Audio/Video Transport Working Group
Internet-Draft
Intended status: Standards Track
Expires: July 28, 2013

A. Clark
Telchemy
G. Zorn, Ed.
Network Zen
C. Bi
STTRI
Q. Wu, Ed.
Huawei

January 24, 2013

RTCP XR Report Block for Concealment metrics Reporting on Audio Applications
draft-ietf-xrblock-rtcp-xr-loss-conceal-04.txt

Abstract

This document defines two RTCP XR Report Blocks that allows the reporting of loss concealment metrics for audio applications of RTP.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

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."

This Internet-Draft will expire on July 28, 2013.

Table of Contents

1. Introduction

1.1. Loss Concealment and Concealment Seconds Metrics Reporting Block

At any instant, the audio output at a receiver may be classified as either 'normal' or 'concealed'. 'Normal' refers to playout of audio payload received from the remote end, and also includes locally generated signals such as announcements, tones and comfort noise. Concealment refers to playout of locally-generated signals used to mask the impact of network impairments or to reduce the audibility of jitter buffer adaptations.

This draft defines two new concealment related block types to augment those defined in [RFC3611] for use in a range of RTP applications.

The first block type provides metrics for actions taken by the receiver to mitigate the effect of packet loss and packet discard. Specifically, the first metric (On-Time Playout Duration) reports the duration of normal playout of data which the receiver obtained from the sender's stream. A second metric (Loss Concealment Duration) reports the total time during which the receiver played out media data which was manufactured locally, because the sender's data for these periods was not available due to packet loss or discard. A similar metric (Buffer Adjustment Concealment Duration) reports the duration of playout of locally-manufactured data replacing data which is unavailable due to adaptation of an adaptive de-jitter buffer. Further metrics (Playout Interrupt Count and Mean Playout Interrupt Size) report the number of times normal playout was interrupted, and the mean duration of these interruptions.

Loss Concealment Duration and Buffer Adjustment Concealment Duration are reported separately because buffer adjustment is typically arranged to occur in silence periods so may have very little impact on user experience, whilst loss concealment may occur at any time.

The second block type provides metrics for concealment. Specifically, the first metric (Unimpaired Seconds) reports the number of whole seconds occupied only with normal playout of data which the receiver obtained from the sender's stream. The second metric (Concealed Seconds) reports the number of whole seconds during which the receiver played out any locally-generated media data. A third metric (Severely Concealed Seconds) reports the number of whole seconds during which the receiver played out locally-generated data for more than SCS Threshold (ms).

These metrics belongs to the class of transport-related terminal metrics defined in [RFC6792].

1.2. RTCP and RTCP XR Reports

The use of RTCP for reporting is defined in [RFC3550]. [RFC3611] defined an extensible structure for reporting using an RTCP Extended Report (XR). This draft defines a new Extended Report block that MUST be used as defined in [RFC3550] and [RFC3611].

1.3. Performance Metrics Framework

The Performance Metrics Framework [RFC6390] provides guidance on the definition and specification of performance metrics. The RTP Monitoring Architectures [RFC6792] provides guideline for reporting block format using RTCP XR. The Metrics Block described in this document are in accordance with those guidelines.

1.4. Applicability

These metrics are only applicable to audio applications of RTP.

