Impact of Voice Payload Size on Behaviour of the Field Network of the Armed Forces of the Czech Republic

Zuzana Vranova, Antonin Mazalek
Faculty of Military Technology
University of Defence
Brno, Czech Republic
zuzana.vranova@unob.cz; antonin.mazalek@unob.cz

Abstract - **This article deals with a choice of voice payload size and its impact on the behavior of the field network of the Armed Forces of the Czech Republic. The contradictory requirements for a high throughput and at the same time a high quality of transmitted voice are discussed. First, a theoretical analysis of the problem is carried out. The theoretical assumptions were compared with those obtained in experiments in laboratory conditions. The result of the experiments is to determine the optimal size of the voice payload size with regard to the specific characteristics and requirements on the battlefield. According to the results we can recommend an optimal voice payload size with respect to the G.114 Recommendation. Thus up to 80 % of bandwidth saving will be achieved.**

Keywords – VoIP; voice payload size; network bandwidth; battlefield

I. INTRODUCTION

The Army of the Czech Republic has recently completed modernization of communication infrastructure of the field network. According to current trends in the world communications, the backbone of battlefield is based on the transmission of information using the IP protocol, published by Postel [1]. After overcoming primary difficulties, the modernization brought a stable network enabling the use of a number of new advanced services. A major benefit of the field network modernization with regard to its specific character is a possibility of an easy and secure encryption of a data and voice traffic. The basic building components of the backbone are Cisco 2800 and 3800 series routers that use the services of Cisco CallManager Express (CCME) [2] for the providing of voice services on the VoIP principles. Particular communication nodes are connected by radiorelay links realizing the E1 interface [3].

With today's technical possibilities of digitized battlefield, a bigger focus is getting on the network bandwidth. In addition, a number of applications require a real-time operation. With regard to this trend, the E1 interface with a capacity of 2 Mb/s represents a weak point of the network. Therefore, it is necessary to find a solution to make the transmission of the required information more efficient.

There are many possibilities on how to solve the problem always with regard to the specifics of the analyzed network. Choudhury and Gibson [4] deal with the issue of effectiveness of multimedia transport in wireless access networks. Specifically, the work deals with the possibility of optimizing the efficiency of transmission, depending on the choice of transmission rate and payload length in IEEE 802.11 network. Bhanu, Chandrasekaran, and Balakrishnan [5] solve security of the required quality of services in VoIP 802.11 environment. A Bandwidth Data rate Moderation (BDM) algorithm has been proposed which correlates the data rate specified in IEEE 802.11b with the free bandwidth. Payload size is one of the input parameters of the BDM algorithm. The way to make the data transfer at a tactical command level more efficient by a header compression technique is addressed in the article of Yoon, Park, Lee, Kim, and Jee [6]. Nowadays, modern VoIP networks place attention to safety issues and therefore the voice encryption. Epiphaniou [7] addresses among others the influence of the used encryption and voice payload size on the delay and packet loss. All considered changes need to be addressed with regard to the required voice quality. For example, the voice quality expressed in Mean Opinion Score (MOS) in dependence on voice payload size and packet loss is engaged in the publication of Becvar, Vondra, and Novak [8].

Unlike those publications, we examine specific real network with sufficiently dimensioned HardWare (HW) because of increased requirements on network reliability. We need not, therefore, deal with the impact of packet loss on speech quality. In addition, the network uses external encryption equipment with sufficient packet throughput. These facts allow us to simplify the initial conditions for the experiments.

The weak point of the analyzed network are the radiorelay links realizing the E1 interface. So, the main goal of our work is to optimize efficiency of transmission of VoIP data on E1 links. Furthermore, we use the fact that the E1 interface uses specific data link layer protocols different from the protocols considered in the above publications.

The organization of this paper is as follows. In Section 2, we present factors affecting the level of savings in bandwidth. In Section 3, there is explained a principle of bandwidth calculation for one call in VoIP environment when voice payload size is explored. In Section 4, results of voice payload size optimization in battlefield conditions are shown. In Section 5, we outline our conclusions and we mention our future work.

