IP Telephony Service Transparency Test

Renshou Dai

December 29, 2004

1 Introduction

At the time of writing this paper, VoIP (Voice-over-IP) is no longer a buzz word used only by professionals. The concept is widely disseminated into the general public as well. But the concept of MoIP (Modem-over-IP) is entirely a different story, even to professionals working in the data and telecom industry.

Carrying a real-time, continuous interactive voice streams across a packet-switched network is not as a trivial problem as was originally envisioned. Voice quality problems related to VoIP have been well known and well tested in the last few years. Most notable problems of VoIP (from user perspective) are degraded clarity (as a result of vocoder compression, packet loss and voice jitter etc), increased latency (delay) and more noticeable echo problems (exacerbated by the longer delay).

Much of the attention has been focused in the past on the well-publicized voice quality problems. The service transparency and MoIP issue is a much less-known and less-discussed problem. There might be two reasons for this. The first one is, if we can't even improve the network to a level that will appease the "dumb" human ears, how can we expect the packet-network to handle the much stringent voice-band modem signals that can't tolerate any amount of packet loss, jitter and compression? The second reason is, if a user already has broadband digital connection to a home, why would he/she still need to transmit and receive the legacy FAX and modem signals through the telephone wire?

The counter argument to the first reason is this. Transporting voice-band modem signals across the packet network may in fact be easier than VoIP, if implemented properly. One way is by modem relay. In this way, the modem counterparts are built into the gateways and IADs. At the ingress of a gateway, the voice-band modem signals are demodulated and the bits content are transported as a special packet. At the egress of the receiving gateway, the bits are remodulated back into the original voice-band signals. This is how ITU-T T.38 FoIP (Fax-over-IP) works. Despite its theoretical elegance, the practical implementation of this scheme poses a serious burden to the equipment designers. Designing tens or even hundreds of different legacy modem types into a highly-integrated and cost-sensitive embedded device such as an IAD is economically prohibitive, if not technically impossible. A more sensible scheme is by voice-band data. Keep in mind that the fundamental challenge of VoIP or MoIP is to balance the conflicting requirements between shorter delay and fewer packet losses. In VoIP, the focus should be on shortening the delay and permitting certain amount of packet loss and voice jitter (sudden delay variation as a result of jitter buffer resizing inside the IAD or gateway). For MoIP, the long delay is not an issue, but not any amount of packet loss or jitter is permitted. In voice-band data mode, the first task of a gateway or IAD is call discrimination. The gateway must decide if a call is modem/data call or voice call by detecting the precursor signature tone(s). If in modem call mode, the gateway should employ the largest static jitter buffer size as possible (this will increase the delay, but no harm generally to modem signals) at the receiving side, and disable all voice-processing features such as low-bit-rate vocoder compression, silence suppression or voice activity detection and echo cancellation etc at the transmitting side. With large jitter buffer size, the risk of packet loss is minimized. With static jitter buffer size, the voice jitter is eliminated. By disabling all the "voice-enhancing and bandwidth-conserving" features except the simple G711 PCM companding, the modem signal is handled in the same way as it is handled by the conventional PSTN network.

As for why people still need voice-band modem service since they already have broadband digital connection, the answer is simply backward compatibility, some regulatory requirements and convenience to the users. When a user subscribes to a VoIP service, he/she generally does not want to be restricted to voice call only. He/she may also expect the "legacy" simple data service to be available too. Ideally, no matter what transport or switching (circuit-switched PSTN or packet-switched IP network) technology is used inside the network, a user should not see any difference in terms of service. The burden of assuring such service transparency should be the concern of the service providers and equipment vendors. It should not be turned into a dilemma for a user.

