Influence of Security Mechanisms on the Quality of Service of VoIP

Peter Backs[1] · Norbert Pohlmann[2]

[1]Sirrix AG
Lise-Meitner-Allee 4, D-44801 Bochum
p.backs@sirrix.com

[2]Institute for Internet Security
University of Applied Sciences Gelsenkirchen
Neidenburger Str. 43, D-45877 Gelsenkirchen
norbert.pohlmann@informatik.fh-gelsenkirchen.de

Abstract

While Voice over IP (VoIP) is advancing rapidly in the telecommunications market, the interest to protect the data transmitted by this new service is also rising. However, in contrast to other internet services such as email or HTTP, VoIP is real-time media, and therefore must meet a special requirement referred to as Quality-of-Service to provide a comfortable flow of speech. Speech quality is worsened when transmitted over the network due to delays in transmission or loss of packets. Often, voice quality is at a level that even prevents comprehensive dialog. Therefore, an administrator who is to setup a VoIP infrastructure might consider avoiding additional decreases in voice quality resulting from security mechanisms, and might leave internet telephony unprotected as a result. The inspiration for this paper is to illustrate that security mechanisms have negligible impact on speech quality and should in fact be encouraged.

1 Introduction

Telephony systems are the most important communication media in modern society[1]. Nonetheless, most users still tolerate the inherent security weaknesses in voice transmission over telephones. Indeed, circuit-switched telephone communication is not encrypted and the call participants are not authenticated. Circuit-switched telephony security is based on physical protection of the telephone line, which leaves communication data completely exposed to an attacker gaining physical access. The same situation applies to VoIP; the attacker must have access to the telephony device of a call participant or the network media carrying the voice data. In comparison to classic telephone technology, however, access to transport media is achieved far more easily. So-called Spoofing Attacks can be used to fake the identification (ID) of a caller, allowing the redirection of VoIP traffic through the system of an attacker to be easily eavesdropped. Another threat is inherent to internet routing technology, as IP packets are not meant to be transported over fixed routes from one host to another. Instead, routes are determined dynamically and beyond the control of communicating peers. In this way, a malicious system could be found along the route of a packet, exposing voice content if not secured properly. Furthermore, the ability to intercept

1 A dedicated study ascertained by University of Applied Science Wiesbaden, Clarity Voice Commerce Research Center, Fachverband für Sprachtechnologien, Voice Commerce Application Committee (VASCom) and Technologiestiftung Hessen (TSH) points out, that telephone communication remains the most important communication channel between a company and its customers.

N. Pohlmann, H. Reimer, W. Schneider (Editors): Securing Electronic Business Processes, Vieweg (2008), 1-6

voice data traffic is not dependent on special hardware, as is the case with ISDN, for example. Instead, a standard PC and freely-available software is sufficient for very effective attacks. To sum up, the danger of a successful attack against VoIP is even greater than that inherent to classic telephone technology.

1.1 Securing Voice-over-IP

Towards this end, several solutions have been developed that offer the security required. Typically, there are two approaches applied towards realizing these solutions: the establishment of a Virtual Private Network (VPN) and the employment of VoIP-specific protocols.

The first of these approaches, VPN, protects any kind of IP traffic independently of the communication service employed. There exist a variety of VPN solutions on the market. Most are based on the IPSec standards from IETF [IPSEC_1], [IPSEC_2], while another very popular product is the open, but not standardized, OpenVPN [OVPN].

A VPN normally extends protection from one site to another. The IPSec standards cover end-to-end protection, but the protocol is not so widely used. The advantage of site-to-site protection is that transmitted data is encrypted transparently for the VoIP participants, but connections are only protected between VPN nodes.

On the other hand, VoIP-specific concepts were developed to specifically protect VoIP connections. The Real-time Transport Protocol (RTP) is used to transfer real-time data, such as a voice stream. The Secure RTP (SRTP) standard extends RTP capabilities by encrypting and signing its payload. Similarly, the standard Session Initiation Protocol (SIP), used to signal calls, has been extended to a more secure variant, SIPS. In contrast to a VPN, SRTP offers end-to-end protection of voice data. However, a special client is required to support SIPS and SRTP. As well, a cryptographic key infrastructure must be deployed including all VoIP participants and not only the viewer VPN nodes.