 ISBN: 978-1-61208-286-8

II. FACTORS AFFECTING THE LEVEL OF SAVINGS IN BANDWIDTH

Nowadays, there are many views how to optimally solve existence of two contradictory requirements on the most efficient use of bandwidth on one hand, and, on the other hand, provision of the high quality voice in the VoIP environment. The rule is that a codec type significantly affects the quality of provided services and the required bandwidth. Its parameters, such as a codec sample interval (CSI), a codec sample size (CSS), are given the same as a value of Mean Opinion Score (MOS) for the default type of codec. Based on the codec, CSI is a sample interval at which the codec operates and CSS is the number of bytes captured by the Digital Signal Processor (DSP) at each codec sample interval. From these two parameters, we can derive another parameter. It is a codec bit rate (CBR) (1)

$$CBR = CSS / CSI \text{ [b/s]}. \quad (1)$$

MOS is a result of subjective or objective methods of speech quality assessment as Bestak, Vranova, and Ondryhal refer [9]. Its values and interpretation are shown in Table 1.

TABLE I. MOS SPEECH QUALITY CLASSIFICATION

MOS	Quality	Note
5	Excellent	Excellent speech quality
4	Good	Quality comparable to analogue phone networks
3	Fair	Fair speech quality
2	Poor	Poor speech quality
1	Bad	Unintelligiblecall

If we use two of the most common codec types, thus G.711 and G.729 codecs, the MOS can reach values up to 4.1 and 3.9.

Note: Due to complexity of nowadays networks, the MOS parameter is no longer sufficient and has been replaced by R-factor, which is more suitable for IP networks. The R-factor is the output of the E-model (according to ITU-T G.107).

In the next step, it is needed to consider a possibility of application of header compression mechanisms or Voice Activity Detection (VAD) techniques. Header compression is possible only when the point-to-point communication occurs. By default, the use of the efficient compression of IP/UDP/RTP headers is used according to Casner and Jacobson in RFC 2508 [10] and Bormann, et al. in RFC 3095 [11]. It reduces the headers up to 2 or 4 bytes. It is not suitable to use this mechanism directly over the Ethernet. However, we can use encapsulation using a Point-to-Point Protocol (PPP) according to Simpson [12]. PPP provides standard services for transmission of different formats of datagrams through serial lines (RFC 1661). PPP is one of the Layer 2 protocols which can also compress the headers in some cases.

VAD is a mechanism that certainly saves bandwidth. However at the same time, it places greater demands on implementation in the terminal equipment. In many cases, the active VAD can result in losses of beginnings of call segments, which greatly decrease quality of voice services. For more complex services, the VAD implementation can even bring errors and cause problems with operation of more complex services. We had such specific experience when we deployed faxes in the field VoIP network.

Another factor that must be considered when optimizing the size of the payload is loading of a router. A large number of short packets are right for real-time applications with respect to small delay but it increases an overhead and computational demands imposed on the routers. A large number of shorter packets and necessity of their processing in nodes paradoxically leads gradually to an increase of the total delay in transmission, in the next step to packet losses and thus to further deterioration of the voice quality and to the flooding of the nodes and a link failure.

Other parameters describing transmission conditions in the particular VoIP network are Voice Payload Size (VPS) (By), Voice Payload Interval (VPI) (ms) and Packets Per Seconds (PPS). The voice payload size represents the number of bytes (or bits) that are filled into a packet. The voice payload size must be a multiple of the codec sample size. The voice payload interval is processed length of a voice segment inserted into one sent packet. PPS represents the number of packets that need to be transmitted every second in order to deliver the codec bit rate. Just these parameters can be varied to optimize network throughput and the voice quality, respectively to look for the best compromise with regard to specifics of the analysed network.

With regard to the facts mentioned above, there can be defined initial conditions for optimization of transfer characteristics of the analysed field network as follows: the mechanism of headers compression is not used, VAD is set, the G.729 codec is used. A computing power of routers with regard to the adequate field network dimensioning is not discussed. A weak point of the field network is the E1 lines. Of possible solutions, we focus on the influence of payload size on the behaviour of the analysed network in our work. We will examine effectiveness of transfer on Layer 3 with respect to the payload size. For now, we will not consider the influence of Layer 2 to which is applied the High Level Data Link Control (HDLC) protocol. Whether the results are relevant even after consideration of the characteristics of HDLC protocol will be verified in the next part of the project.

A. *Characteristics of the G.729codec*

The G.729 codecs are often deployed in VoIP networks in order to save the bandwidth. For example, Cisco technology uses them as a default set in the outgoing direction from the network. In order to save the bandwidth, the G.729 codec is also used as a default set in the analysed army field network.

