

SIMULATION OF DIFFERENT ROUTER BUFFER SIZES WHICH INFLUENCES ON VOIP JITTER DELAY WITHIN THE ROUTED NETWORK

Saša Klampfer¹, Jože Mohorko², Žarko Čučej², Amor Chowdhury¹

¹Margento R&D d.o.o., Maribor, Slovenia

²University of Maribor, Faculty of Electrical Engineering and Computer Science, Maribor, Slovenia

Key words: Jitter, VoIP, delay, queuing discipline, buffer, QoS, voice quality

Abstract: This paper is continuation of /1/ where we tested different queuing methods and their influence on VoIP Jitter delay. Because we combine different queuing disciplines /1/ has also waiting queues be combined, in many cases even doubled. Such method has circumstances in case of VoIP round trip delay and also in case of packet delay variation where they rapidly increase. Such manner is not acceptable in VoIP case where we require small delays in all aspects. This is the reason why we in this paper presented aspects, how to reduce before mentioned delays. One approach is connected with reducing buffer length but anyway it mustn't be too short. The proper solution will also be presented in continuation.

Simulacija vpliva različnih velikosti pomnilnika usmerjevalnikov na zakasnitev VoIP prometa znotraj usmerjanega omrežja

Ključne besede: Jitter, VoIP, zakasnitev, mehanizmi uvrščanja, pomnilnik, kvaliteta storitev

Izvilleček: Članek predstavlja nadaljevanje raiskav, ki smo jih objavili v /1/, kjer smo preizkušali različne metode uvrščanja ter njihov vpliv na VoIP zakasnitev ter vpliv na Jitter. Ker smo v /1/ kombinirali različne mehanizme uvrščanja, so bile s tem kombinirane tudi čakalne vrste, v nekaterih primerih pa so bile celo podvojene. Pristop združevanja mehanizmov uvrščanja in s tem tudi čakalnih vrst ima lahko predvsem vpliv na čas obhoda VoIP paketa in na variacijo zakasnitev, kjer se le te lahko opazno povečajo. Takšni scenariji v omrežjih s časovno občutljivimi aplikacijami niso dobrodošli, še posebej ne v primeru VoIP prometa, kjer težimo k čim manjšim zakasnitvam v vseh pogledih. To je pa hkrati tudi glavni razlog, zakaj smo se v tem delu odločili prikazati pristop, kako zmanjšati tovrstne zakasnitve pri kombiniranju mehanizmov uvrščanja. Eden izmed takšnih pristopov je zmanjšanje velikosti pomnilnika uporabljenega pri čakalnih vrstah. Pri zmanjševanju pomnilnika je potrebno paziti tudi na minimalno dolžino, saj se lahko zaradi premajhnega pomnilnika zavrže večje število paketov, kar pa je v nasprotju z našimi zahtevami in željami. Primerno rešitev predstavljamo v nadaljevanju tega prispevka.

1. Introduction

DEMAND upon VoIP quality speech has been expressed from VoIP vendors long time ago. Since then have been presented many solutions, some of them concerned improved voice encoder schemes, other are connected with queuing schemes and third one are connected with buffer length. The last mentioned also represents essential element of this research where we try to prove how buffer length influence on VoIP delay within the network. Definition says that delay is caused when packets of data (voice) take more time than expected to reach their destination. This causes some disruption in the voice quality. However, if it is dealt with properly, its effects can be minimized. When packets are sent over a network towards a destination machine/phone, some of them might be delayed. Reliability features in the voice quality mechanism sees to it that a conversation is not deadlocked waiting for a packet that went to have a walk somewhere in the green. In fact, there are many factors affecting the journey of packets from source to destination, and one of them is the underlying network. The delayed

packet may come late or may not come at all, in last case it is lost. QoS (Quality of Service) considerations for voice are relatively tolerant towards packet loss, as compared to text. If you lose a word or a zero in your balance, your text might mean something completely different! If you lost a "hu" or a "ha" in a speech, it does not make a really big impact, except some hitch in voice quality. Besides, voice smoothing mechanism regulates it, so that you don't feel the bump. When a packet is delayed, you will hear the voice later than you should. If the delay is not big and is constant, your conversation can be acceptable. Unfortunately, the delay is not always constant, and varies depending on some technical factors. This variation in delay is called jitter, which causes damage to voice quality. That could such damage be minimal we make simulations where results proves how buffer length influences to VoIP jitter.