1.2 Quality of Service in Voice over IP

When referring to quality in VoIP, quality of data transmission over the network and quality of speech is important. Quality of transmission refers to the parameters of the network voice media is transmitted over, including end-to-end delay, variation in delay, packet loss and bandwidth. In contrast to quality of speech, quality of transmission way can be objectively characterized by these parameters. Voice quality is the quality of the transmitted audio data in comparison to a face to face dialog. The only way to rate this quality is by personal impression. Voice quality is mainly influenced by the voice codec, the algorithm used to compress the media stream. Both factors are connected, as quality of data transmission directly influences the quality of speech. A high end-to-end delay also impacts on quality of speech, when flow in dialog is disturbed for instance.

2 Impact of Security on Quality of Service

The security aspects of these implementations are based on well known cryptographic procedures and met with general acceptance. However, research has yet to be conducted to ascertain the overall impact when using VoIP with these security implementations. This paper addresses this gap in knowledge by analyzing and comparing common VoIP security implementations as to their effect on VoIP usage. Besides organizational and technical efforts, such as manual keying (if no cryptographic key infrastructure has been deployed), security mechanisms affect the Quality-of-Service (QoS) of a VoIP connection.

Parameters of the QoS that are affected include end-to-end delay, variation of delay (jitter), packet loss and bandwidth consumption.

The impact on QoS is easy to measure across a VPN when specific tools are used. These tools act as VoIP extensions sending test calls to one other. When set in different networks connected by a VPN tunnel, the impact on quality can be determined by comparing against those results obtained when the VPN functionality is disabled (i.e. the reference measurement). Figure 1 shows the testing environment consisting of two separated networks connected by a routing network. Security gateways at the network boundary feature VPN functionality. OpenVPN and OpenSWAN were chosen as a representative selection of most common VPN solutions.

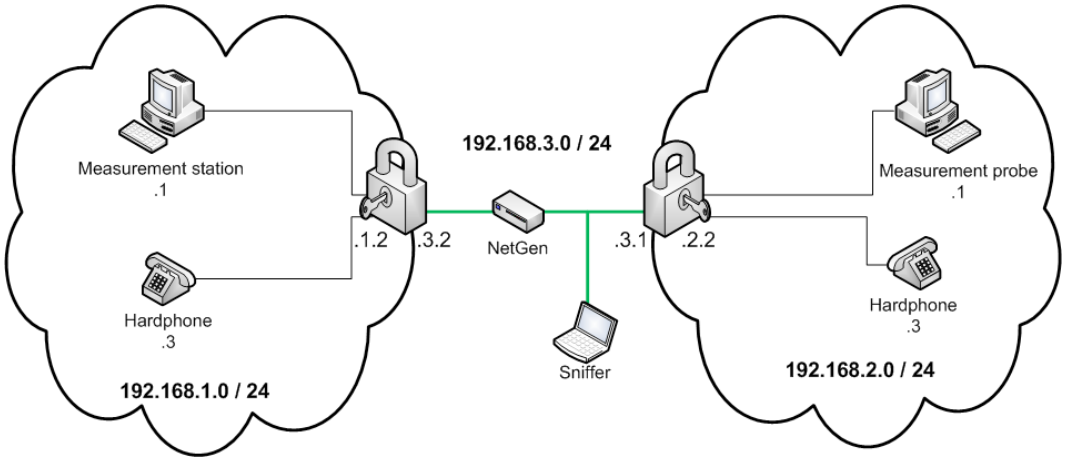

Figure 1: Network configuration for testing VPN based security solutions.

In fact, the above mentioned tools are only useful when analysing transparent security concepts. If the security solution has to be implemented by the VoIP participants, the test tools must implement them as well. As this is not the case, no measurements can be obtained when the VoIP-specific SRTP standard is employed. However, the behaviour of SRTP implementations is more predictable due to its greater simplicity. Firstly, cryptographic calculations are distributed over the VoIP clients as opposed to being concentrated to a single point as in the case of a VPN. Furthermore, stream cipher algorithms encouraged by the SRTP standard are supposed to work fast, even on embedded devices, so as to expect that additional computing would be minimized. Last, but not least, SRTP does not tunnel the entire VoIP packet into a new one, but merely encrypts the voice data and appends a checksum instead. In contrast, VPN tunnelling adds additional headers which can be of even bigger size than the VoIP payload itself, resulting in bandwidth more than double of the unencrypted VoIP stream. The simpler concept of SRTP implies faster processing. As a result, the impact of SRTP implementations should be considered less than that experienced when applying VPN protection.

2.1 Delay, Jitter and Packet Loss

To investigate the validity of these assertions, the influence on the QoS stemming from the various security processes was measured. It was concluded that none of the security implementations tested had a significant impact on the delay, jitter or packet loss. The end-to-end delay increased less than two milliseconds. According to the ITU-T, a delay in a voice connection of 150 milliseconds or less still allows for optimal voice quality. Hence, the delay introduced when protecting the signal is negligible. Jitter and packet loss measurements were essentially equal in protected versus unprotected connections.