2. Terminology

2.1. Standards Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC2119 [RFC2119].

3. Loss Concealment Block

3.1. Report Block Structure

Loss Concealment metrics block

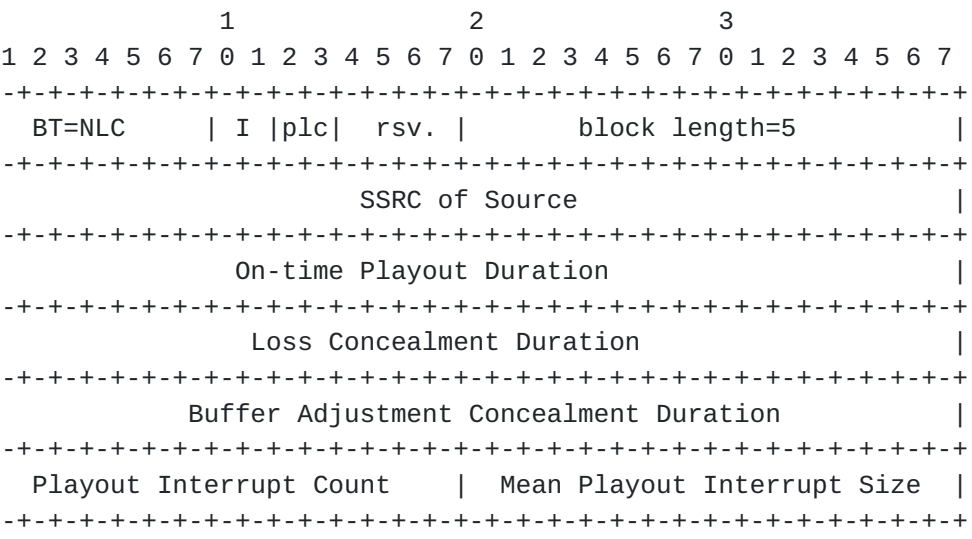

Figure 1: Report Block Structure

3.2. Definition of Fields in Loss Concealment Report Block

Block type (BT): 8 bits

A Loss Concealment Metrics Report Block is identified by the constant NLC.

[Note to RFC Editor: please replace NLC with the IANA provided RTCP XR block type for this block.]

Interval Metric flag (I): 2 bit

This field is used to indicate whether the Loss Concealment metrics are Sampled, Interval or Cumulative metrics:

I=10: Interval Duration - the reported value applies to the most recent measurement interval duration between successive metrics reports.

I=11: Cumulative Duration - the reported value applies to the accumulation period characteristic of cumulative measurements.

I=01: Sampled Value - the reported value is a sampled instantaneous value.

Packet Loss Concealment Method (plc): 2 bits

This field is used to identify the packet loss concealment method in use at the receiver, according to the following code:

bits 014-015

0 = silence insertion

1 = simple replay, no attenuation

2 = simple replay, with attenuation

3 = enhanced

Other values reserved

Reserved (resv): 4 bits

These bits are reserved. They MUST be set to zero by senders and ignored by receivers (See [RFC6709] section 4.2).

block length: 16 bits

The length of this report block in 32-bit words, minus one. For the Loss Concealment block, the block length is equal to 5.

SSRC of source: 32 bits

As defined in Section 4.1 of [RFC3611].

On-time Playout Duration (ms): 32 bits

'On-time' playout is the uninterrupted, in-sequence playout of valid decoded audio information originating from the remote endpoint. This includes comfort noise during periods of remote talker silence, if VAD is used, and locally generated or regenerated tones and announcements.

An equivalent definition is that on-time playout is playout of any

signal other than those used for concealment.

On-time playout duration MUST include both speech and silence intervals, whether VAD is used or not. This duration is reported in millisecond units.

If the measured value exceeds 0xFFFFFFFD, the value 0xFFFFFFFE MUST be reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFFFFFF MUST be reported.

Loss Concealment Duration (ms): 32 bits

The duration, in milliseconds, of audio playout corresponding to Loss-type concealment.

Loss-type concealment is reactive insertion or deletion of samples in the audio playout stream due to effective frame loss at the audio decoder. "Effective frame loss" is the event in which a frame of coded audio is simply not present at the audio decoder when required. In this case, substitute audio samples are generally formed, at the decoder or elsewhere, to reduce audible impairment.

If the measured value exceeds 0xFFFFFFFD, the value 0xFFFFFFFE MUST be reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFFFFFF MUST be reported.

Buffer Adjustment Concealment Duration (ms): 32 bits

The duration, in milliseconds, of audio playout corresponding to Buffer Adjustment-type concealment, if known.

If the measured value exceeds 0xFFFFFFFD, the value 0xFFFFFFFE MUST be reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFFFFFF MUST be reported.

Buffer Adjustment-type concealment is proactive or controlled insertion or deletion of samples in the audio playout stream due to jitter buffer adaptation, re-sizing or re-centering decisions within the endpoint.

Because this insertion is controlled, rather than occurring randomly in response to losses, it is typically less audible than loss-type concealment. For example, jitter buffer adaptation events may be constrained to occur during periods of talker silence, in which case only silence duration is affected, or

sophisticated time-stretching methods for insertion/deletion during favorable periods in active speech may be employed.