The G.729 codec is specified in ITU-T Recommendation G.729 [13]. Coding of speech is at 8 kbit/s using Conjugate Structure – Algebraic Code Excited Linear Prediction (CS-ACELP). A reduced – complex vision of the G.729 algorithm is specified in Annex A to Rec. G.729. The speech coding algorithms in the main body of G.729 and in G.729 Annex A are fully interoperable with each other. An implementation that signals or accepts use of the G.729 payload format may implement either G.729 or G.729A. The G.729 and G.729 Annex A codecs were optimized to represent speech with high quality where G.729 Annex A transmits some speech quality for an approximate 50% complexity reduction. For all data rates, the sampling frequency (and RTP timestamp clock rate) is 8000 Hz. The Voice Activity Detector (VAD) and the Comfort Noise Generator (CNG) algorithm in Annex B of G.729 are recommended for digital simultaneous voice and data application. An example of G.729 and G.729A bit packing according to ITU-T Recommendation G.729 is in Figure 1.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|L|      L1     |   L2    |    L3   |       P1      |P|    C1   |
|0|             |         |         |               |0|         |
| |0 1 2 3 4 5 6|0 1 2 3 4|0 1 2 3 4|0 1 2 3 4 5 6 7| |0 1 2 3 4|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       C1      |  S1   | GA1 |  GB1  |    P2   |      C2       |
|          1 1 1|       |     |       |         |               |
|5 6 7 8 9 0 1 2|0 1 2 3|0 1 2|0 1 2 3|0 1 2 3 4|0 1 2 3 4 5 6 7|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   C2    |  S2   | GA2 |  GB2  |
|    1 1 1|       |     |       |
|8 9 0 1 2|0 1 2 3|0 1 2|0 1 2 3|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1. Example of G.729 and G.729A bit packing.

The transmitted parameters of a G.729/G.729A are 10 ms frame, consisting of 80 bits. The mapping of these parameters is given on the Figure 1. The bits of each 32-bit word are numbered from 0 to 31, with the most significant bit on the left and numbered 0. The octets (bytes) of each word are transmitted with the most significant octet first.

III. CALCULATION OF ONE CALL BANDWIDTH IN VOIP ENVIRONMENT

In theoretical calculations of bandwidth required for one call, it is necessary to know particular operating technology, including protocols employed in each level of the communication model.

Bandwidth calculation formulas in general [14]:

$$TPS = L2H + OH + VPS \text{ [By]}, \tag{2}$$

$$PPS = CBR / VPS \text{ [1/s]}, \tag{3}$$

$$Bandwidth = TPS * PPS \text{ [b/s]}, \tag{4}$$

where: L2H is Layer 2 header
TPS is Total Packet Size
OH is overhead of IP/UDP/RTP headers.

Since different systems differ in Layer 2 protocols the most frequently, an operational efficiency is evaluated at Layer 3. Here, we are dealing with overhead of IP/UDP/RTP headers. Layer 2 has often a major impact on the efficiency of transmission. We considered the effect of Layer 2 in the previous work [15], when we analysed the effect of network load on the voice quality. We interconnected two LAN networks using 2811 Cisco routers. The routers were connected through the Catalyst 2950 Switch via Ethernet in a mode of full duplex at 10 Mb/s. This link represents a critical point of the network because of its slow speed. We calculated the theoretical value of how many simultaneous calls can be transferred over a tested line, so we needed to know the bandwidth required for one call.

In this case, we used (see Figure 2):

- G.711 codec
- Sampling frequency = 8000 Hz
- VPI = 30 ms
- VPS = 240 By
- Layer 3 headers = IP header 20 By + UDP header 8 By + RTP header 12 By = 40By
- Layer 2 (Ethernet) header = 38 By (8 By of preamble, 14 By of the source and destination MAC addresses and the length, 4 By of CRC, 12 By of interframe gap).

Note: Possibility of VAD usage and header compression is not considered.

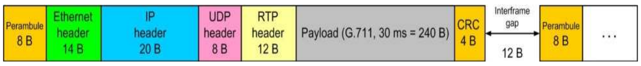

Figure 2. Necessary overhead for transmissionof one 30 ms voice segment.

According to (2), (3) and (4), we need bandwidth for transmission of one voice channel:

TPS = (38 + 40 + 240) = 318 By,

PPS = 64000 / 240 * 8 = 33.333 1/s,

One call bandwidth = 318 * 8 * 33.333 = 84800 b/s.

We get interesting results to the effectiveness of bandwidth usage when using the G.729 codec and the same VPI = 30 ms, so VPS = 30 By:

TPS = (38 + 40 + 30) = 108 By,

PPS = 8000 / 30 * 8 = 33.333 1/s,

One call bandwidth = 108 * 8 * 33.333 = 28800 b/s.

