Impact of Security on Speech Quality

Miroslav Vozňák[1] – Filip Řezáč[1] – Antonio Nappa[2] - Alessandro Rozza[2]

[1] CESNET, z.s.p.o.
Zikova 4, 160 00 Prague
Email: miroslav.voznak@vsb.cz , filip.rezac.st@vsb.cz

[2] University of Milan
Via Comelico 39/41, I-20135 Milan
Email: nappa@security.dico.unimi.it , rozza@dico.unimi.it

Abstract

This paper deals with impact of secured environment on speech quality of IP telephony. There are presented the results of the analyzing of voice over secure communication links based on TLS. The using of secure network environments can affect a speech quality. There is the performance comparision of cipher alghorithms and description how the used security mechanisms influence the final R-factor. The presented results are based on numerous of experiments which have been performed in real IP network.

1. Introduction

Virtual Private Network (VPN) is a technology to construct a private network over public networks. OpenVPN [1] is one of the most popular software-based VPN products and has high flexibility. The usability of OpenVPN is high because offers a open-source, cost-effective and widely testet solution, not requiring expert knowledge. Software VPN products are popular, because they don't need any appliance and OpenVPN provides such solution which is based on matured protocols. The OpenVPN security model is based on SSL (Secure Socket Layer), the industry standard for secure communications via IP network. OpenVPN implements transport secure network extension using the TLS protocol (Transport Layer Security).

On the other side the using of OpenVPN increases an overhead which is affected by encryption and this overhead can influence overall speech quality [2]. This paper contains a description of OpenVPN and its possibilities regarding a configuration, than there is explained a core of the matter which is the splitting of a RTP packet to equally divided blocks.

2. OpenVPN and encryption

TLS ensures a secured connection which is encrypted and decrypted with the keys negotiated during a phase of keys exchange. The key exchange and authentication algorithms are typically public key algorithms but subsequent data exchange is usually done by symmetric ciphers because of considerably faster processing. Of course, symmetric encryption is more suitable for IP telephony and this paper deals only with this type of ciphers [3]. TLS involves three main phases such as negotiation of supported algorithms, keys exchange and authentication and in the end symmetric encryption of transmitted data.

The endpoint establishing VPN tunnel are declared one as server and the other as client. Before establishing the VPN, the client first reaches the server on a specific port, whereas the server doesn't need to reach the client. Configuration files are located in directory /etc/openvpn as server.conf or client.conf. The tunnel can be established on UDP or TCP, unfortunately TCP protocol is more widespread although UDP is more effective because of real-time applications. The most important information in configuration files is the type of cipher alghoritm because it affects the number of blocks and overhead as is shown in fig. 1.

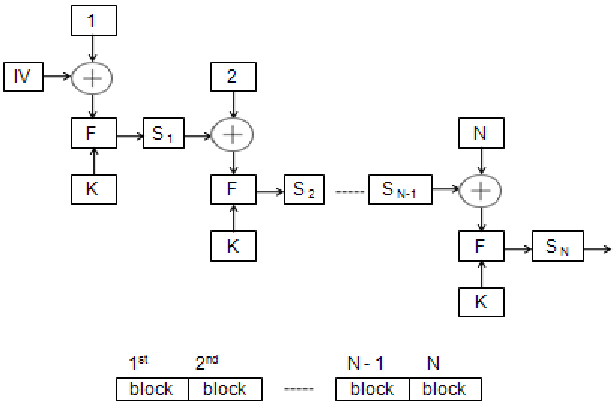


Figure 1. The number of blocks affected by CBC mode.