Concealment events which cannot be classified as Buffer Adjustment- type MUST be classified as Loss-type.

Playout Interrupt Count: 16 bits

The number of interruptions to normal playout which occurred during the reporting period.

If the measured value exceeds 0xFFFD, the value 0xFFFE MUST be reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFF MUST be reported.

Mean Playout Interrupt Size (ms): 16 bits

The mean duration, in ms, of interruptions to normal playout which occurred during the reporting period.

If the measured value exceeds 0xFFFD, the value 0xFFFE MUST be reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFF MUST be reported.

4. Concealment Seconds Block

This sub-block provides a description of potentially audible impairments due to lost and discarded packets at the endpoint, expressed on a time basis analogous to a traditional PSTN T1/E1 errored seconds metric.

The following metrics are based on successive one second intervals as declared by a local clock. This local clock does NOT need to be synchronized to any external time reference. The starting time of this clock is unspecified. Note that this implies that the same loss pattern could result in slightly different count values, depending on where the losses occur relative to the particular one-second demarcation points. For example, two loss events occurring 50ms apart could result in either one concealed second or two, depending on the particular 1000 ms boundaries used.

The seconds in this sub-block are not necessarily calendar seconds. At the tail end of a session, periods of time of less than 1000ms shall be incorporated into these counts if they exceed 500ms and shall be disregarded if they are less than 500ms.

4.1. Report Block Structure

Concealed Seconds metrics block

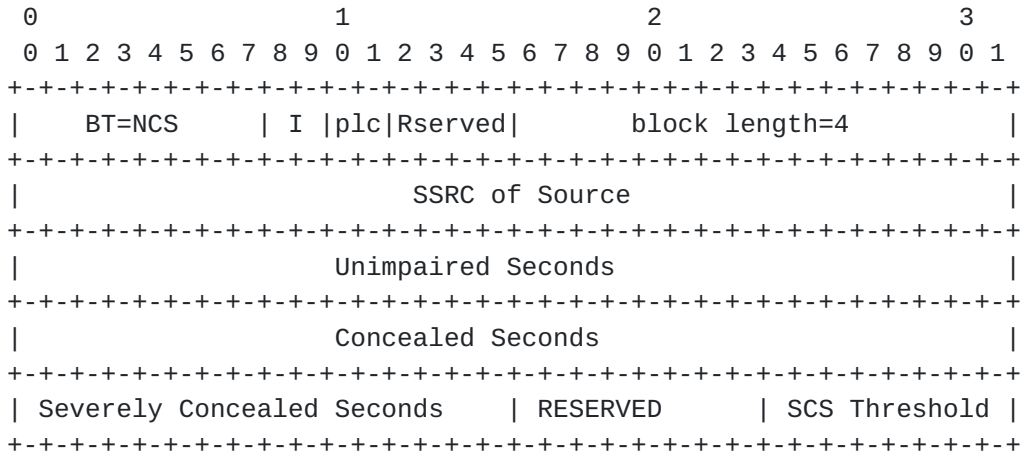

Figure 1: Report Block Structure

4.2. Definition of Fields in Concealed Seconds Metrics Block

Block type (BT): 8 bits

A Concealed Seconds Metrics Report Block is identified by the constant NCS.

[Note to RFC Editor: please replace NCS with the IANA provided RTCP XR block type for this block.]

Interval Metric flag (I): 2 bit

This field is used to indicate whether the Concealment Seconds metrics are Sampled, Interval or Cumulative metrics:

I=10: Interval Duration - the reported value applies to the most recent measurement interval duration between successive metrics reports.

I=11: Cumulative Duration - the reported value applies to the accumulation period characteristic of cumulative measurements.

I=01: Sampled Value - the reported value is a sampled instantaneous value.

Packet Loss Concealment Method (plc): 2 bits

This field is used to identify the packet loss concealment method in use at the receiver, according to the following code:

bits 014-015

0 = silence insertion

1 = simple replay, no attenuation

2 = simple replay, with attenuation

3 = enhanced

Other values reserved

Reserved (resv): 4 bits

These bits are reserved. They MUST be set to zero by senders and ignored by receivers (See [RFC6709] section 4.2).