Used queuing scheme in simulation is described in the second section, meanwhile the third section presents VoIP jitter and his causes. Fourth section presents used G. 729 voice encoder scheme, fifth section presents the OPNET

Modeler simulation tool, which is used for proving our assumptions and decisions. In sixth section is presented simulated wired network structure, meanwhile obtained simulation results are presented in seventh section. Finally section concludes the paper.

2. Used queuing discipline introduction

Weighted fair queuing (WFQ): In situations where we want to provide constant response time and keep delays within range, without assigning an excessive bandwidth, we require optimal solution, so called weighted fair queuing. WFQ is an algorithm/mechanism which introduces bit-wise fairness and allows each queue to be served fairly (Figure 1). Fairness is provided by a mechanism which counts bytes. For the simplest example, we can observe two queues of the same length. The first contains 100 packets and the second only 50. In this case WFQ will work in the following manner; firstly it will take two packets from the first queue then one packet from the second queue and keep repeating this until fairness is achieved.



Fig. 1: *Weighted fair queuing (WFQ) discipline*

Such an algorithm creates service fairness for each participating queue. Low level priority traffic cannot then disturb travel through the network, which is a good compromise for each participating traffic flow within the network. WFQ also has other benefits such as configuring cost minimization, because it is capable of automatically accommodating dynamic network changes. Because of its good qualities, it is used on the majority of serial interfaces configured for E1 operation speeds (2048 Mbps). Weight is also, in this case, defined using the IP priority amount defined in the ToS field of the IP packet. For IP priorities, settings within the range from 0 (best effort) to 5 (IP quality speech) are in use, whilst meanwhile settings 6 and 7 are reserved. The algorithm then uses these data to calculate, how many additional services need to be provided for predicting each individual queue. It can use each available bandwidth that could interfere with low-priority traffic flows, in an order that does not present high-priority traffic flow. In the basis is this principle, an opposite comparison with time-division multiplexing (TDM) which reserves all available bandwidths and leaves it unused if the traffic flow is not presented. WFQ is suitable for operating with IP priority settings, as is resource reservation protocol (RSVP), and is also capable of managing round-trip delay problem. Such queuing clearly improves an algorithms' SNA, logical link control (LCC) and transmission control protocol (TCP) and is, at the same time, capable of accelerating slow features and remov-

ing congestion within the network. Results become more predictable over the whole routing path, meanwhile delays can be multiplied reduced/decreased in comparison with other queuing disciplines (CQ, PQ, FIFO) /2, 3/.

3. VoIP Jitter and his causes introduction

Jitter is a variation in packet transit delay caused by queuing, contention and serialization effects on the path through the network. In general, higher levels of jitter are more likely to occur on either slow or heavily congested links. It is expected that the increasing use of "QoS" control mechanisms such as class based queuing, bandwidth reservation and of higher speed links such as 100 Mbit Ethernet, E3/T3 and SDH will reduce the incidence of jitter related problems at some stage in the future, however jitter will remain a problem for some time to come. In continuation we will introduce the main causes of jitter /4/. In particular, three types of jitter are well-known where belongs 'Constant Jitter (CJ)' marked as group 'A', and which is in correlation with packet to packed delay variation. The second type of jitter marked as group 'B' is 'Transient Jitter (TJ)' which is in connection with substantial incremental delay which could be incurred by a single packet. The last one, marked as group 'C' is short term delay variation which is usually commonly associated with congestions and route changes. Into group 'A' belongs the following causes; load sharing among multiple access links or IP service providers, load sharing within an IP service, and internal load sharing within routers. Into group 'B' belongs such causes which influences on 'TJ' jitter, and those are: sending system packet scheduling, LAN congestions, router firewall, routing table updates, route flapping and timing drift. Into last group belong access link congestion and router firewall conditionally. What is more important in our investigation is Jitter Buffer, respectively buffer length influence on such jitter.

