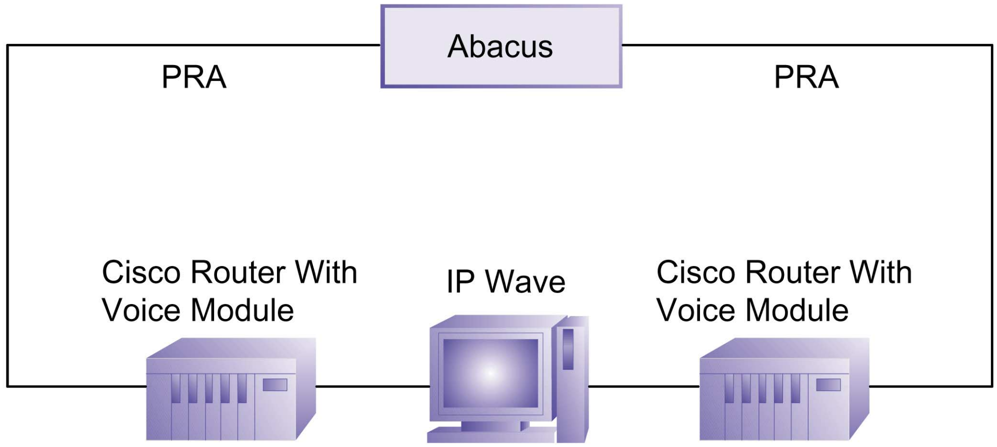


Figure 16. Impairment Configuration

IP Wave was configured with the following frame loss — 0%, 1%, and 3%. G.711 and G.729 were used in the gateway to see how each coding algorithm could cope with the frame loss. The following table summarizes the results.

Coding Algorithm	Frame Loss	PSQM		
		Minimum	Average	Maximum
G.711	0%	0	0.0	1.6
	1%	0	0.4	2.4
	3%	0	1.0	2.9
G.729	0%	0.7	0.8	2.3
	1%	0.7	1.3	2.9
	3%	0.8	1.9	3.0

With 0% frame loss, G.711 had a PSQM score of 0, which represented little or no signal distortion. With G.729, since the bit rate is 8 kbps, it introduced some signal distortion even without frame loss. However, the impact of frame loss on both coding algorithms seemed comparable. A 1% frame loss caused a PSQM increase of 0.4 for G.711 and 0.5 for G.729. Similarly, a 3% frame loss caused PSQM increases of 1.0 and 1.1 respectively for G.711 and G.729.

Figure 17 shows the PSQM values for G.711 with a 3% frame loss.