The figure above illustrates the splitting of one RTP packet to N blocks. Every block has the same lenght which contains in case of AES (Advanced Encryption Standard) 128 bits although the key size can be 192 or 256 bits but the block has the same size 128 bits. If another alghoritms are applied such as DES (Data Encryption Standard), Triple DES or BF (Blow Fish), then the block size is set to value 64 bits. A complete list of supported cipher alghoritms can be obtained as a result of command *openvpn --show-ciphers*. The following ciphers and cipher modes are available for use with OpenVPN. Each cipher shown below may be used as a parameter to the *--cipher option*. The default key size is shown as well as whether or not it can be changed with the *--keysize* directive. Using a CBC mode is recommended.

DES-CBC 64 bit default key (fixed)
RC2-CBC 128 bit default key (variable)
DES-EDE-CBC 128 bit default key (fixed)
DES-EDE3-CBC 192 bit default key (fixed)
DESX-CBC 192 bit default key (fixed)
BF-CBC 128 bit default key (variable)
RC2-40-CBC 40 bit default key (variable)
CAST5-CBC 128 bit default key (variable)
RC2-64-CBC 64 bit default key (variable)
AES-128-CBC 128 bit default key (fixed)
AES-192-CBC 192 bit default key (fixed)
AES-256-CBC 256 bit default key (fixed)

CBC means Cipher-block chaining, in this mode of operation, each block of plaintext is XORed with the previous ciphertext block and afterward is encrypted, that is why an initialization vector IV must be used in the first block, see figure 1.

3. Used techniques of measurement

The presented results in next chapter are based on series of measurements which has been performed in real network with OpenVPN and IxChariot [4], a scheme is shown in figure 2.

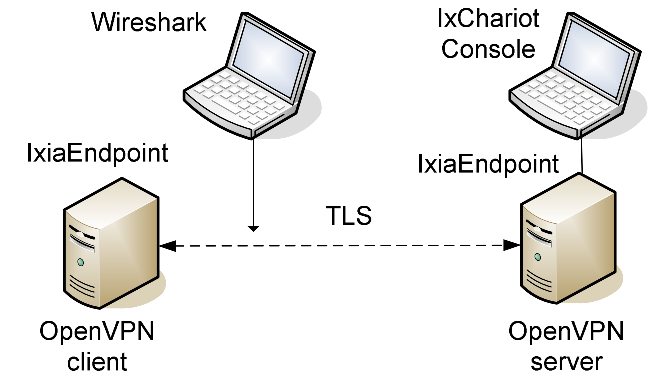


Figure 2. Logical scheme of testbed.

Whole traffic carried out between OpenVPN client and server was captured by Wireshark [5] and individual packets were analyzed. IxChariot is a software, produced by Ixia, which consists of the IxChariot console and IxChariot endpoints. The IxChariot console allows a selection of several test configurations, such as the used codec, timing, number of concurrent calls, test duration and so on. The test is initialized at the console, the conditions are uploaded into endpoints and consequently the test is performed. The results are sent back to the console. There was observed an influence of OpenVPN-TLS on overhead which has been increased and hence the required bandwith has been affected.

4. Bandwith requirements

The basic steps of speech processing on a transmission side are encoding and packetizing [6], [7]. RTP packets are sent in dedicated times and a difference between them depends on timing [8]. This process of packetizing is given by the following basic equation:

$$\Delta t = \frac{P_S}{C_R} \tag{1}$$

where Δt [s] is timing in seconds, P_S [b] is a payload size and C_R [kbps] represents a codec rate. The timing can be derived from content of RTP packet as a difference of two consecutive timestamps, see relation (2). Typical value of a sampling frequency is 8 KHz.

$$\Delta t = \frac{\text{timestamp}_{\{N+1\}} - \text{timestamp}_{\{N\}}}{\text{sampling\_frequency}} \tag{2}$$

There is necessary to express a size of packet at application layer which might be defined by the following formula:

$$S_{AL} = H_{RTP} + P_S \tag{3}$$

where S_{AL} [b] is the expected size that consists of RTP header H_{RTP} [b] and payload size P_S [b]. Equation (4) determines a size of frame S_F [b] at link layer.

$$S_F = S_{AL} + \sum_{j=1}^{3} H_j \tag{4}$$

S_F [b] includes a packet at application layer and the sum of lower located headers of OSI model where H_1 [b] is media access layer header, H_2 [b] internet layer header and H_3 [b] is transport layer header.