Sage Instruments designed this IP Telephony Service Transparency Test (abbreviated as IP-TSTT from here on) to specifically verify a gateway or IAD's ability to handle a MoIP call or a VoIP call. More specifically, the test sends standard-based precursor signature tones by emulating a FAX call, a V34/V90 modem call or a voice call to test the gateway's call discrimination capability. After the signature tones, the test then measures the round-trip delay as the delay is a good indicator of the internal jitter buffer size. The test then starts simplex, full-duplex or half-duplex packet network impairments test by measuring in detail each individual packet loss event and its duration, and each individual voice jitter (gapping or jerking) event and its duration. The signal level change (gain) is also measured, and the codec type (PCM, ADPCM or VOCODER) is also detected. With this test, a user can quickly determine whether or not the device and network under test can handle various MoIP or VoIP calls.

2 Overview of IP-TSTT

2.1 Director and responder format

Like most of Sage's other tests such as SMOS [1], this IP-TSTT comes in an automated director and responder format for the ease of use. After connecting to the device and network under test, the "dumb" responder simply waits for an incoming call from the director to start a test. The detailed test parameters are sent over from the director using slow but reliable in-band telemetry. The responder is generally placed in an "inconspicuous" location and does not need any intervention from the user. The director end is under the direct control of the user. After selecting the user options, a user starts the test at the director end. The director calls the responder, and the test proceeds automatically. The test results measured at the responder end are sent back to the director end with in-band telemetry that is designed to be slow but reliable in case of severe packet impairments. The test results are normally presented at the director end only.

The director to responder direction is defined as the forward direction. The responder to director direction is defined as the reverse direction.

2.2 IP-TSTT measurements

The IP-TSTT measures the following parameters:

1. Round-trip delay: [0,8000] ms, ± 1 ms.
2. Codec type: reported as PCM, ADPCM or VOCODER.
3. Gain: [-80,40] dB, ± 1 dB.
4. Number of packet loss events: [0,4095].
5. Total accumulated packet loss duration: $[0{,}4 \times 10^9]$ ms.
6. Number of compressive jitter events: [0,511].
7. Total accumulated compressive jitter duration: $[0{,}4 \times 10^9]$ ms.
8. Number of expansive jitter events: [0,511].
9. Total accumulated expansive jitter duration: $[0{,}4 \times 10^9]$ ms.
10. Total measurement duration $[0{,}4 \times 10^9]$ ms.

A packet loss event occurs when the receiving-end gateway plays out a substituted packet in place of a packet whose arrival time exceeds the maximum jitter buffer size. The substituted packet is either a silence, a previous packet or some form of interpolation if a future packet is available. No matter what, IP-TSTT will register such event, and accumulate the total packet loss duration. With PCM or ADPCM codecs, the measured packet loss duration will be close to the actual packet size within ± 1 ms. With low-bit-rate vocoders, each individual packet loss duration measurement precision is ± 5 ms.

A voice jitter event occurs when the receiving-end gateway dynamically readjusts its jitter buffer size, therefor causing a sudden delay variation at the audio signal side. When the gateway decreases the jitter buffer size, a compressive jitter event (sudden shortening of delay) will be measured and its duration is accumulated. When the gateway increases the jitter buffer size, an expansive jitter event (sudden elongation of delay) will be registered and its duration is accumulated. The measurement precision for each individual jitter event is ± 1 ms.

The packet loss and jitter are commonly expressed in a percentage format. These percentage amounts are calculated by dividing the total accumulated packet loss or jitter durations by the total measurement duration. Such percentage numbers will be presented to users.

3 The three testing modes

When operating the IP-TSTT, a user needs to specify the following test parameters:

1. Test duration: [10,4095] seconds.
2. Test signal level: [-30,0] dBm.
3. Test mode: Fax-modem, V34/V90-modem, or Voice-mode.

The three test modes deserve further detailed explanations.

