

PWE3
Internet-Draft
Expires: April 19, 2004

Y(J) Stein
I. Druker
RAD Data Communications
October 20, 2003

The Effect of Packet Loss on Voice Quality for TDM over Pseudowires
draft-stein-pwe3-tdm-packetloss-01.txt

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of [Section 10 of RFC2026](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/1id-abstracts.txt>.

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

This Internet-Draft will expire on April 19, 2004.

Copyright Notice

Copyright (C) The Internet Society (2003). All Rights Reserved.

Abstract

The effect of packet loss on voice quality has been the subject of detailed study in the VoIP community, but these results are not directly applicable to speech channels carried in TDM pseudowires, as being studied in the PWE WG. The present document presents an analysis of packet loss for the TDM over PW case, and demonstrates that packet loss of a few percent can be tolerated when appropriate packet loss concealment techniques are employed.

Internet-Draft

PWE3 TDM Packet Loss

October 2003

Table of Contents

1.	Introduction	3
2.	TDM Pseudowires	4
3.	Effect of Packet Loss on TDM Pseudowires	5
4.	Measures of Voice Quality	6
5.	Packet Loss Replacement Algorithms	7
6.	Experimental Results	8
7.	Discussion	10
8.	References	12
	Authors' Addresses	13
	Full Copyright Statement	14

1. Introduction

There are several sources of packet loss in PSNs. Packets are discarded upon detection of bit errors, but with modern fiber optic technology such errors are rare in core networks. Routers must drop packets when congested, and may do so when they sense congestion is imminent. Real-time streams may have an additional source of packet loss, namely rejection of a packet that has successfully arrived at the destination, but has been overly delayed. Non-real-time data communications are not overly effected by packet loss, due to the possibility of retransmission; but real-time constraints usually prohibit retransmission, and hence packet loss leads to noticeable quality degradation.

Packet loss in voice traffic can cause in gaps or artifacts that result in choppy, garbled or even unintelligible speech. Market acceptance of TDM transport over pseudowires will depend on service providers being able to offer meaningful voice quality guarantees, while deploying networks with some reasonable amount of packet loss. Hence packet loss concealment (PLC) mechanisms may need to be employed.

We study here the effect of packet loss on the perceived quality of speech occupying a timeslot in a TDM bitstream that is transported via a structured TDM pseudowire. In [Section 2](#) we briefly explain TDM emulation, and in [Section 3](#) we survey known results regarding the effect of packet loss on VoIP and TDM pseudowires. [Section 4](#) elucidates voice quality measurement, while [Section 5](#) suggests several packet loss concealment algorithms for the TDM case. In [Section 6](#) we outline the numeric results of a few experiments we have carried out, the consequences of which are discussed in [Section 7](#).

2. TDM Pseudowires

The public telephone system uses TDM (e.g. T1, E1) to carry multiple telephone-quality audio channels. Since TDM networks dedicate highly synchronous circuits to voice calls, there is never packet loss, and even individual bit slips are tightly controlled. Telephony customers have grown accustomed to telephone service quality, and are not amenable to lower quality unless there are other advantages (e.g. mobility or significantly lower price).

TDM bitstreams may be transported over packet-switched networks via structure-agnostic [[SAToP](#)] or structure-aware [[TDMoIP](#),[CESoPSN](#)] pseudowires. As discussed in the introduction, packet loss is to be expected in any packet switched network; however, its effect on most data traffic is minimal since retransmission mechanisms compensate for it with no ill effects other than a reduction in effective data transfer rate. Unfortunately, real-time traffic such as TDM can not tolerate the added latency incurred by retransmission. TDM pseudowires will thus suffer from packet loss in the underlying PSN and the telephony channels will accordingly be of lower perceived quality.

Interworking devices based on structure-agnostic techniques are inherently unaware of the individual telephone channels, and are thus limited to simplistic treatment of packet loss, such as replacing all missing bits with ones. Structure-aware emulation is intrinsically more robust to packet loss as it necessarily reconstitutes the TDM framing, and in addition this knowledge of frame structure makes

possible more sophisticated treatment of packet loss. In the following we shall assume structure-aware emulation is employed.