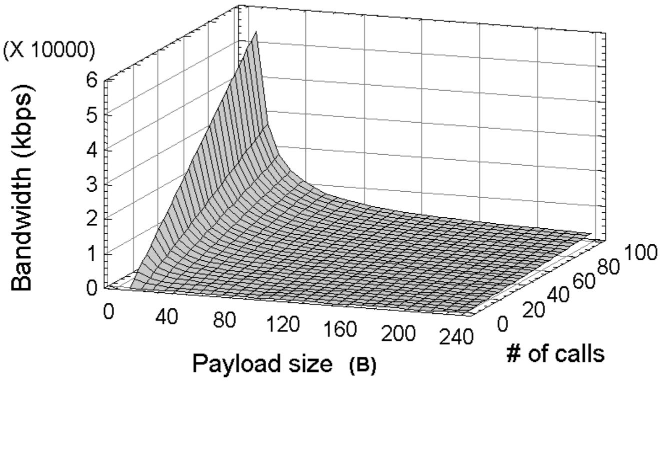


Figure 3. Bandwith as a function of payload size and concurrent calls.

Figure 3 illustrates the relation between bandwith, payload size and number of concurrent calls.

$$BW_M = \sum_{i=1}^{M} \frac{S_{Fi}}{\Delta t_i} \tag{5}$$

Total bandwith BW_M [kbps], which is required in case of M concurrent calls, we express in relation (5) and if we apply the relations (1), (2) and (4) to the relation (5) so we obtain the following result (6).

$$BW_M = M \cdot C_R \cdot \left(1 + \frac{H_{RTP} + \sum_{j=1}^{3} H_j}{P_S} \right) \tag{6}$$

We have to realize that TLS is located between two layers of OSI model, between application and transport layer and therefore we apply S_{TLS} instead of S_{AL}. This replacement should be done in respect of explained location of TLS and we define a new parameter S_{TLS}, size at TLS layer. S_{TLS} is expressed in relation (7).

$$S_{TLS} = C_0 + \left\lceil \frac{S_{AL}}{B_S} \right\rceil \cdot B_S \tag{7}$$

We use a symbol $\lceil x \rceil$ to denote the ceiling function where $\lceil x \rceil = \min\{n \in \mathbb{Z} \mid x \le n\}$, it means that ceiling function of x gives the smallest integer greater than or equal to x. The ceiling function was defined by M. Schroeder in 1991 [9] and the symbol was coined by K. Iverson in 1994. The parameter B_S represents a block size which has been explained in figure 1, its value is 64 or 128 bits and depends on applied cipher alghoritm (AES, DES, Triple DES or Blow Fish). C_0 is a constant and equals to zero in case of clear TLS unfortunately OpenVPN adds supplementary overhead that is included in C_0. The value has been achieved by performed experiments, see picture 2. We can claim that this constant C_0 = 83 bytes in case of block size 128 bits and C_0 = 75 bytes in case of block size 64 bits.

5. Achieved results

Relations stated in previous chapter have been confirmed by experiments, following figures illustrate how required bandwith is affected by TLS and OpenVPN.