3.1 Fax modem simplex mode

If the Fax-modem mode is chosen, this IP-TSTT will emulate a Fax call by sending the unique Fax call precursor tones as specified in ITU-T T.30 [2]. More specifically, after calling the responder, the director sends the CNG (calling tone), which is simply a 1100Hz tone with a cadence of 0.5 second ON and 3 seconds OFF. After answering the call and detecting the CNG tone, the responder sends the CED (called terminal identification) answer tone, which is just a continuous 2100Hz tone for 3.6 seconds. The director stops sending CNG after detecting the end of the CED from responder. Either the CNG or the CED tone should trigger the device under test into the Fax mode. After these tones, the IP-TSTT then proceeds with following test sequence:

1. Perform coordinated round-trip delay measurement between the director and responder.

2. Director transmits test parameters to responder via in-band telemetry.

3. Perform simplex packet network impairment test in the forward direction only since a Fax page is generally transmitted one way only. The director transmits the unique test signal, and the responder performs packet loss, jitter, gain measurements and codec detection with this unique signal. The reverse direction is quiet.

4. After reaching the specified test duration, the responder sends back measurement results via in-band telemetry, and then disconnects the call.

5. After receiving the results, the director then presents the results and disconnects the call.

3.2 V34/V90 modem full duplex mode

If the V34/V90 modem mode is chosen, this IP-TSTT will emulate a typical computer modem call by sending the unique precursor signals described in ITU-T V.8 [3]. More specifically, after calling, the director sends the CI (call indicator signal), which is an FSK modulated signal (per ITU-T V.21 [4]) of certain bit sequence with a cadence of 1 s ON and 1 s OFF. After answering the call and detecting the CI signal, the responder sends the ANSam (modified answer tone) for 3.6 seconds. The ANSam tone consists of a sinewave of 2100Hz with phase reversals at an interval of 450 ms and amplitude-modulated by another sinewave of 15Hz (amplitude modulation index 0.2). This ANSam tone should disable all echo cancellers along the way and trigger the device under test into a full-duplex modem mode. The director stops sending CI after detecting the end of ANSam from responder, and then proceeds the following test sequence:

1. Perform coordinated round-trip delay measurement between the director and responder.

2. Director transmits test parameters to responder via in-band telemetry.

3. Perform full-duplex packet impairment test on both directions simultaneously since the V34/V90 modem works in full-duplex mode. Both the director and responder transmits the test signals. However, the forward and reverse signals do occupy non-overlapping frequency bands so that an undisabled echo canceller will not disrupt the signals and that IP-TSTT will work under 2-wire POTS interface. Both director and responder also performs packet loss, jitter, gain measurements and codec detection while transmitting the test signals.

4. After reaching the specified test duration, the responder sends back measurement results via in-band telemetry, and then disconnects the call.

5. After receiving the results, the director then presents the results and disconnects the call.

3.3 Voice half duplex mode

In voice mode, the director, after calling, sends the Hoth noise per ITU-T P.800 [5]. Hoth noise represents a typical background noise in a telephone handset. The responder, after answering the call and detecting the Hoth noise, sends a TPT (test-progress-tone) of 750Hz for 3.6 seconds. The director stops sending Hoth noise after detecting the end of the TPT. It is hoped that Hoth noise and 750Hz TPT will not trigger the device under test into any special Fax or modem mode. The device should stay in a normal voice call mode. After the precursor signals, the IP-TSTT proceeds with the following test sequence:

1. Perform coordinated round-trip delay measurement between the director and responder.
2. Director transmits test parameters to responder via in-band telemetry.
3. Perform simplex forward direction packet impairment test. The director transmits the unique test signal, and the responder performs packet loss, jitter, gain measurements and codec detection with this unique signal. The reverse direction is quiet.
4. After reaching the specified test duration, the responder sends back measurement results via in-band telemetry. The director receives the forward direction test results.
5. Perform simplex reverse direction packet impairment test. The responder transmits the unique test signal, and the director performs packet loss, jitter, gain measurements and codec detection with this unique signal. The forward direction is quiet.
6. After reaching specified test duration, both ends disconnect the call.