Thus by changing of the codec type will be saved about 2/3 of the bandwidth with slight deterioration of the voice quality (measured by MOS or R - Factor).

IV. RESULTS OF VOICE PAYLOAD SIZE OPTIMIZATION IN FIELD NETWORK CONDITIONS

In the previous section, we set the initial conditions for the theoretical part and following experiments. Relations (2-4) allowing calculation of required bandwidth for a single call at Layer 3 are the starting point. VPS is one of variable parameters. Cisco routers allow setting of different values of VPS within 10 to 240 ms [10]. Figure 3 shows the dependence of the required bandwidth for one call with the G.729 codec on selected VPS.

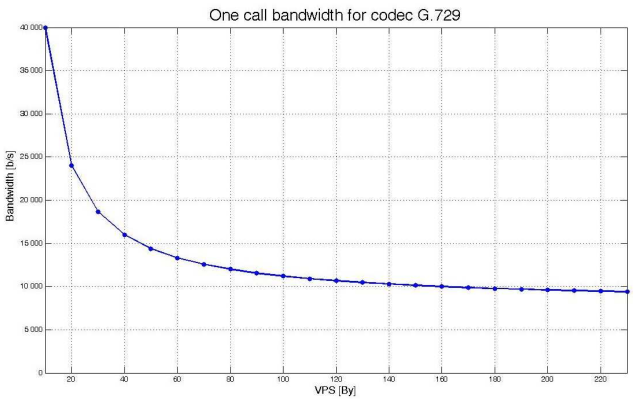

Figure 3. Dependence of one call bandwidth on VPS – G.729 codec.

The curves are not linear. Comparing low values of VPS to higher ones, it is evident that the significantly larger bandwidth to transfer one call is needed for low values. From the values of VPS = 100 ms, use efficiency of transmission line is almost the same. In this case, the required bandwidth for one call ranges around value of 11,2 kbit/s.

The default VPS value for Cisco equipment is VPS = 20 ms which corresponds to the bandwidth = 24 kbit/s. Thus, an appropriate setup of VPS can significantly affect the level of use of the transmission line.

Dependence of number of packets generated by end-terminal on the value of VPS is shown on Figure 4.

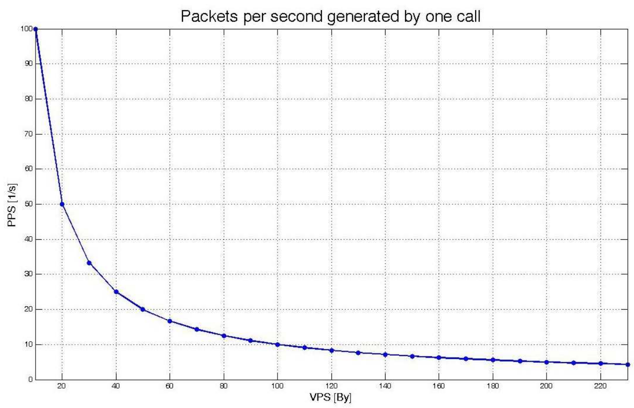

Figure 4. Dependence of number of sent packets on VPS.

Again, dependence is not linear. The VPS increase causes a significant reduction in the computational load of network nodes. As stated previously, we solve two contradictory requirements, saving of the bandwidth and minimum delay to obtain the good voice quality. It is logical that increase of the VPS takes effect in the increase of round trip delay. Round trip delay consists of the sum of the specific delays. VPS resize directly affects two of them and that packet delay and delay in buffers of end-terminals. (The buffers eliminate jitter.)

The VPS increase of 10 By corresponds to the VPI change to 10 ms. Therefore, delay in one channel (packet delay and delay in buffers) is increased by 20 ms, in the loop by 40 ms. G.114 recommendation sets the maximum round trip delay = 300 ms. The increase by 40 ms for each VPS increase by 10 By cannot be ignored. The dependence of rise of round trip delay on VPS is shown on Figure 5.

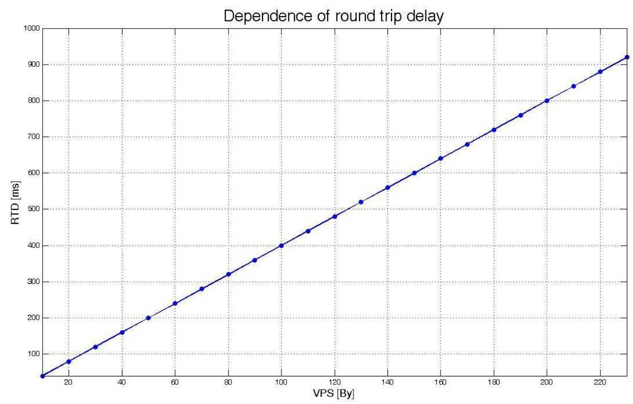

Figure 5. Dependence of rise of round trip delay on VPS.