Block Length: 16 bits

The length of this report block in 32-bit words, minus one. For the Concealment Seconds block, the block length is equal to 4.

SSRC of source: 32 bits

As defined in Section 4.1 of [RFC3611].

Unimpaired Seconds: 32 bits

A count of the number of unimpaired Seconds that have occurred.

An unimpaired Second is defined as a continuous period of 1000ms during which no frame loss or discard due to late arrival has occurred. Every second in a session must be classified as either OK or Concealed.

Normal playout of comfort noise or other silence concealment signal during periods of talker silence, if VAD [VAD] is used, shall be counted as unimpaired seconds.

If the measured value exceeds 0xFFFFFFFD, the value 0xFFFFFFFE MUST be reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFFFFFF MUST be reported.

Concealed Seconds: 32 bits

A count of the number of Concealed Seconds that have occurred.

A Concealed Second is defined as a continuous period of 1000ms during which any frame loss or discard due to late arrival has occurred.

Equivalently, a concealed second is one in which some Loss-type concealment has occurred. Buffer adjustment-type concealment SHALL not cause Concealed Seconds to be incremented, with the following exception. An implementation MAY cause Concealed Seconds to be incremented for 'emergency' buffer adjustments made during talkspurts.

Loss-type concealment is reactive insertion or deletion of samples in the audio playout stream due to effective frame loss at the audio decoder. "Effective frame loss" is the event in which a frame of coded audio is simply not present at the audio decoder

when required. In this case, substitute audio samples are generally formed, at the decoder or elsewhere, to reduce audible impairment.

Buffer Adjustment-type concealment is proactive or controlled insertion or deletion of samples in the audio playout stream due to jitter buffer adaptation, re-sizing or re-centering decisions within the endpoint.

Because this insertion is controlled, rather than occurring randomly in response to losses, it is typically less audible than loss-type concealment. For example, jitter buffer adaptation events may be constrained to occur during periods of talker silence, in which case only silence duration is affected, or sophisticated time-stretching methods for insertion/deletion during favorable periods in active speech may be employed. For these reasons, buffer adjustment-type concealment MAY be exempted from inclusion in calculations of Concealed Seconds and Severely Concealed Seconds.

However, an implementation SHOULD include buffer-type concealment in counts of Concealed Seconds and Severely Concealed Seconds if the event occurs at an 'inopportune' moment, with an emergency or large, immediate adaptation during active speech, or for unsophisticated adaptation during speech without regard for the underlying signal, in which cases the assumption of low-audibility cannot hold. In other words, jitter buffer adaptation events which may be presumed to be audible SHOULD be included in Concealed Seconds and Severely Concealed Seconds counts.

Concealment events which cannot be classified as Buffer Adjustment- type MUST be classified as Loss-type.

For clarification, the count of Concealed Seconds MUST include the count of Severely Concealed Seconds.

If the measured value exceeds 0xFFFFFFFD, the value 0xFFFFFFFE MUST be reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFFFFFF MUST be reported.

Severely Concealed Seconds: 16 bits

A count of the number of Severely Concealed Seconds.

A Severely Concealed Second is defined as a non-overlapping period of 1000 ms during which the cumulative amount of time that has been subject to frame loss or discard due to late arrival, exceeds the SCS Threshold.

If the measured value exceeds 0xFFFD, the value 0xFFFE MUST be reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFF MUST be reported.

Reserved: 8 bits

These bits are reserved. They MUST be set to zero by senders and ignored by receivers (See [RFC6709] section 4.2).

SCS Threshold: 8 bits

The SCS Threshold defines the amount of time corresponding to lost or discarded frames that must occur within a one second period in order for the second to be classified as a Severely Concealed Second. This is expressed in milliseconds and hence can represent a range of 0.1 to 25.5 percent loss or discard.

A default threshold of 50ms (5% effective frame loss per second) is suggested.

5. SDP Signaling

[RFC3611] defines the use of SDP (Session Description Protocol) [RFC4566] for signaling the use of XR blocks. XR blocks MAY be used without prior signaling.