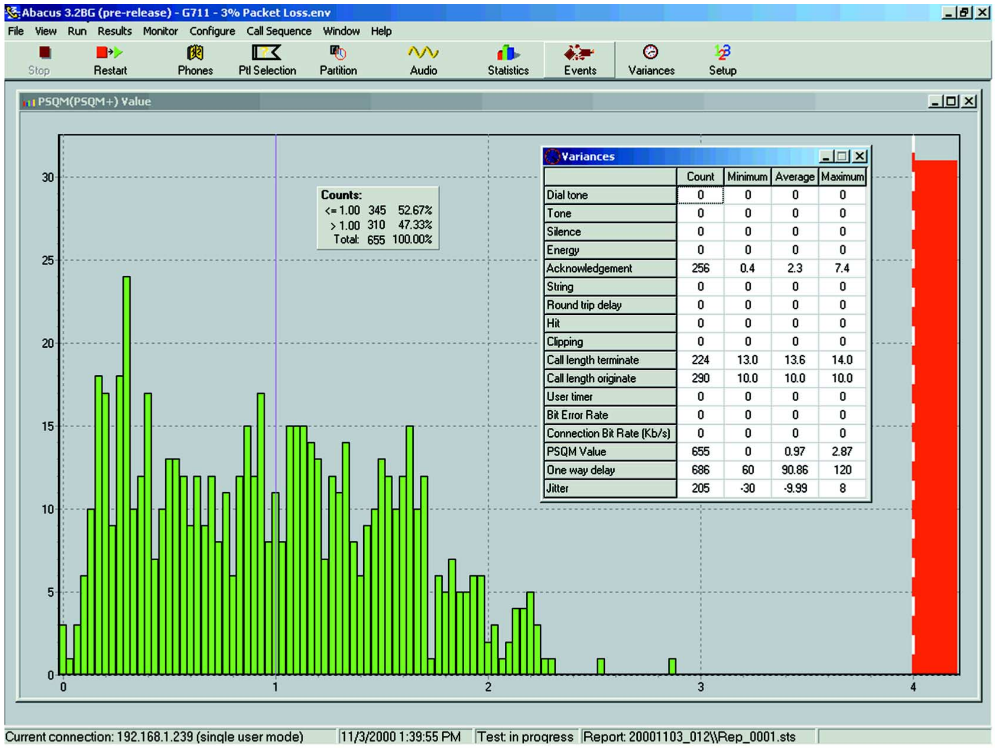


Figure 17. PSQM Values With 3% Frame Loss

The same tests were repeated again. This time, there was no frame loss but various delay impairments were introduced. First, a 50 ms delay was introduced. Then a 50 ms jitter was added and finally a 100 ms jitter was used. The following is a summary of the results:

Coding Algorithm	**Delay Impairments**	**PSQM**		
		Minimum	Average	Maximum
G.711	50 ms	0	0.3	1.3
	50 ms jitter	0	0.2	2.2
	100 ms jitter	0	0.6	2.8
G.729	50 ms	0.7	.08	2.2
	50 ms jitter	0.7	.09	2.7
	100 ms jitter	0.7	1.3	2.8

These test results show that a 50 ms delay and a 50 ms jitter had minimal impact on the PSQM score.[7] The 100 ms jitter caused an increase of 0.3 and 0.5 on G.711 and G.729 respectively. This could be interpreted to indicate that the routers decided to allow some frame loss to not introduce too much delay by fully compensating for the jitter.

Conclusions

VoIP is an IP application that has stringent performance requirements. The performance of the IP network has a direct impact on voice quality. This document identifies the transmission impairment factors that should be measured. These include frame loss rate, delay, and jitter. In particular, QoS is an important component of the IP network. When there is resource contention, such as network congestion, it is important for the network to provide better service to real-time traffic such as VoIP at the expense of data traffic. This document also examines various quality measures including MOS, PSQM, PAMS, PESQ, and the E-Model. All of these measures are useful and can be used in combination.

When testing VoIP, there are different tests that should be performed. These include IP network analysis, end-to-end voice testing, DTMF, fax, and modem testing, impairment simulation, and signaling stress testing. This document discusses the various test configurations.

7. Standard PSQM scoring does not include the effect of delay. However, as previously mentioned, voice quality will degrade if the one-way delay exceeds 150 ms due to talker overlap.

Glossary

Adaptive Differential Pulse Code Modulation (ADPCM)

Process by which analog voice samples are encoded into high-quality digital signals.

ADPCM

See Adaptive Differential Pulse Code Modulation.

Advanced Intelligent Network (AIN)

Telephone network architecture defined by Bell that separates service logic from switching equipment. This allows the addition of services with minimal impact to traffic switches.

AIN

See *Advanced Intelligent Network*.

ANI

Answer Number Indication. Also known as Caller ID. The calling number (number of calling party).

ARQ

Admission request.

Backbone Network

Core high bandwidth links concentrating traffic from access links.

Call Detail Record (CDR)

Description of a call—initiation, duration, services used, termination, and other attributes. Used for billing and traffic management within and between telephone carriers.

CCS

See *Common Channel Signaling*.

CDR

See *Call Detail Record*.

CELP

See *Code-Excited Linear Predictive Coding*.

Central Office (CO)

A telephone company office that connects subscriber local loops within an area to trunk lines. A CO may also connect trunks from other offices.

Channel Associated Signaling (CAS)

A system where signaling information is carried within the bearer channel. Contrasts with AIN.

Circuit-Switched Network

A network where a dedicated physical circuit is established, maintained, and terminated for each communication session. A traditional method for connecting telephone calls.

CLEC

See *Competitive Local Exchange Carrier.*

Coder/Decoder (CODEC)

Hardware or software that converts analogue signals (e.g., video, audio) to or from digital form for transmission or storage. May perform compression or other optimizations such as silence suppression.

Code-Excited Linear Predictive Coding (CELP)

A voice compression algorithm used for low bit-rate voice encoding (i.e., 8 kbps). Used in ITU-T Recommendations G.728, G.729, and G.723.1.

Common Channel Signaling

A signaling system in which control signals for many separate data channels or circuits are carried over a common channel that itself carries no data (e.g., SS7).

Competitive Local Exchange Carrier (CLEC)

A company that builds and operates communication networks in metropolitan areas and provides its customers with an alternative to the local telephone company.

Compression

Reducing the size of a data set to lower the bandwidth or space required for transmission or storage.

