Identifying the Factors that Affect Voice Quality

Document ID: 108718

Contents

Introduction
What are the factors that affect voice quality on SPA3000?
Related Information

Introduction

This article is one in a series to assist in the setup, troubleshooting, and maintenance of Cisco Small Business products.

Q. What are the factors that affect voice quality on SPA3000?

A.

Voice Quality perceived by the subscribers of the IP Telephony service should be indistinguishable from that of the PSTN. Voice Quality can be measured with methods such as Perceptual Speech Quality Measurement (PSQM) (1–5 – lower is better) and Mean Opinion Score (MOS) (1–5 – higher is better).

This table displays speech quality metrics associated with various audio compression algorithms:

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Bandwidth</th>
<th>Complexity</th>
<th>MOS Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>G.711</td>
<td>64 kbps</td>
<td>Very Low</td>
<td>4.5</td>
</tr>
<tr>
<td>G.726</td>
<td>16, 24, 32, 40 kbps</td>
<td>Low</td>
<td>4.1 (32 kbps)</td>
</tr>
<tr>
<td>G.729a</td>
<td>8 kbps</td>
<td>Low - Medium</td>
<td>4</td>
</tr>
<tr>
<td>G.729</td>
<td>8 kbps</td>
<td>Medium</td>
<td>4</td>
</tr>
<tr>
<td>G.723.1</td>
<td>6.3, 5.3 kbps</td>
<td>High</td>
<td>3.8</td>
</tr>
</tbody>
</table>

Factors that Affect Voice Quality

• Audio Compression Algorithm
Speech signals are sampled, quantized, and compressed before they are packeted and transmitted to the other end. For IP Telephony, speech signals are usually sampled at 8000 samples per second with 12–16 bits per sample. The compression algorithm plays a large role in determining the Voice Quality of the reconstructed speech signal at the other end. The SPA supports the most popular audio compression algorithms for IP Telephony: G.711 a-law and –law, G.726, G.729a and G.723.1. The encoder and decoder pair in a compression algorithm is known as a codec. The compression ratio of a codec is expressed in terms of the bit rate of the compressed speech. The lower the bit rate, the smaller the bandwidth required to transmit the audio packets. Voice Quality is usually lower with lower bit rate. However, Voice Quality is usually higher as the complexity of the codec gets higher at the same bit rate.

• Silence Suppression?
The SPA applies silence suppression so that silence packets are not sent to the other end in order to conserve more transmission bandwidth. Instead, a noise level measurement can be sent periodically during silence suppressed intervals so that the other end can generate artificial comfort noise that mimics the noise at the other end using a CNG or comfort noise generator.

• Packet Loss
Audio packets are transported by UDP which does not guarantee the delivery of the packets. Packets...
may be lost or contain errors which can lead to audio sample drop−outs and distortions and lowers the perceived Voice Quality. The SPA applies an error concealment algorithm to alleviate the effect of packet loss.

• **Network Jitter**
The IP network can induce varying delay of the received packets. The RTP receiver in the SPA keeps a reserve of samples in order to absorb the Network Jitter, instead of playing out all the samples as soon as they arrive. This reserve is known as a Jitter Buffer. The bigger the Jitter Buffer, the more jitter it can absorb and the bigger the delay it can introduce. Therefore the jitter buffer size should be kept to a relatively small size whenever possible. If jitter buffer size is too small, then many late packets may be considered as lost and thus lowers the Voice Quality. The SPA can dynamically adjust the size of the jitter buffer according to the network conditions that exist during a call.

• **Echo**
Impedance mismatch between the telephone and the IP Telephony gateway phone port can lead to near−end echo. The SPA has a near end echo canceller with at least 8 ms tail length to compensate for impedance match. The SPA also implements an echo suppressor with comfort noise generator (CNG) so that any residual echo will not be noticeable.

• **Hardware Noise**
Certain levels of noise can be coupled into the conversational audio signals due to the hardware design. The source can be ambient noise or 60Hz noise from the power adaptor. The SPA hardware design minimizes noise coupling.

• **End−to−End Delay**
End−to−end delay does not affect Voice Quality directly but is an important factor in determining whether subscribers can interact normally in a conversation taking place over an IP network. Reasonable delay figure should be about 50–100ms. End−to−end delay larger than 300ms is unacceptable to most callers. The SPA supports end−to−end delays well within acceptable thresholds.

### Related Information

• Technical Support & Documentation – Cisco Systems

---

Updated: Dec 12, 2008

Document ID: 108718