A jitter buffer is designed to remove the effects of jitter from the decoded voice stream, buffering each arriving packet for a short interval before playing it out. This substitutes additional delay and packet loss (discarded late packets) for jitter. A fixed jitter buffer maintains a constant size whereas an adaptive jitter buffer has the capability of adjusting its size dynamically in order to optimize the delay/discard tradeoff. Both fixed and adaptive jitter buffers are capable of automatically adjusting to changes in delay. For example if a step change in delay of 20 milliseconds occurs then there may be some short term packet discards resulting from the change however the jitter buffer would be quickly realigned. In many cases the jitter buffer can be considered as a time window with one side (the early side) aligned with the recent minimum delay and the other side (the late side) representing the maximum permissible delay before a packet would be discarded. Another approach could be used with modeling the jitter buffer, named jitter buffer emulation. In this case we directly determine how many packets would be discarded as a result of jitter. This

has the advantage of being able to directly observe the time distribution of discard events and eliminates the step of trying to relate a jitter metric to a discard rate. If used as a jitter metric this would require a standardized jitter buffer to be used as proposed in /4, 7, 8, 9, and 10/.

4. G. 729 voice encoder scheme

The G.729 speech coder is an 8 kbps Conjugate-Structure Algebraic-Code-Excited Linear Prediction (CS-ACELP) speech compression algorithm approved by ITU-T. G.729. It offers high quality, robust speech performance at the price of higher complexity. It requires 10 ms input frames and generates frames of 80 bits in length. Within the G.729 coder the processing signals are 10 ms frames with additional 5 ms look-ahead. The total delay introduced by algorithm is 15 ms. Since G.729 is based on the Code-Excited Linear Prediction (CELP) model, each produced 80 bit frame, contains linear prediction coefficients, excitation code book indices, and gain parameters that are used by the decoder in order to reproduce speech. The input/output of this algorithm is 16 bit / 8 kbps linear PCM audio stream that is converted from/to a compressed data stream /5/.

5. Opnet modeler

OPNET Modeler's cutting-edge technology provides an environment for designing protocols and technologies as well as testing and demonstrating designs in realistic scenarios prior to production. OPNET Modeler is used by the world's largest network equipment manufacturers to enhance the design of network devices, technologies such as VoIP, TCP, OSPFv3, MPLS, IPv6, etc. Modeler represents a leading simulation tool used for modeling, and simulating communication networks, and their applications. At the same time it offers study analyze of telecommunication infrastructures, individual apparatus, protocols, and applications. This is a graphically-oriented simulation tool, which uses project, node and process editors for building communication models. The project editor offers graphic topological description, whilst the node editor is used for describing protocols. The process editor is an upgrade of C language, which uses a finite-state machine for algorithm and protocol descriptions. OPNET includes many standard communication models for constructing wired, radio, and optical and satellite communication structures. This tool also offers 2D and 3D animation for accompanying changes within the network's structure /6/.

6. Simulation structure

Network topology used in our simulation scenarios is shown in Figure 2. It represents only a test board structure for proving our assumptions about buffer length influence on jitter delay. Our main goal in these simulations is oriented towards improving the network's performances, regarding to the VoIP jitter.

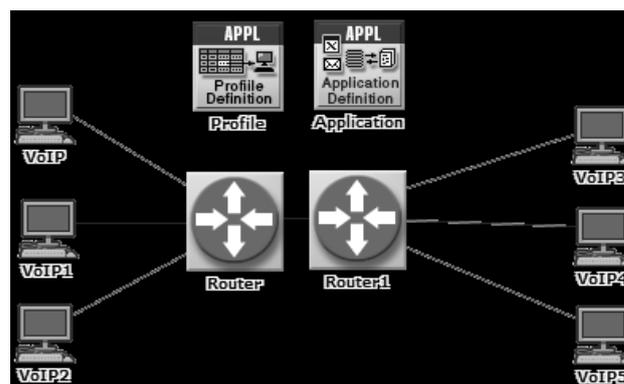


Fig. 2: Network simulation structure

The network structure consists of two sixteen port routers mutual connected where are on each router connected three VoIP users, commonly six, with generated 56 full meshed VoIP traffic flows. All clients are connected over 10BaseT link on routers, where are also routers mutual connected with the same connection type. VoIP application is defined using the 'Application' node shown on the top-right side in Figure 2. With 'Profile' node (beside applications node) are defined a client profiles, respectively what should User Equipment (UE) do, or better say, which application is such UE capable to use.