Computer Telephony Integration (CTI)

Applications or technology combining telecommunications equipment and services with computer applications. For example, using Caller ID to bring up a client data record from a database.

Conjugate Structure Algebraic Code Excited Linear Prediction (CS-ACELP)

CELP voice compression algorithm providing 8 kbps, or 8:1 compression, standardized in ITU-T Recommendation G.729.

Connectivity

The ability to connect physically and logically between devices to exchange data.

Concatenation

Combining multiple small frames for transmission to reduce overhead due to lower communication layers.

CPE

See *Customer Premises Equipment*.

CPL

Call Processing Language.

CRTP

Compressed Real-Time Transmission Protocol. See *RTP.*

CS-ACELP

See *Conjugate Structure Algebraic Code Excited Linear Prediction*.

CSM

Call switching module.

CTI

See *Computer Telephony Integration*.

Customer Premises Equipment (CPE)

Equipment at the end user's premises (e.g., PC, router, terminal, telephone, etc.); may be provided by the end user or the service provider.

Dedicated Circuit

A transmission circuit leased by one customer for exclusive use all the time. Also called a *private line* or *leased line*.

Delay

In the context of telephony or circuit switching, the amount of time a call spends waiting to be processed. In the context of network transfers, the time to traverse a network or network segment. Differential delay is the difference in transit time between data packets taking separate transmission paths.

Dial Peer

An addressable call endpoint. In VoIP, there are two types of dial peers: POTS and VoIP.

DNIS

Dialed number identification service. The called number.

Digital Signal Processor (DSP)

A high-speed co-processor designed to perform real-time signal manipulation (e.g., conversion between analogue and digital).

DLC

Digital Loop Carrier. A PSTN distribution system with fiber links from the carrier office to a distribution node from which conventional analogue phone loops emanate to individual subscribers.

DNS

See *Domain Name System*.

Domain Name System (DNS)

IETF protocols and services used to map hierarchically structured names to IP addresses. In the context of VoIP, used to convert H.323 IDs, URLs, or other identifiers to IP addresses as well as for locating gatekeepers and gateways.

DS-0

Digital Signal 0. North American Digital Hierarchy signaling standard for transmission at 64 kbps. Also the worldwide standard transmission rate (64 kbps) for PCM digitized voice channels.

DS-1

Digital Signal 1. North American Digital Hierarchy signaling standard for transmissions at 1.544 Mbps. Supports 24 simultaneous DS-0 signals. The term is often used interchangeably with T-1, although DS-1 signals may be exchanged over other transmission systems.

DS-3

Digital Signal 3. North American Digital Hierarchy signaling standard for transmissions at 44 Mbps.

DSP

See *Digital Signal Processor.*

DTMF

See *Dual Tone Multi-Frequency.*

Dual Tone Multi-Frequency (DTMF)

A standard set of tones generated from superimposing two sine waves. Used for telephony signaling (e.g., a touch tone pad).

Dynamic Host Configuration Protocol (DHCP)

IETF protocol to support dynamic allocation of IP addresses to PCs and other hosts to avoid the logistics and overhead of configuring static addresses in each individual machine.

E1

European equivalent of T1 but operating at 2.048 Mbps.

E.164

The international public telecommunications numbering plan. A standard set by ITU-T that addresses telephone numbers.

Ear and Mouth (E and M, E&M) Signaling

Trunk signaling between a PBX and CO used to seize a line, forward digits, release the line, etc.

Endpoint

An H.323 terminal or gateway. An endpoint can call and be called. It generates and/or terminates the information stream.

Echo Cancellation

When transmitting a signal, some of the energy may be reflected back to the transmitter. For some types of full duplex communication, this will interfere with a real signal being sent to the transmitter. A full duplex device can eliminate some of this noise in a received signal by applying a correction signal derived from its transmitted signal.

ETSI

See *European Telecommunications Standards Institute*.

European Telecommunications Standards Institute (ETSI)

European standards body. ETSI ETR-328 is the full rate European ADSL specification.

Foreign Exchange Office (FXO)

A remote Telephone Company Central Office used to provide local telephone service over dedicated circuits from that office to the user's local central office and premises.

Foreign Exchange Station (FXS)

User premises to which a foreign exchange circuit is connected.

FRF.11

Frame Relay Forum implementation agreement for Voice over Frame Relay (v1.0 May 1997). This specification defines multiplexed data, voice, fax, DTMF digit-relay, and CAS/Robbed-bit signaling frame formats, but does not include call setup, routing, or administration facilities.