[3.](#) Effect of Packet Loss on TDM Pseudowires

The precise effect of packet loss on voice quality, and the development of PLC algorithms have been the subject of detailed study in the VoIP community. Their results can be summarized as follows:

- 1) One percent packet loss causes perceived voice quality to drop from near toll-quality to cell-phone quality.
- 2) Above two percent, packet loss is the dominant cause of voice quality deterioration, compressed and uncompressed speech becoming comparable in quality.
- 3) Packet size is not a significant factor (at least for lengths typically employed in VoIP).
- 4) By using appropriate packet loss concealment algorithms (PLC) five percent packet loss of uncompressed speech can be comparable to cell-phone quality.

These results are not directly applicable to audio channels in TDM transport. This is because VoIP packets typically contain between 80 samples (10 milliseconds) and 480 samples (60 milliseconds) of the speech signal, while multichannel TDM packets may contain only a single sample, or perhaps a very small number of samples, of each audio channel. PLC for the TDM emulation case is seen to be much more justifiable, since the gaps are always much smaller than speech

events. In contrast, loss of a single VoIP packet, and certainly of several packets, can result in irreparable loss of entire phonemes.

An alternative viewpoint emphasizes that a packet carrying TDM over a PSN contains data from multiple voice channels, as compared with a VoIP packet of similar size that contains audio from a single source. Since TDM emulation has natural data interleaving, each channel is less influenced by loss events.

[4. Measures of Voice Quality](#)

Perceived voice quality is a psychophysical quantity that depends on the physiology and psychology of the listener. The most universally accepted subjective measure of voice quality is the mean opinion score (MOS) defined by the ITU-T for telephone quality speech in [\[P.800\]](#), and by the ITU-R for higher fidelity audio in [\[BS.1116-1\]](#). It is found by averaging the reported opinion scores of multiple listeners, each of whom rates the audio on a five point quality scale, with MOS=1 signifying unintelligibility, and MOS=5 meaning excellent quality. Due to the 4 KHz bandwidth limitation and the logarithmic amplitude characteristics of the 64 Kbps DS0 digital channel, telephony voice is rated lower than 5, with 4 to 4.5 being considered "toll-quality". MOS ratings of 3.5 to 4 are considered

acceptable to many listeners, and cellular telephone audio is deemed acceptable at about MOS=3.5 due to the added convenience of mobility. Speech quality lower than MOS=3 is considered acceptable only for certain applications, such as encrypted military communications.

The problem is that MOS is based on subjective scoring, and so is time consuming and costly to measure. Objective measures, i.e. ones that can be computed by signal processing algorithms based on the signal samples, are preferable if they correlate well with the subjective measures. The ITU-T has standardized two such measures for telephony quality speech, known as PSQM [[P.861](#)] and PESQ [[P.862](#)], while the ITU-R has sanctioned PEAQ [[BS.1387](#)] for higher fidelity radio quality audio. These objective measures utilize models of the biological auditory system and have been shown to correlate well with subjective measurements of MOS.

PSQM was developed for lab comparison of different speech codecs and does not take such factors as delay or packet loss into account. PESQ specifically performs end-to-end speech quality assessment and was therefore chosen for our experiments.

[5](#). Packet Loss Replacement Algorithms

In this section we discuss algorithms for concealing the loss of a packet. For concreteness we will assume in the following discussion that packets carry single samples of each TDM timeslot. The extension to multiple samples is relatively straightforward, and turns out not to drastically change our results.

The simplest ploy to implement is to blindly insert a constant value in place of any lost speech samples. Since we can assume that the input signal is zero-mean (i.e. contains no DC component) minimal distortion is attained when this constant is chosen to be zero. This is in fact precisely what happens when a G.711 mu-law codec receives a word containing all-ones, as would be the case if AIS were to be received (but unfortunately is not the case for A-law).

A slightly more sophisticated technique is to replace the missing sample with the previous one. This method is justifiable in the VoIP case where the quasistationarity of the speech signal means that the missing buffer is expected to be similar to the previous one. Even in the single sample case it is decidedly better than replacement by zero due to the typical low-pass characteristic of speech signals, and to the fact that during intervals with significant high frequency content (e.g. fricatives) the error is less noticeable.