Table 1: Number of users

Application	Users	Class	Bandwidth
VoIP	6	1	10 Mbit/s

Simulation stricter is very simple because we only want to show how important part plays proper buffer length on voice quality in sense of jitter etc. For that case we prepare two different scenarios. In first is buffer length set to 256 MB (Mbytes). Such buffer represents the memory used to store packets awaiting processing or currently processed by the forwarding CPU. In a centralized scheme, processed packets remain in that memory until sent out on an appropriate interface. In the distributed scheme, processed packets are then forwarded to the appropriate slot memory. In second scenario is such buffer reduced on only 16 MB which contribute better results presented in continuation.

7. Simulation results

As we mentioned above, we use two different scenarios with the same network structure and different buffer lengths. We also try to figure out which length is the proper one, respectively which gives better delay results.

The minimal value which can be chosen in simulation tool is 8MB. For tested behavior of such value we again created new scenario which has been after successfully executed simulation compared with other two scenarios. The analysis shows that decreasing buffer length less than 16MB does not effect on delay parameters within the network, and at

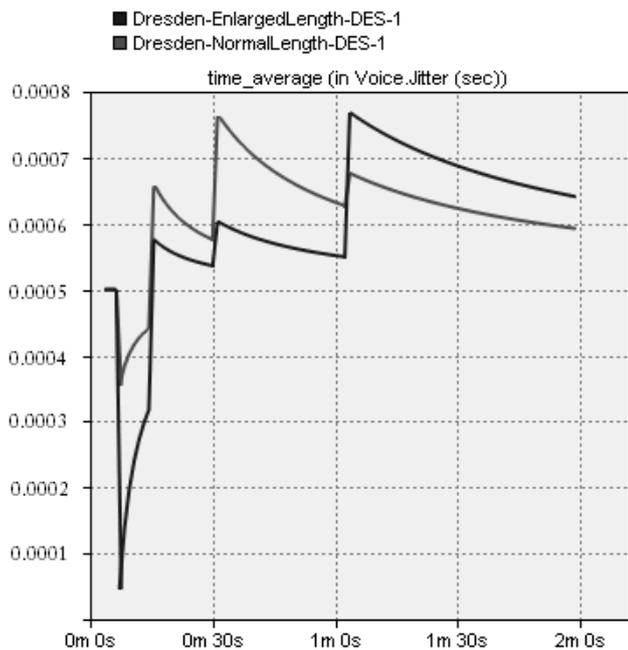


Fig. 3: Voice Jitter of VoIP application transmitted through two different buffer sizes (16MB buffer length -red, and 256MB buffer length - blue) in time average statistic

the same time, results in case of use 8MB buffer are quite similar to results when 16MB buffer is used. Comparison between 8MB and 16MB buffer length and their effect to Voice jitter is shown in figure 3. As we expected, large buffer increase voice jitter in second half of simulated period. Within first half of period largest buffer gives better results compare smaller buffer, because we have in first half of period smaller traffic compare with second half of simulation period. Since buffer is full, voice jitter rapidly increases as is shown in Figure 3 and then such approach fails, because VoIP traffic requires minimal delay within the network. Such jitter for both cases is fully acceptable, because highest value didn't reach 0.8ms, but is clearly to see affect of buffer length.

The end-to-end delay for both scenarios is presented in Figure 4. The end-to-end delay is thus the sum of the delays experienced at each hop on the way to the destination. Each such delay in turn consists of two components, a fixed component which includes the transmission delay at a node and the propagation delay on the link to the next node, and a variable component which includes the processing and queuing delays at the node.

In our case is such delay minimal, but situation can be changed using different simultaneous applications within the same network where can congestions appear and so on. Our simulated network has enough available bandwidth to transmit all VoIP traffic between communicating nodes, and this is the main reason why is end-to-end, respectively packet delay so small. It takes maximal value around 60,53ms for large buffer and around 60,22ms for small buffer.