FRF.12

The FRF.12 Implementation Agreement (also known as FRF.11 Annex C) was developed to allow long data frames to be fragmented into smaller pieces and interleaved with real-time frames. In this way, real-time voice and non real-time data frames can be carried together on lower speed links without causing excessive delay to the real-time traffic.

Frame

A collection of data sent as a unit. Normally used in the context of layer two of the OSI protocol stack.

Frame Loss Rate

The measurement of loss, over time, of data frames as a percentage of the total traffic transmitted.

G.711

Audio codec over 48, 56, and 64 kbps PCM half-duplex channels (normal telephony). Encoded voice is already in the correct format for digital voice delivery in the PSTN or through PBXs; also referred to as “clear channel” coding. Characteristics: high quality, high bandwidth, and minimum processor load.

G.722

Audio codec over 48, 56, and 64 kbps channels.

G.723, G.723.1

Audio codec over 5.3 and 6.3 kbps channels. Selected by the VoIP Forum for use with VoIP. Based on CELP. Characteristics: low quality, low bandwidth, and high processor load due to the compression.

G.726

40/32/24/16 kbps ADPCM codec. Characteristics: good quality, medium bandwidth, and low processor load due to minimal compression.

G.728

Audio codec over 16 kbps channels using LD-CELP. Characteristics: medium quality, medium bandwidth, and very high processor load to greater compression.

G.729, G.729a

Audio codec over 8 kbps channels using CELP. Adopted by the Frame Relay Forum for voice over Frame Relay. Characteristics: medium quality, low bandwidth, and high processor load.

Gatekeeper

An H.323 entity on the LAN that maintains a registry of devices (e.g., H.323 terminals, gateways, and MCUs) to provide address translation services. The devices register with the gatekeeper at startup and request admission to a call from the gatekeeper.

Gateway

In general, a gateway translates between similar services using different protocols to support inter-operation. In the VoIP context, a gateway allows H.323 terminals to communicate with non-H.323 terminals. Different types of gateways may be involved. A Signaling Gateway may be needed to convert IP signaling protocol (H.323, SIP) to PSTN signaling protocols (SS7, ISDN-D). A Media Gateway may be needed to convert IP media protocols (H.323, RTP) to PSTN media protocols (ISDN-B, DS0, DS1). Other types of gateways may be needed to connect to cellular phone networks, etc.

H.225.0

Call Control. An ITU standard that governs H.323 session establishment and packetization. H.225 describes several different protocols: RAS, use of Q.931, use of RTP, and message formats.

H.225.0 RAS

Registration, admission, and status. The RAS signaling function performs registration, admissions, bandwidth changes, status, and disengagement procedures between the VoIP gateway and the gatekeeper.

H.245

An ITU standard that governs H.323 endpoint control, including the opening and closing of channels for media streams, capability negotiation, and more.

H.248

See *MeGaCo* and *MGCP.*

H.261

Video codec for audio visual services at multiples of 64 kbps.

H.263

Specifies a codec for video over the PSTN.

H.323

An ITU-T standard that describes packet-based video, audio, and data conferencing over unreliable networks (i.e., QoS is not guaranteed). H.323 is an umbrella standard that describes the architecture of the conferencing system and refers to a set of other standards (H.245, H.225.0, and Q.931) to describe its actual protocol. H.323 is an extension of ITU standard H.320, which is geared to ISDN. Related standards are:

- H.332: Conferences
- H.235: Security — authentication, encryption, integrity, non-repudiation
- H.246: Interface with PSTN
- H.450.1, .2, .3: Supplementary services, call transfer, call diversion
- H.261, H.263, ...: Video
- H.320, H.321, H.324: ISDN, PRI/ATM, PSTN analogue

H.323 Terminal

Network nodes that provide real-time, two-way communications with another H.323 terminal (e.g., computer-based video conferencing systems).

IA 1.0

VoIP Forum Implementation Agreement 1.0 selecting protocol options for interoperable VoIP.

IEEE

See *Institute of Electrical and Electronic Engineers*.

IETF

See *Internet Engineering Task Force*.

IGMP

See *Internet Group Management Protocol*.

Institute of Electrical and Electronic Engineers (IEEE)

A voluntary organization which, among other things, sponsors standards committees and is accredited by the American National Standards Institute. IEEE Project 802 drives many network standards oriented to the physical and logical link layers.

Interactive Voice Response (IVR)

A voice prompted telephony service that prompts users through a series of choices via a touchtone keypad.