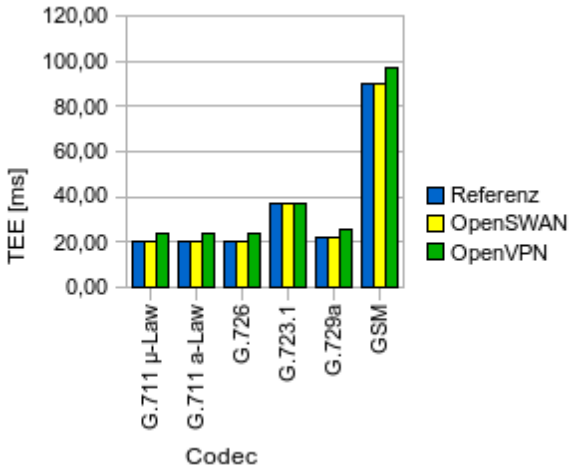

Figure 2: End-to-end delay (TEE) for different voice codecs and security solutions.

2.2 Bandwidth

In contrast to delay, jitter and packet loss, bandwidth requirements increased more substantially. Depending on the voice coding and framing parameters, as well as the security software employed, the bandwidth consumed by a protected VoIP connection is more than double that of an unprotected one. Therefore, if the available bandwidth is limited, protection could be a restricting factor. Two of the factors, the voice transmission settings and the security software, can be tweaked so as to tune down the bandwidth requirements. To begin with, the VoIP software can be set up to submit more than one voice frame per RTP packet, with the effect that packet overhead (and, therefore, security overhead) is reduced. That is, if two frames are sent in one RTP packet, the overhead is halved. This reduces the bandwidth but causes an additional delay due to the accumulation of an extra frame for each transmission. Another option to tune the bandwidth is the choice of the voice codec. This way, the additional bandwidth required for the cryptographic functionality can be compensated by the use of a low-bandwidth codec. In most cases, however, this also has the effect that voice quality decreases and the sensitivity to packet loss increases. The security settings can be adjusted to further reduce the bandwidth of the signal. The most influential of these stems from the security protocol employed. Whereas SRTP merely adds a signature to the RTP packet, VPN solutions encapsulate the entire RTP packet, resulting in more overhead. Furthermore, SRTP encourages the use of stream cipher encryption. Stream ciphers tend to process streamed data more efficiently, as they do not make use of fixed block sizes. The IPSec and OpenVPN protocols both rely on block ciphers, which pad the payload to fit the block size and, therefore, increase the overhead. SRTP is most often the optimal choice to reduce the bandwidth requirements of a secure connection.

Table 1: Bandwidth consumption [Kbps] of voice codecs at different packeting settings, security protocols and security protocol cipher block size (BS).

	Voice stream,	Segm. /	RTP	IPSec		OpenVPN		SRTP
Codec	Circuit-switched	packet	n/a	BS 16	BS 8	BS 16	BS 8	n/a
G.711	64	1	96	147,2	140,8	166,4	153,6	102,4
		2	80	105,6	102,4	115,2	108,8	83,2
G.723.1	5,3	1	16	32	29,867	38,4	34,133	18,133
		2	10,667	20,267	18,133	21,333	20,267	11,733
	6,3	1	17,067	36,267	32	38,4	36,267	19,2
		2	11,733	20,267	19,2	23,467	21,333	12,8
G.726	16	1	48	96	89,6	115,2	102,4	54,4
		2	32	60,8	54,4	64	60,8	35,2
	24	1	56	108,8	102,4	115,2	108,8	62,4
		2	40	67,2	60,8	70,4	67,2	43,2
	32	1	64	121,6	108,8	128	121,6	70,4
		2	48	73,6	70,4	83,2	76,8	51,2
	40	1	72	121,6	115,2	140,8	128	78,4
		2	56	80	76,8	89,6	83,2	59,2
G.729a	8	1	40	96	83,2	102,4	96	46,4
		2	24	48	44,8	57,6	51,2	27,2
GSM	13	1	29,2	54,4	51,2	64	57,6	32,4
(FullRate)		2	21,2	33,6	32	38,4	35,2	22,8
iLBC	13,33	1	24	40,533	38,4	46,933	42,667	26,133
		2	18,667	26,667	25,6	29,867	27,733	19,733
	15,2	1	31,2	60,8	54,4	64	57,6	34,4
		2	23,2	36,8	33,6	38,4	36,8	24,8

As table 1 illustrates, the bandwidth consumed by a protected VoIP stream may be more than double the bandwidth of an unprotected stream even when using the same speech codec. The differences are even more significant when comparing the influence of codecs. In any case, the figure above should allow one to estimate the volume of VoIP traffic to expect and be a useful reference for tweaking it.