We will declare a packet lost following the reception of the following packet. Hence when loss needs to be concealed, both the sample prior to the missing one, and that following it can be assumed to be available. This enables us to estimate the missing sample value by interpolation, the simplest type of which is linear interpolation, whereby the missing sample is replaced by the average of the two surrounding values. More complex interpolation, such as quadratic interpolation or splines can be used as well, but for the purposes of this analysis we will restrict ourselves to the linear case.

More sophisticated methods of packet concealment are based on model-based prediction. Standardized speech compression algorithms have had integral packet loss concealment methods for some time, and more recently the ITU-T has standardized a packet loss concealment method for uncompressed speech [[G.711App1](#)]. For the purposes of our experiments we need only to estimate the value of a single missing sample (or more generally a small number of missing samples), and so relatively simple modeling is sufficient. We used an interpolation model based on second order statistics of the previous N samples; we call this method Statistically Enhanced Interpolation (STEIN). In the simulations below we took N=30 samples. Details and derivation of this algorithm will be reported elsewhere.

6. Experimental Results

In order to quantify the anecdotal results we have observed in real-world deployments, we have carried out a controlled experiment to measure the effect of packet loss on voice quality. We first describe the methodology we employed.

The speech data was selected from English and American English subsets of the ITU-T P.50 Appendix 1 corpus [P.50App1] and consisted of 16 speakers, eight male and eight female. Each speaker spoke either three or four sentences, for a total of between seven and 15 seconds. The selected files were filtered to telephony quality using modified IRS filtering and downsampled to 8 KHz.

A uniform random number generator was used to generate packet loss. Packet loss of 0, 0.25, 0.5, 0.75, 1, 2, 3, 4 and 5 percent were tested. In the simulations reported here we explicitly disallowed loss of successive packets; bursty packet loss (where the probability of groups of missing samples is much higher than would be expected from the average packet loss rate) was also simulated but is not reported here.

For each file the four methods of lost sample replacement were applied and the PESQ scores evaluated. A graph depicting the PESQ derived MOS as a function of packet loss for the four lost packet replacement algorithms cases is available in ps and pdf formats at <http://www.dspscsp.com/tdmoip/pl.ps> and <http://www.dspscsp.com/tdmoip/pl.pdf> respectively.

We obtained the following qualitative and quantitative results.

- 1) For all cases the MOS resulting from the use of zero insertion is less than that obtained by replacing with the previous sample, which in turn is less than that of linear interpolation, which is slightly less than that obtained by statistical interpolation.
- 2) Unlike the artifacts speech compression methods may produce when subject to buffer loss, packet loss here effectively produces additive white impulse noise. The subjective impression is that of static noise on AM radio stations or crackling on old phonograph records. For a given PESQ, this type of degradation is more acceptable to listeners than choppiness or tones common in VoIP.

Internet-Draft

PWE3 TDM Packet Loss

October 2003

3) If MOS>4 (full toll quality) is required, then the following packet losses are allowable:

zero insertion - 0.05 %

previous sample - 0.25 %

linear interpolation - 0.75 %

STEIN - 2 %

4) If MOS>3.75 (barely perceptible quality degradation) is acceptable, then the following packet losses are allowable:

zero insertion - 0.1 %

previous sample - 0.75 %

linear interpolation - 3 %

STEIN - 6.5 %

5) If MOS>3.5 (cell-phone quality) is tolerable, then the following packet losses are allowable:

zero insertion - 0.4 %

previous sample - 2 %

linear interpolation - 8 %

STEIN - 14 %

[7.](#) Discussion

When structure-agnostic TDM transport is used, the only option for handling packet loss in TDM over PW is to generate Alarm Indication Signal (AIS) whenever a packet is lost. This results in insertion of constant values, which has been seen to result in extremely low tolerance to packet loss.

Structure-aware transport methods, may employ "frame replay", which increases the perceived voice quality and has the added benefit that CAS signaling integrity is guaranteed.

The linear and statistically enhanced interpolation methods can only be employed for structure-aware TDM transport, since only then are the timeslot signal values readily available for manipulation. This rules out unframed transport and non-byte-oriented transport (including some methods of transporting T1 links). In addition, complex encapsulations that impede the extraction of required samples, may hinder the use of these methods.