International Electrotechnical Commission (IEC)

An international standards body.

International Organization for Standardization (ISO)

An international standards body, commonly known as the *International Standards Organization*. ISO is known for its seven layer OSI model of tiered communication systems.

International Telecommunications Union Telecommunications Standards Sector (ITU-TSS)

The new name for CCITT. An international standards body that is a committee of the ITU; a UN treaty organization.

Internet Engineering Task Force (IETF)

An open standards organization driving the Internet RFC(Request For Comment) process. The IPCDN (IP over Cable Data Network) working group has produced various RFCs dealing with management SNMP MIBs (Simple Network Management Protocol Management Information Base) in support of DOCSIS.

Internet Group Management Protocol (IGMP)

A network-layer protocol for managing multicast groups on the Internet.

Interexchange Carrier (IXC) Interexchange Common Carrier

A long-distance telephone company offering circuit-switched, leased-line, or packet-switched service or some combination of these services.

Internet

(Note the capital "I.") The largest internet in the world consisting of large national backbone nets (such as MILNET, NSFNET, and CREN) and a myriad of regional and local campus networks all over the world. The Internet uses the Internet protocol suite. To be on the Internet, you must have IP connectivity (i.e., be able to Telnet to or ping other systems). Networks with only e-mail connectivity are not actually classified as being on the Internet.

Internet Protocol (IP, IPv4, IPv6)

A Layer 3 (network layer) protocol that contains addressing information and some control information that allows packets to be routed. Documented in RFC 791. Comprises many other protocols, notably, IP multicast and various routing protocols. IPv4 supports 32 bit addresses and is predominant. IPv6 uses 128 bit addresses and is expected to eventually displace IPv4.

Internet Service Provider (ISP)

A company that provides users and companies with a connection to the Internet.

Internet Telephony

A generic term used to describe various approaches to running voice telephony over IP.

Internet Telephony Service Provider (ITSP)

A company that provides users with Telephony Services and applications via the Internet.

Internetwork

A collection of networks interconnected by routers that function (generally) as a single network. Sometimes called an *internet*, which is not to be confused with the *Internet*.

Intranet

A private network inside a company or organization that uses the same kinds of software that you would find on the public Internet, but that is only for internal use. As the Internet has become more popular, many of the tools used on the Internet are being used in private networks. For example, many companies have Web servers that are available only to employees.

IP, IPv4, IPv6

See *Internet Protocol*.

IPDC

IP Device Control (family of protocols, IETF work in progress, see also MGCP).

ITU

See *International Telecommunications Union*.

IVR

See *Interactive Voice Response*.

IXC

See *Interexchange Carrier.*

Latency

The delay between the time when a device receives a frame and when the frame is forwarded out of the destination port.

LDAP

Lightweight Directory Access Protocol. An Internet standard for accessing Internet directory services.

LDCELP

See *Low-delay CELP.*

LEC

See *Local Exchange Carrier.*

Lifeline POTS

A minimal telephone service designed to extend a “lifeline” to the telephone system in case of emergency, particularly when electric power is lost.

LLC

See *Logical Link Control (LLC).*

LNP

Local Number Portability.

Local Exchange Carrier (LEC)

See *Competitive Local Exchange Carrier.*

Loop

Twisted-pair copper telephone line connecting from the PSTN to a client’s premises. Loops may differ in distance, diameter, age, and transmission characteristics depending on the network.

Logical Link Control (LLC)

The LLC layer is the upper sub-layer of the OSI Data Link layer. It controls the assembling of data link layer frames and their exchange between data stations, independent of how the transmission medium is shared.

Low-Delay CELP (LDCELP)

CELP voice compression algorithm providing 16 kbps, or 4:1 compression. See *ITU-T G.728*.

MCU

See *Multipoint Control Unit*.

Mean Opinion Score (MOS)

A system of grading the voice quality of telephone connections. The MOS is a statistical measurement of voice quality, derived from a large number of subscribers judging the quality of the connection.

Media Gateway Control Protocol

A protocol used by central controllers to monitor events and manage terminal units and gateways. The objective is to separate signaling/call control from voice traffic to facilitate service/feature upgrades by upgrading a central controller rather than all end devices and gateways. Intended for carrier-grade scalability. Defined in IETF RFC 2705. Like H.323, it is comprised of agents and a gatekeeper. Unlike H.323 and SIP, end systems are controlled by a network server; hence, there is no direct peer-to-peer connection.