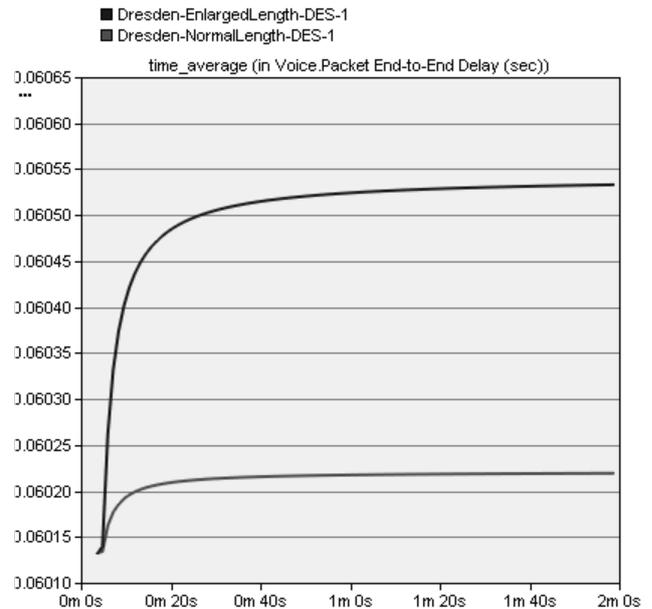


Fig. 4: Voice Packet End-to-End delay for two different buffer sizes (16MB buffer length - red, and 256MB buffer length - blue) in time average statistic

Packet Delay Variation (PDV) is defined by RFC 3393 as the difference in end-to-end delay between selected packets in a flow with any lost packets being ignored. The means of selection is not specified in RFC 3393, but e.g. could be the packets which gave the biggest variation in delay in a selected time period. The delay is specified from the start of the packet being transmitted at the source to the end of the packet being received at the destination. A component of the delay which does not vary from packet to packet can be ignored, hence if the packet lengths are the same and packets always take the same time to be re-assembled at the destination then the packet arrival time at the destination

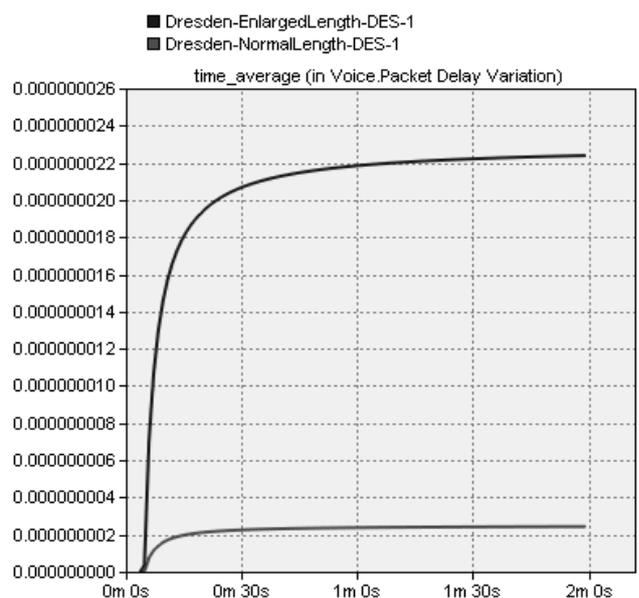


Fig. 5: Voice packet delay variation for two different buffer sizes (16MB buffer length - red, and 256MB buffer length - blue) in time average statistic

could be used instead of the time the end of the packet is received. From obtained results presented on Figure 5 we can assume that is before mention difference for PDV on minimal level, because maximal value don't reach more than 2,3 nano-seconds.

From above Figure we can assume that has been enough bandwidth present within the network for VoIP application transmission and because of that is data loss on minimal level.