5.1. SDP rtcp-xr-attrib Attribute Extension

This section augments the SDP attribute "rtcp-xr" [RFC3611] by providing two additional values of "xr-format" to signal the use of the report block defined in this document.

```
xr-format =/ xr-conceal-block
xr-format =/ xr-conc-sec-block

xr-conceal-block = "loss-conceal"
xr-conc-sec-block = "conc-sec" ["=" thresh]

thresh      = 1*DIGIT          ; threshold for SCS (ms)
DIGIT          = %x30-39
```

5.2. Offer/Answer Usage

When SDP is used in offer-answer context, the SDP Offer/Answer usage defined in [RFC3611] applies.

6. IANA Considerations

New block types for RTCP XR are subject to IANA registration. For general guidelines on IANA considerations for RTCP XR, refer to [RFC3611].

6.1. New RTCP XR Block Type values

This document assigns two block type values in the IANA "RTCP XR Block Type Registry":

Name: NLC
Long Name: Loss Concealment Block
Value <NLC>
Reference: Section 3.1

Name: NCS
Long Name: Concealment Seconds Block
Value <NCS>
Reference: Section 4.1

[Note to RFC Editor: please replace <NLC> and <NCS> with the RTCP XR block type assigned by IANA for this block.]

6.2. New RTCP XR SDP Parameters

This document also registers two new parameters in the "RTCP XR SDP Parameters Registry":

- "loss-conceal"

- "conc-sec"

6.3. Contact information for registrations

The contact information for the registrations is:

Qin Wu (sunseawq@huawei.com)
101 Software Avenue, Yuhua District
Nanjing, Jiangsu 210012
China

7. Security Considerations

It is believed that this proposed RTCP XR report block introduces no new security considerations beyond those described in [RFC3611]. This block does not provide per-packet statistics so the risk to confidentiality documented in Section 7, paragraph 3 of [RFC3611] does not apply.

8. Contributors

Geoff Hunt wrote the initial draft of this document.

9. Acknowledgements

The authors gratefully acknowledge reviews and feedback provided by Bruce Adams, Philip Arden, Amit Arora, Bob Biskner, Kevin Connor, Claus Dahm, Randy Ethier, Roni Even, Jim Frauenthal, Albert Higashi, Tom Hock, Shane Holthaus, Paul Jones, Rajesh Kumar, Keith Lantz, Mohamed Mostafa, Amy Pendleton, Colin Perkins, Mike Ramalho, Ravi Raviraj, Albrecht Schwarz, Tom Taylor, and Hideaki Yamada.

10. References

10.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", March 1997.

[RFC3550] Schulzrinne, H., "RTP: A Transport Protocol for Real-Time Applications", RFC 3550, July 2003.

[RFC3611] Friedman, T., Caceres, R., and A. Clark, "RTP Control Protocol Extended Reports (RTCP XR)", November 2003.

[RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", July 2006.

10.2. Informative References

[RFC6390] Clark, A. and B. Claise, "Framework for Performance Metric Development", RFC 6390, October 2011.

[RFC6709] Carpenter, B., Aboba, B., and S. Cheshire, "Design Considerations for Protocol Extensions", RFC 6709, September 2012.

[RFC6792] Hunt, G., "Monitoring Architectures for RTP", RFC 6792, November 2012.

[VAD] "http://en.wikipedia.org/wiki/Voice_activity_detection".

Appendix A. Change Log

Note to the RFC-Editor: please remove this section prior to publication as an RFC.

A.1. draft-ietf-xrblock-rtcp-xr-loss-conceal-04

The following are the major changes to previous version :

- Merge Concealment Seconds draft into this draft (i.e.,Loss Concealment draft).

- Updated references.

Authors' Addresses

Alan Clark
Telchemy Incorporated
2905 Premiere Parkway, Suite 280
Duluth, GA 30097
USA

Email: alan.d.clark@telchemy.com

Glen Zorn (editor)
Network Zen
77/440 Soi Phoomjit, Rama IV Road
Phra Khanong, Khlong Toie
Bangkok 10110
Thailand

Phone: +66 (0) 87 502 4274
Email: gwz@net-zen.net

Claire Bi
Shanghai Research Institure of China Telecom Corporation Limited
No.1835,South Pudong Road
Shanghai 200122
China

Email: bijy@sttri.com.cn

Qin Wu (editor)
Huawei
101 Software Avenue, Yuhua District
Nanjing, Jiangsu 210012
China

Email: sunseawq@huawei.com