MeGaCo IETF Working Group

Responsible for MGCP definition and evolution. Began with IPDC from Level3, Lucent, and others. Also involves SGCP (Simple Gateway Control Protocol) from Telcordia/ Bellcore and Cisco. MGCP resulted from combining SGCP and IPDC protocols to create RFC 2705. It was followed by MDCP (Media Device Control Protocol from Lucent) and ITU work to modularize H.323 and define inter-module protocols. The merging of MGCP, MDCP, and ITU to produce MeGaCo protocol (also known as *H.GCP* and *H.248*), has been submitted to ITU for approval.

MGCP

See *Media Gateway Control Protocol*.

MIB

See *Management Information Base*.

MOS

See *Mean Opinion Score*.

Moving Picture Experts Group (MPEG)

A voluntary body that develops standards for digital compressed moving pictures and associated audio.

Multicast

A process of transmitting PDUs from one source to many destinations. The actual mechanism for this process might be different for different LAN and WAN technologies.

Multipoint Control Unit (MCU)

In VoIP, an MCU is an endpoint that supports three or more terminals and gateways participating in a multipoint conference.

Multipoint-Unicast

A process of transferring PDUs (Protocol Data Units) where an endpoint sends more than one copy of a media stream to different endpoints. This might be necessary in networks that do not support multicast.

NAT

Network Address Translation.

Network Time Protocol (NTP)

A protocol built on top of TCP that assures accurate local time-keeping with reference to radio and atomic clocks located on the Internet. This protocol is capable of synchronizing distributed clocks within milliseconds over long time periods.

Node

An H.323 entity that uses RAS to communicate with the gatekeeper. For example, an endpoint may be a terminal, proxy, or a gateway.

NTP

See *Network Time Protocol*.

Off-Hook

The active condition of a Switched Access or Telephone Exchange Service line.

OH

See *Overhead*.

On-Hook

The idle condition of a Switched Access or Telephone Exchange Service line.

Overhead

(OH) Bits in a frame or cell that are required for framing, CRC, routing, etc.

P.50

ITU-T Recommendation (1993), Artificial voices.

P.56

ITU-T Recommendation (1993), Objective measurement of active speech level.

P.501

ITU-T Recommendation (1996), Test signals for use in telephonometry.

P.561

ITU-T Recommendation (1996), In-service, non-intrusive measurement devices for voice service measurement.

P.800

ITU-T Recommendation (1996), Methods for the subjective determination of transmission quality.

P.830

ITU-T Recommendation (1996), Subjective performance assessment of telephone-band and wideband digital codecs.

P.861

ITU-T Recommendation (1996), Objective quality measurement of telephone-band (300 - 3400 Hz) speech codecs.

Packet

A collection of data sent as a unit. Normally used in the context of layer three of the OSI protocol stack. A packet may be fragmented and sent in multiple frames as required by the underlying layer two facilities.

Packet Loss Rate

The measurement of loss, over time, of data packets as a percentage of the total traffic transmitted.

Packet Switching

A WAN switching method in which network devices share a single point-to-point link to transport packets from a source to a destination across a carrier network.

PacketCable

An MCNS and Cable Labs initiative principally intended to carry packetized voice and fax over DOCSISTM capable cable systems. Services include voice mail, call placement, call management, PSTN interfaces (SS7), and other functions common to traditional voice carriers.

PBX

See *Private Branch Exchange*.

PCM

See *Pulse Code Modulation*.

PDU

See *Protocol Data Unit*.

PHS

See *Payload Header Suppression*.

Plain Old Telephone System (POTS)

Traditional analogue telephone service that uses voice bands. Sometimes used as a descriptor for all voice-band services.

Point to Point Protocol (PPP)

A protocol used to encapsulate various network protocols, typically to interconnect two networks or a remote user and a network, over a link or circuit that is accessible to only two parties (typically a serial link).

Private Branch Exchange (PBX)

A small telephone network for customer premises. Provides local connectivity, switching, and connections to the wide area voice network.

Proxy

In the general sense, a proxy is an agent that performs operations on behalf of another entity. In the context of VoIP, proxies are special gateways that relay one H.323 session to another.

PSTN

See *Public Switched Telephone Network*.

Public Switched Telephone Network

An umbrella term that represents the carriers that make up the worldwide telephone services. See also *POTS*.

Pulse Code Modulation (PCM)

The transmission of analog information in digital form through the sampling and encoding of samples with a fixed number of bits.

Q.931

An ITU standard that describes ISDN call signaling and setup. The H.225.0 standard uses a variant of Q.931 encapsulated within TCP to establish and disconnect H.323 sessions.