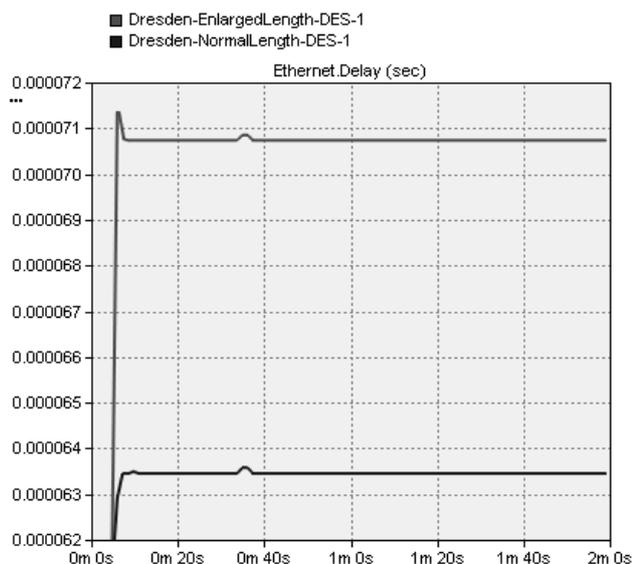


Fig. 6: Ethernet delay for two different buffer sizes (16MB buffer length - red, and 256MB buffer length - blue) in time average statistic

Also in case of Ethernet delay are differences between large and small buffer noticeable but they are on 7.5 us level. Such delay is in practice imperceptible.

8. Conclusion

During this paper we show the reader how one of few possibilities affect on jitter in VoIP case, and how such solution affect on other parameters within the network, such as end-to-end delay and packet delay variation. From simulation results is clearly to see buffer length influence on VoIP traffic delay. As we mentioned in second section, such jitter can be additionally reduced with use of jitter buffer emulation, fixed jitter buffer and adaptive jitter buffers. This is also the issue for further research analysis on this area.

References

- /1./ Saša Klampfer, Jože Mohorko, Žarko Čučej, "Influence of hybrid queuing disciplines on VoIP traffic delay", InfoCOM international conference
- /2./ Saša Klampfer, Jože Mohorko, Žarko Čučej, "IP packet queuing disciplines as basic part of QoS assurance within the network", Informacije MIDEM, junij 2009, letn. 39, št. 2(130)
- /3./ Internetworking Technology Handbook - Quality of Service (QoS), Cisco Systems
- /4./ Cole, R. Rosenbluth J. Voice over IP Performance Monitoring, AT&T Preprint September 2000

- /5./ G. 729 Data Sheet
- /6./ OPNET Modeler Technical Documentation
- /7./ TIPHON 22TD047 Problems with the behavior of Jitter Buffers and their influence on the end-to-end speech quality, source KPN Research, March 2001
- /8./ RFC1889 Real Time Control Protocol
- /9./ ITU-T Y.1541 Network Performance Objectives for IP Based Services
- /10./ ITU-T SG12 D74 IP Phones and Gateways: Factors impacting speech quality, France Telecom, May 2002

Saša Klampfer has graduated in 2007, at the Faculty of Electrical Engineering and Computer Science in Maribor with diploma degree thesis: »Network simulations in OPNET Modeler«. In 2009 he finish master degree study programe with masters thesis entitled: »Expert system for automatic analysis of tactical radio network properties« at the same Faculty. His research activity is focused on the simulation areas and on design of telecommunications (UMTS) and computer networks (Fast ethernet, Wi-Fi) and on the influence how individual applications affects on specific network in terms of time, utilization, packet loss, etc.
sasa.klampfer@margento.com

Jože Mohorko received his B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the University of Maribor, Slovenia, in 1990, 1994 and 2002, respectively. He is a docent and researcher at the Faculty of Electrical Engineering and Computer Science in Maribor, University of Maribor. His research interests include image processing, modeling, simulating and evaluating telecommunication infrastructures.

Žarko Čučej is a full professor for automatic control and robotics, and telecommunications at the Faculty of Electrical Engineering and Computer Science of the University of Maribor. His recent research interests include signal processing and industrial data networks.

Amor Chowdhury has graduated in 1994, at the Faculty of Electrical Engineering and Computer Science in Maribor, from the discrete regulation area with diploma degree thesis entitled: »Discrete regulation systems design on personal computer«. In 1997 he receives masters degree at the same faculty with masters degree thesis entitled: »The theory of robust synthesis«. Four years later (2001) he has complited Phd. study programe with Phd. Thesis entitled: »Regulation systems robust synthesis including performance criteria«. He is a great innovator and researcher what confirms numbers of given and registered patents, for example 11643 Terminal Device for Monitoring, Maintenance and Service Data patent.