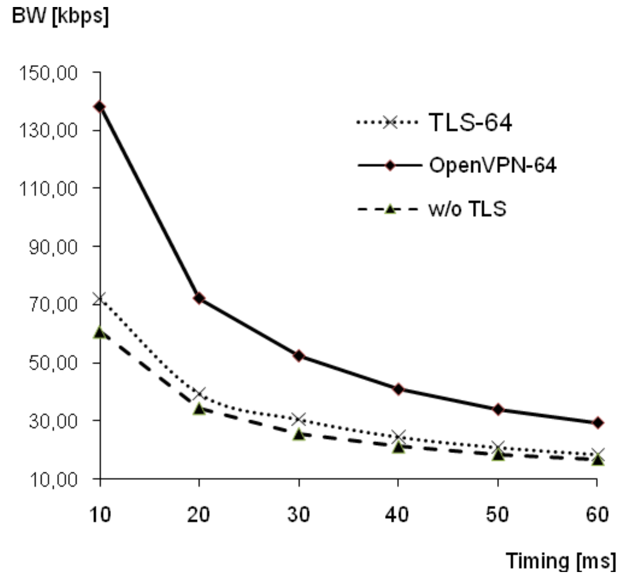


Figure 4. Comparision of required bandwith for codec G.729 without TLS, with TLS and OpenVPN, Bs =64 bits

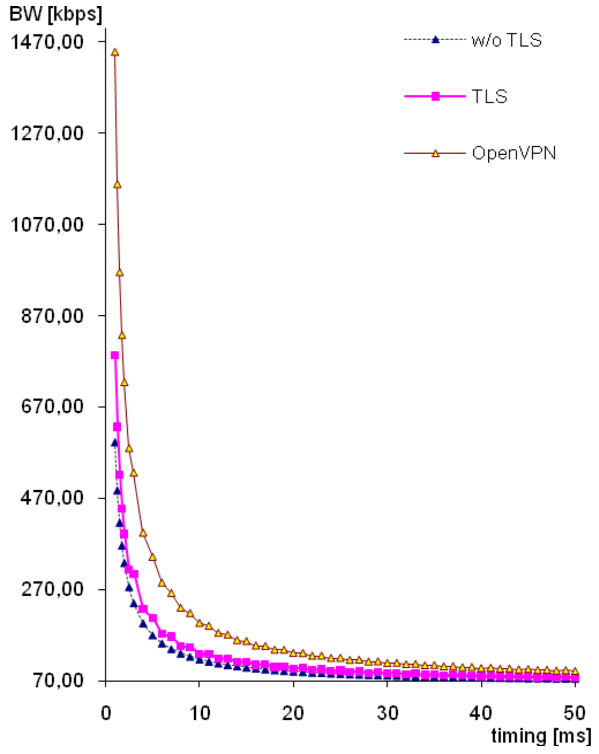
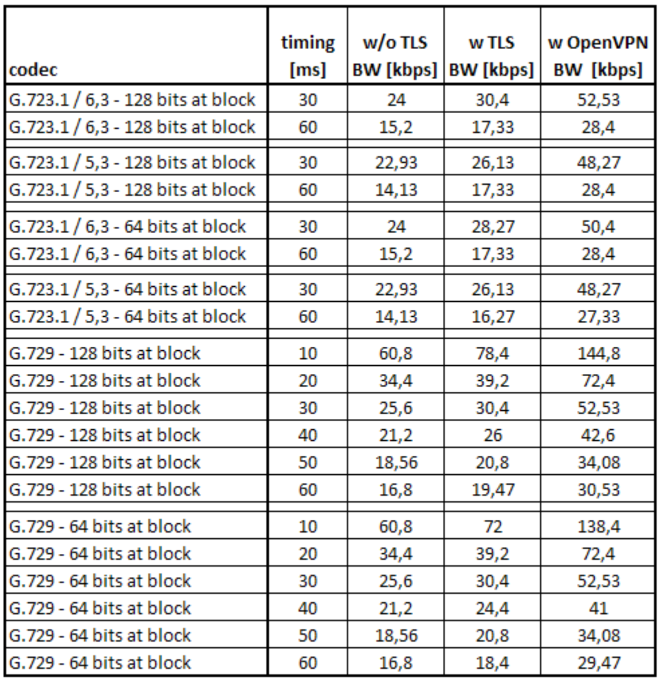


Figure 5. Comparision of required bandwith for codec G.711 without TLS, with TLS and OpenVPN, Bs =128 bits

The first column of table 1 contains codec G.729 and both variants of G.723.1. Block size has a length eiher 64 bits or 128 bits. This table provides the results for Ethernet without TLS, with TLS and with OpenVPN.