QoS

See *Quality of Service*.

QSIG

A signaling system between a PBX and CO, or between PBXs used to support enhanced features such as forwarding and follow me.

Quality of Service (QoS)

In communications, an umbrella term that refers to the application of constraints to favor certain types of traffic and, potentially in some contexts, ensure a given level of service quality and availability. Typically intended to constrain errors of latency or jitter while ensuring a set bandwidth.

RAS

Registration Authentication Status. A specification within H.323 that allows for session authentication and authorization. This is what validates the call. See *H.225.0*.

RRQ

Registration request.

Remote Switching Module (RSM)

The term *Central Office* designates the combination of the Remote Switching Unit and its Host.

Real-Time Transport Protocol (RTP)

Real-time Transport Protocol, IETF RFC1889. A real-time, end-to-end streaming protocol utilizing existing transport layers for data that has real-time properties. RTP provides for payload type identification, sequence numbering, time-stamping, and delivery monitoring of real-time applications. The H.225.0 standard describes how to use RTP to handle the packetization of video and audio in H.323.

RTP Control Protocol (RTCP)

IETF RFC1889. A protocol to monitor the QoS and to convey information about the participants in an ongoing session; provides feedback on total performance and quality so that modifications can be made. Monitors RTP connections including timing reconstruction, loss detection, security, and content identification. RTCP provides support for real-time conferencing for large groups, including source identification and support for gateways (like audio and video bridges) and multicast-to-unicast translators.

Resource Reservation Protocol (RSVP)

IETF RFC2205-2209. A general purpose signaling protocol that allows network resources to be reserved for a connectionless data stream, based on receiver-controlled requests. Applications running on IP end systems can use RSVP to indicate to other nodes the nature (bandwidth, jitter, maximum burst, and so on) of the packet streams that they want to receive.

RSVP

See *Resource Reservation Protocol*.

RTCP

See *RTP Control Protocol*.

RTP

See *Real-Time Transport Protocol*.

RTSP

Real-Time Streaming Protocol. A protocol used to interface to a server that will provide real-time data.

SAP

Session Announcement Protocol. A protocol used by multicast session managers to distribute a multicast session description to a large group of recipients.

SDP

See *Session Description Protocol*.

Session Description Protocol (SDP)

RFC 2327. Describes the data payload for SAP, SIP, and RTSP sessions. It is text-based for easy processing and extensibility. Describes media stream type and number (e.g., audio + video), and is used to set up H.323, Internet radio, game session, chat, etc. Defines originator, unicast, multicast, broadcast destination, and UDP port numbers. Defines features for the session (e.g., codecs, call control capabilities). Defines scheduling, beginning, end times, and repetitions for broadcasts.

Session Initiation Protocol (SIP)

RFC 2543. Used to set up a unicast session between two endpoints. This is a text-based format inspired by HTTP that is much simpler than H.323. It runs over any transport protocol (e.g., UDP, TCP, ATM AAL5, IPX, X.25).

SGCP

Simple Gateway Control Protocol. A simple UDP-based protocol for managing endpoints and connections between endpoints.

SIP

See *Session Initiation Protocol*.

SS7

Signaling System 7. A standard CCS system used with BISDN and ISDN that was developed by Bellcore. SS7 is a packet signaling network that runs parallel to the circuit-switched network that transports actual voice traffic. It is used for control / management—setup, teardown, control calls, etc. Control traffic is out-of-band (i.e., on separate lines, but within the same devices, which makes it less prone to problems of congestion since it avoids heavy traffic).

T1

One implementation of DS-1 services utilizing 4 wires and Bipolar Alternate Mark Inversion (AMI) encoding. Requires repeaters every 6,000 ft.

T.120

An ITU standard that describes data conferencing H.323. It enables the establishment of T.120 data sessions inside of an existing H.323 session.

TCP, UDP

Internet standard Transport Layer protocols.

Time Division Multiplexing (TDM)

Multiple data streams, possibly in different directions, sharing a physical medium by isolation within reserved time intervals. Bandwidth is allocated to each channel regardless of whether the station has data to transmit.

TOS

See *Type of Service*.

Transmission Control Protocol (TCP)

A transport-layer Internet protocol that ensures successful end-to-end delivery of data packets without error.

Type of Service (TOS)

One byte of an IP datagram reserved for qualifying the desired QoS. Three bits define the IP precedence (priority). One bit requests “low delay.” One bit requests “high throughput.” One bit requests “high reliability.” Two bits are unused. The precedence bits can be used by a router or switch to favor traffic when congestion is evident.