What is the computational burden of these interpolations? Assuming a processor with hardware companding and that can perform an addition and a shift in a single cycle (e.g. a DSP processor), linear interpolation requires a single cycle per timeslot per sample loss event, or $8000 L$ instruction cycles per second, where L is the packet loss percentage. An entire 30 channel E1 link will thus require $0.24 L$ MIPS, and an entire 24 channel T1 link $0.192 L$ MIPS. For example at 2% packet loss, an average processing power of 1 MIPS will suffice for 208 E1 trunks or 260 T1 trunks. Even using a processor that requires 10 instructions to process an interpolation, dedicating 1 MIPS will enable fixing 20 E1s or 26 T1s.

The statistically enhanced interpolation method requires the computation of energy, single and dual lag autocorrelations, which for a history buffer of N samples involves approximately $3N$ multiplications and additions. For processors that can perform

multiply and accumulate operations in a single cycle (e.g. DSP processors) this translates to $0.024 N L$ MIPS per timeslot ($0.72 N L$ MIPS per E1 or $0.576 N L$ MIPS per T1), when computation is only carried out when needed. Alternatively, the required autocorrelations could be continuously gathered (using telescoping series methods) at the price of three multiply and accumulate operations per input sample, or 0.024 MIP per channel, to which one must add a small amount of additional computation per packet loss event.

The duration over which the autocorrelations are computed must be chosen long enough for the signal statistics to be significant, but

not so long that the statistics would be expected to change significantly during normal speech. Numbers in the range 10 to 100 are reasonable. For example, using $N=30$ and once again assuming 2% packet loss, the processing drain for non-telescoping computation would be 0.432 MIPS per E1 and 0.3456 MIPS per T1.

Although statistically enhanced interpolation is consistently better than simple linear interpolation, the additional MIPS is only be justifiable when the packet loss rate is sufficiently high.

8. References

[BS.1116-1] ITU-R Recommendation BS.1116-1 (1994-1997) Methods for the Subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound

[BS.1387] ITU-R Recommendation BS.1387 (1998) Method for Objective Measurements of Perceived Audio Quality

[CESoPSN] [draft-vainshtein-cesopsn-06.txt](#) (2003) TDM Circuit Emulation Service over Packet Switched Network, A. Vainshtein et al, work in progress

[G.711App1] ITU-T Recommendation G.711 - [Appendix I](#) (1999) A high quality low-complexity algorithm for packet loss concealment with G.711

[P.50App1] ITU-T Recommendation P.50 - [Appendix I](#) (1998) Artificial Voices - Test Signals

[P.800] ITU-T Recommendation P.800 (1996) Methods for Subjective Determination of Transmission Quality

[P.861] ITU-T Recommendation P.861 (1998) Objective Quality Measurement of Telephone-band (300-3400 Hz) Speech Codecs

[P.862] ITU-T Recommendation P.862 (2001) Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrow-band Telephone Networks and Speech Codecs

[SATO] [draft-ietf-pwe3-satop-00.txt](#) (2003) Structure Agnostic TDM over Packet, A. Vainshtein and Y. Stein, work in progress

[TDMoIP] [draft-anavi-tdmoip-05.txt](#) (2003) TDM over IP, Yaakov (Jonathan) Stein et al, work in progress

Stein & Druker

Expires April 19, 2004

[Page 12]

Internet-Draft

PWE3 TDM Packet Loss

October 2003

Authors' Addresses

Yaakov (Jonathan) Stein
RAD Data Communications
24 Raoul Wallenberg St., Bldg C
Tel Aviv 69719
ISRAEL

Phone: +972 3 6455389
EMail: yaakov_s@rad.com

Ilya Druker
RAD Data Communications
24 Raoul Wallenburg St., Bldg C

Tel Aviv 69719
ISRAEL

Phone: +972 3 7657061
EMail: ilya_d@rad.com

Stein & Druker

Expires April 19, 2004

[Page 13]

Internet-Draft

PWE3 TDM Packet Loss

October 2003

Full Copyright Statement

Copyright (C) The Internet Society (2003). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any

kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.