Table 1. Values of required bandwith in various environments



codec	timing [ms]	w/o TLS BW [kbps]	w TLS BW [kbps]	w OpenVPN BW [kbps]
G.723.1 / 6,3 - 128 bits at block	30	24	30,4	52,53
G.723.1 / 6,3 - 128 bits at block	60	15,2	17,33	28,4
G.723.1 / 5,3 - 128 bits at block	30	22,93	26,13	48,27
G.723.1 / 5,3 - 128 bits at block	60	14,13	17,33	28,4
G.723.1 / 6,3 - 64 bits at block	30	24	28,27	50,4
G.723.1 / 6,3 - 64 bits at block	60	15,2	17,33	28,4
G.723.1 / 5,3 - 64 bits at block	30	22,93	26,13	48,27
G.723.1 / 5,3 - 64 bits at block	60	14,13	16,27	27,33
G.729 - 128 bits at block	10	60,8	78,4	144,8
G.729 - 128 bits at block	20	34,4	39,2	72,4
G.729 - 128 bits at block	30	25,6	30,4	52,53
G.729 - 128 bits at block	40	21,2	26	42,6
G.729 - 128 bits at block	50	18,56	20,8	34,08
G.729 - 128 bits at block	60	16,8	19,47	30,53
G.729 - 64 bits at block	10	60,8	72	138,4
G.729 - 64 bits at block	20	34,4	39,2	72,4
G.729 - 64 bits at block	30	25,6	30,4	52,53
G.729 - 64 bits at block	40	21,2	24,4	41
G.729 - 64 bits at block	50	18,56	20,8	34,08
G.729 - 64 bits at block	60	16,8	18,4	29,47

6. Impact on R-factor

Lack of bandwith causes a loss in the first place hence the estimation its impact on R-factor is explained in this chapter. . Packet loss distribution can be modelled using a Markov process. A multi-state Markov Model is used to measure the distribution of lost or discarded packets or frames, and to divide the call into “bursts” and “gaps”. The call quality is calculated separately in each state and then combined using a perceptual model, such as in VQmon [13]. The mentioned VQmon does incorporate G.107 compliant implementation of the E-Model. However, we applied a very simple method described in the last revision of G.107 from 2005 [14]. The impairment factor values I_E under packet-loss were tabulated for particular codecs. Robustness Factor B_{pl} is defined as codec-specific value. B_{pl} can be described as the robustness of the codec to packet-loss. Both values are listed in Appendix I of ITU-T G.113 and are available for several codecs. If we consider the Packet-loss Probability as P_{pl}, the I_{E-EF} impairment factor can be calculated using the formula:

$$I_{E-EF} = I_E + (95 - I_E) \cdot \frac{P_{pl}}{\frac{P_{pl}}{BurstR} + B_{pl}} \qquad (8)$$

$BurstR$ is the so-called Burst Ratio, when packet loss is random $BurstR$=1 and when packet loss is bursty $BurstR$>1. For packet loss distributions corresponding to a 2-state Markov model with transition probabilities p between a "found" and a "loss" state, and q between the "loss" and the "found" state, the Burst Ratio can be calculated as:

$$BurstR = \frac{1}{p+q} \qquad (9)$$

Once the I_{E-EF} factor is calculated it is not difficult to determine R-factor as an output of E-Model using implicit values of recommendation ITU-T G.107 which are R_0=94,7688 , Is =1,4136 and A=0, hence we could modify the basic formula of R-factor computing:

$$R = 93,3553 - I_D - I_{E-EF} \qquad (10)$$

The model used to estimate I_D is described in [15]. Where it is explained that the effects of delay are well known and easily modelled. Delays of less than 175 ms have a small effect on conversational difficulty, then $I_D = 4 \cdot T$ where T is the delay in ms.