4 User guidance

4.1 Guidance on selecting the test parameters

Generally, the test signal level should be high enough to maximize the signal-to-noise ratio. However, the level should not be so high as to cause clipping and other nonlinear effects. The IP-TSTT test signal has a peak-to-RMS ratio about 7 dB. As a rule of thumb, a -10 dBm transmit level should be used for all "standard" telephony interfaces such as T1/E1 and analog 2-wire POTS.

The test duration should be chosen according to the purpose of testing. In Fax modem mode, a single page transmission may take 30 to 90 seconds. So a test duration between 30 and 90 seconds should be chosen. However, if multiple page transmission needs to be emulated, then the test duration needs to increased accordingly. For both V34/V90 modem and voice modes, the test duration should be hundreds of seconds.

The choice of testing mode is quite self-explanatory. The purpose of the test determines the mode. If one wants to verify the Fax transmission capability, one chooses Fax modem mode. If one wants to find out how much packet impairments there are for a voice call, one chooses the voice mode.

4.2 Guidance on interpreting the results

In Fax and V34/V90 modem modes, the round-trip delay measurement is less critical. However, theoretically, the longer the delay (¿300 ms), the better. If implemented properly, the IAD or gateway under test should use the maximum jitter buffer size (hence longer delay) so as to minimize the potential of packet loss. IP-TSTT should not measure and report any amount of packet loss and jitter. Any measurable packet loss or jitter indicates a failure of the device's capability of handling a data call. The signal gain should be in the range of [-20,0] dB. In Fax mode,the codec type must be PCM or ADPCM. In V34/V90 modem mode, the codec type must be PCM.

For the voice mode, the round-trip delay measurement is very critical. The shorter the delay, the better the voice quality (<300 ms). Certain amount of packet loss and jitter can be tolerated (< 3% packet loss, for example), but to assure voice quality, fewer impairments are always better.

5 Specifications of the test signals

The test signals used for measuring packet impairments are "broadband complex" signal. The bandwidth is 500 Hz, and the peak-to-RMS ratio is about 7 dB. The principle test signal occupies the band of [304.6875 796.875] Hz. The secondary test signal (used only during the full-duplex V34/V90 mode in the reverse direction) occupies the band of [1304.6875 1796.875] Hz. Both signals have a period of 128 ms. The gain measurement is integrated across the whole 500 Hz band occupied by the test signal.

6 Exceptional error information

This IP-TSTT has the same level of robustness as SMOS [1]. If this test cannot proceed successfully, the following causes should be investigated:

1. Operator error: the responder has not been set up properly before starting the director.
2. Audio connectivity problem. Make sure there are audio connections on both directions after call establishment. Some simple send-tone and measure-tone type of tests on both directions should be performed before starting an automated director-responder type of test.
3. Signal level is set too high or too low. This is particularly a problem for some "non-standard" dry-circuit audio connections, as the unknown phone handset impedance and voltage level makes it hard to choose the proper signal level. Several rounds of trial and error may be required to make the test work.
4. Excessive amount of packet impairments. If the packet impairment exceeds the equivalent of 30% packet loss in PCM mode, the test may fail at one of the following 3 stages: the delay measurement stage, the forward-command telemetry stage and the reverse results telemetry stage. In this case, one may argue, the failure to complete the test itself is an information. The information indicates that there are too many impairments in the network and device under test.

References

[1] Renshou Dai, "A white paper on SMOS", Sage Instruments white papers, November, 2001.

[2] "Procedures for document facsimile transmission in the general switched telephone network," *ITU-T Recommendation T.30*, April, 1999.

[3] "Procedures for starting sessions of data transmission over the public switched telephone network," *ITU-T Recommendation V.8*, November, 2000.

[4] "300 bits per second duplex modem standardized for use in the general switched telephone network," *ITU-T Recommendation V.21*, 1993.

[5] "Methods for subjective determination of transmission quality," *ITU-T Recommendation P.8000*, 1996.