VAD

See *Voice Activity Detection*.

Virtual Private Network

A collection of nodes within a larger physical network that are connected in such a way that they appear to be on a separate isolated network. Traffic into and out of a VPN requires explicit routing. VPNs are often encrypted to protect data.

Voice Activity Detection (VAD)

Saves bandwidth by transmitting voice cells only when voice activity is detected (i.e., silence is not encoded and transmitted over the network). Sound quality is slightly degraded, but the connection uses much less bandwidth.

VoIP

See *Voice over IP*.

Voice over IP (VoIP)

An umbrella term for the set of standards emerging to support voice services over packet-based IP networks.

Voice Band

The frequency range from 0 to 4kHz used for analogue (voice, fax, data) signals in conventional POTS.

VPN

See *Virtual Private Network*.

VTSP

Voice telephony service provider.

WEPD

Weighted Early Packet Discard.

WFQ

Weighted fair queuing. Congestion management algorithm that identifies conversations (in the form of traffic streams), separates packets that belong to each conversation, and ensures that capacity is shared fairly between these individual conversations. WFQ is an automatic way of stabilizing network behavior during congestion and results in increased performance and reduced retransmission.

WRED

Weighted Random Early Detection.

Zone

A collection of all terminals, gateways, and Multipoint Control Units managed by a single gatekeeper. A zone includes at least one terminal and may or may not include gateways or MCUs. A zone has only one gatekeeper. A zone may be independent of LAN topology and may be comprised of multiple LAN segments that are connected using routes or other devices.

References

ITU-T Recommendation G.103 (2/1996), Transmission impairments.

ITU-T Recommendation G.103 Appendix I (9/1999), Provisional planning values for the equipment impairment factor Ie.

ITU-T Recommendation G.107 (5/2000), The E-Model, a computational model for use in transmission planning.

ITU-T Recommendation G.114 (2/1996), One-way transmission time.

ITU-T Recommendation G.168 (4/1997), Digital network echo cancellers.

CCITT Recommendation G.711 (1988), Pulse Code Modulation (PCM) of voice frequencies.

ITU-T Recommendation G.711 Appendix I (9/1999), Appendix I: A high quality low-complexity algorithm for packet loss concealment with G.711.

ITU-T Recommendation G.723.1 (1996), Speech coders: Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbps.

CCITT Recommendation G.726 (1990), 40, 32, 24, 16 kbps Adaptive Differential Pulse Code Modulation (ADPCM).

CCITT Recommendation G.728 (1992), Coding of speech at 16 kbps using Low Delay Code Excited Linear Prediction (LD-CELP).

ITU-T Recommendation G.729 (3/1996), Coding of speech at 8 kbps using Conjugate-Structure Algebraic Code Excited Linear Prediction (CS-ACELP).

ITU-T Recommendation H.323 (9/1999), Packet-based multimedia communications systems.

ITU-T Recommendation P.800 (8/1996), Methods for subjective determination of transmission quality.

ITU-T Recommendation P.861 (2/1998), Objective quality measurement of telephone-band (300-3400 Hz) speech codecs.

IETF RFC 1889 (1/1996), RTP: A Transport Protocol for Real-Time Applications.

IETF RFC 1890 (1/1996), RTP Profile for Audio and Video Conferences with Minimal Control.

IETF RFC 2327 (4/1998), SDP: Session Description Protocol.

IETF RFC 2543 (3/1999), SIP: Session Initiation Protocol.

IETF RFC 2705 (10/1999), Media Gateway Control Protocol (MGCP) Version 1.0.

About the Author

Angus Ma (B.Eng., M.Eng., M.B.A.)

Mr. Angus Ma began his career as a software designer for Nortel Networks (formerly Bell-Northern Research). After leaving Nortel, he developed data communications products as well as UNIX-based office systems. In 1986, Mr. Ma launched AHM Technology Corporation, which provides network design, analysis, and troubleshooting services to large corporate clients. Angus is an internationally-known speaker appearing regularly in North America, Europe, and Asia and is a technical editor and author for Learning Tree International. Mr. Ma has worked in data and telecommunications since 1980 and has extensive experience in planning, implementing, maintaining, and analyzing enterprise networks.

Acknowledgements

The author would like to thank Spirent Communications for their support during testing, in particular, Mr. Iain Milnes and Mr. Andrez Chavez for running the PSQM tests.

SPIRENT
Communications