7. Conclusion

The real-time applications are very sensitive to packet loss, and each variation occurring on the network can modify and influence the final result of a real-time data transmission, such as a VoIP Call. On the one hand the defence techniques using cryptography such as OpenVPN reduce the danger of security threats but on the other side they affect the required bandwith of IP telephony which is significantly increased in case of OpenVPN. The presented relations in this paper help us to understand how OpenVPN and TLS can affect the bandwith of calls and how we can optimize the timing.

For example we can show an optimalization at G.723.1 with 6.3 kbps, see table I. If we used a timing 30 ms during packetization, we would require 30,4 kbps in case of TLS against 24 kbps in environment without TLS, but we could achieve better result with timing 60 ms, because we would require 17,33 kbps for TLS against 15,2 kbps without TLS and it is really much better ratio. This presented example is valid for AES due to size of blocks but the new created realations help to optimize any CBC encryption.

The new contribution of this paper is the presented method of bandwith calculation in network using TLS. The achieved results were confirmed in testbed, the bandwith of any particular call was affected by length of cipher block and didn't depend on key size. The results corresponded with relations stated in chapter IV. This paper is an extension of a previous work on the impact of security on the quality of VoIP calls [10]-[12].

8. Appendix and acknowledgments

This work was supported by CESNET, z.s.p.o. in a framework of its research plan financed by MSMT CR. This paper inlcudes the result achieved in a research activity Multimedia transmissions and collaborative environment.

References

[1] OpenVpn,*The OpenVpn Project*. Available: http://openvpn.net/

[2] Vozňák,M. - Neuman, M.: The Monitoring and Measurement of Voice quality in VoIP Environment, 12 p. Technical report 18/2006, CESNET, November 2006. Available http://www.cesnet.cz/doc/techzpravy/2006/voice-quality/

[3] M. Voznak, *Impact of security on speech quality*, Invited lecture at University of Milan, July.2008. Available http://www.dsi.unimi.it/seminario.php?id=383

[4] Ixia, *IxChariot*.Available http://www.ixiacom.com

[5] Wireshark, *Sniffer*. Available http://www.wireshark.org/

[6] M. Halas, B. Kyrbashov, M. Voznak . *Factors influencing voice quality in VoIP technology*, In: 9th International Conference on Informatics' 2007, pp. 32-35, Bratislava, June 2007

[7] M. Vozňák, E. Rocha, B.Kyrbashov. *End-to-end delay in VoIP*. In proceedings Conference RTT 2007, University of Žilina, 2007, p. 466-469, September 2007.

[8] I. Baroňák, M.Halás, M.Orgoň. *Mathematical model of VoIP connection delay*. In: Telecommunications, Networks and systems, Conference in Lisboa, 3-8 September, 2007.

[9] M, Schroeder, Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise. New York: W. H. Freeman, p. 57, 1991.

[10] Vozňák M., Nappa A.*Performance evaluation of VoIP infrastructure*. In FreeVoice, November 2007.

[11] A.Nappa, D. Bruschi, A. Rozza, M.Voznak, *Analysis and implementation of secure and unsecure Voice over IP environment and performance comparison using OpenSER*. Technical report, 84 pages, published at Universita degli studi di Milano, December, 2007.

[12] M. Voznak, A. Rozza,A. Nappa,*Performance comparision of secure and insecure VoIP environments*.TERENA Networking Conference 2008, Brugge, Belgium, 19-22 May, 2008.

[13] Clark,A. *Extension to the E-Model to incorporate the effects of time varying packet loss and recency*. ETSI TIPHON committee, TS 101 329-5 Annex E, July 2001.

[14] ITU Recommendation G.107. *E-model, a computational model for use in transmission planning*. 2005.

[15] Clark,A. *Modelling the Effects of Burst Packet Loss and Regency on Subjective Voice Quality*, 2001,