3 VPN Gateway Load

Another important fact regarding VoIP and VPN security mechanisms is the load on the VPN gateways. As a central node protecting all passing data, the VPN gateway has to perform encryption and verification for all VoIP clients on the network. Measurements showed that VoIP packets, in particular, stress the gateway more than ordinary IP traffic. This is due to the very small packet size in VoIP streams in comparison to, e.g., FTP or HTTP. Due to the higher average load, gateways tend to be quicker to reach their limit in throughput when VoIP packets are handled. The dramatic decrease in throughput is shown in Figure 3. The lines show the end-to-end delay when raising the data throughput of the gateway. At the point of singularity, the gateway can be considered overloaded and at the point of maximum throughput. The difference in the maximum throughput when sending small packets, as opposed to larger packets, is apparent in the thick and thin lines of each colour. The reference, which implements ordinary routing devices instead of securing VPN stations, performs a maximal throughput of roughly 60 Mbps when routing VoIP packets versus 95 Mbps when sending larger packets. The same situation can be noted

when implementing VPN gateways. When using Open VPN, the throughput drops from 40 to 15 Mbps, which is nearly the same when implementing OpenSWAN.

To compensate for this, an administrator might install more powerful gateway hardware or even consider a clustering solution, which also renders failover services.

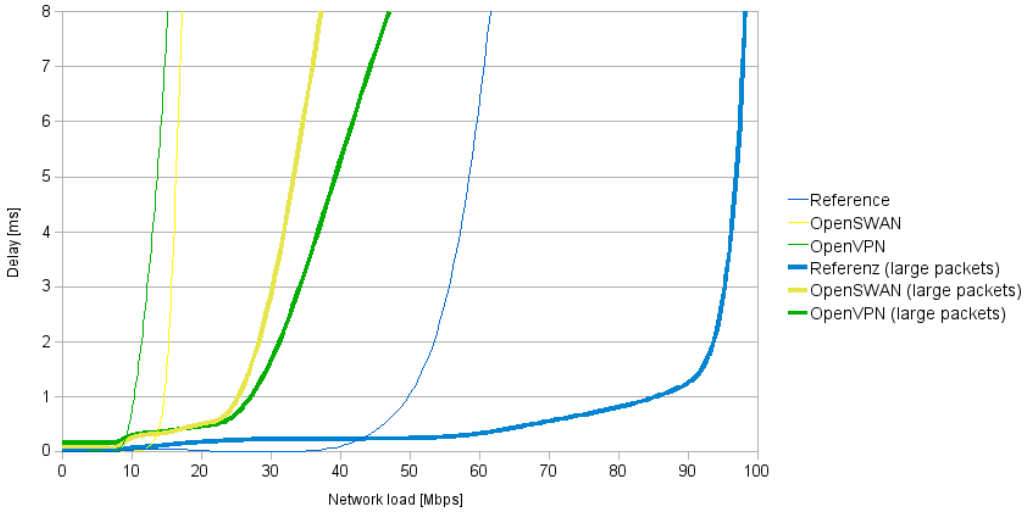

Figure 3: Limit in network load depending on security solution implemented and the network packet size.

4 Conclusion

To sum up, the QoS parameters of a voice stream, including delay, jitter and packet loss, are hardly affected by security measures. Special considerations have to be taken with regard to the consumption of bandwidth by a secure, versus insecure, call. The overhead for protection may have a significant impact on the required bandwidth, but the administrator has several ways to compensate for this. Also, the administrator must consider that a security gateway will overload more easily when processing VoIP as opposed to other internet services; however, this can be resolved easily enough by scaling server capacity. In any case, secure VoIP should not be dismissed due to misgivings of decreased voice quality, but encouraged by the many risks addressed by the protection it provides.

References

[ZDNT03] Fiutak, Martin: Telefontopp t E-Mail, 2003, http://www.zdnet.de/news/tkomm/0,39023151,39118288,00.htm.

[OVPN] OpenVPN community: OpenVPN website, http://www.openvpn.net

[IPSEC98_1] Kent, Stephen / Atkinson, R., RFC 2401: Security Architecture for the Internet Protocol, IETF, November 1998

[IPSEC98_2] Kent, Stephen / Atkinson, R., RFC 2406: IP Encapsulating Security Payload (ESP), IETF, November 1998