With regard to the obtained results, we recommend choosing a compromise between the size of the VPS and the considered delay as follows: VPS = 60 By saves bandwidth by 80% with predictable delay in the loop = 240 ms. Therefore, we keep a sufficient reserve for possible increase of delay in the nodes and transmit delay.

V. CONCLUSION AND FUTURE WORK

When analysing the field network, we focused on its weak points. One of them is the bandwidth of E1 links. Although, the network uses the VAD mechanism and efficient G.729 codec to reduce the load generated by voice streams, we were looking for other options to streamline the voice traffic. We focused on verification of bandwidth depending on the voice payload size. From the calculations, it is clear that for small values of VPS, operation is very inefficient. In the next step, we evaluated a relation between VPS and round trip delay. We recommended VPS = 60 ms as the optimal value with respect to the G.114 Recommendation. Bandwidth saving up to 80 % can be done by this setting compared to the VPI default value = 20 ms. All consideration were carried out only at Layer 3 protocols.

Since transmission protocols of the Layer 2 have a significant impact on efficiency, we will focus our future work on this issue. In particular, we will test the operation of the field network for HDLC and PPP protocols.

REFERENCES

[1] J. Postel, "Internet protocol," Network Working Groupe, Request for Comments: 791, September. 1981.

[2] "Cisco Unified Communications Manager Express System Administrator Guide," Cisco, Document ID: OL-10663-03, March. 2013.

[3] "Physical/electrical characteristics of hierarchical digital interfaces," ITU-T, G.703, November. 2001.

[4] S. Choudhury and J. D. Gibson, "Payload length and rate adaptation for multimedia communications in wireless LANs," in IEEE Journal on Selected Areas in Communications, vol. 25, no. 4, May. 2007, pp. 796–807. ISSN 07338716.

[5] S. Vijay Bhanu, RM. Chandrasekaran, and V. Balakrishnan, "Effective Bandwidth Utilization in IEEE802.11 for VOIP," in Internation Journal of Computer Science and Information Security, vol. 8, no. 1, April. 2010, pp. 68-75. ISSN 1947-5500.

[6] Y. Yoon, S. Park, H. Lee, J. S. Kim, and S. B. Jee, "Header compression method and its performance for IP ovet tactical data link," in International Journal of Energy, Information and Communications, vol. 2, issue 2, May. 2011, pp. 61-77. ISSN 2093-9655.

[7] G. Epiphaniou, "Iterative Block Ciphers' Effects on Quality of Experience for VoIP Unicast Transmissions under Different Coding Schemes," University of Bedfordshire, Luton, UK, 2010 (Thesis).

[8] Z. Becvar, M. Vondra, and L. Novak, "Assessment of Speech Quality in VoIP," In VoIP Technologies. Rijeka: InTech, 2011, pp. 27-44. ISBN: 978-953-307-549-5.

[9] R. Bestak, Z. Vranova, and V. Ondryhal, "Testing of transmission channels quality for different types of communications technologies," in Communications in Computer and Information Science, vol. 166. 2011, pp. 13-23. ISSN 1865-0929.

[10] S. Casner and V. Jacobson, "Compressing IP/UDP/RTP headers for low-speed serial links," Network Working Groupe, Request for Comments: 2508, February. 1999.

[11] C. Bormann, et al., "Robust header compression (ROHC) : Framework and four profiles: RTP, UDP: ESP, and uncompressed," Network Working Groupe, Request for Comments: 3095, July. 2001.

[12] W. Simpson, "The Point-to-Point Protocol (PPP)," Network Working Groupe, Request for Comments: 1661, July. 1994.

[13] "Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear-prediction (CS-ACELP)," ITU-T, G.729, January. 2007.

[14] "Voice over IP – per call bandwidth consumption," Cisco, Document ID: 7934, February. 2006.

[15] Z. Vranova and A. Mazalek, "Testing of the effect of load on voice quality in VoIP networks," Proc. International Conference on Telecomunication Systems – Modeling and Analysis (SCIS 07), Czech technical university in Prague, May. 2012, pp. 198-205. ISBN 978-0-9820958-6-7.