Internet Engineering Task Force (IETF)
Request for Comments: 7867
Category: Standards Track
ISSN: 2070-1721

R. Huang
Huawei

July 2016

RTP Control Protocol (RTCP) Extended Report (XR) Block for Loss Concealment Metrics for Video Applications

Abstract

This document defines a new RTP Control Protocol (RTCP) Extended Report (XR) block that allows the reporting of loss concealment metrics for video applications of RTP.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc7867.

Table of Contents

1. Introduction

Multimedia applications often suffer from packet losses in IP networks. In order to get a reasonable degree of quality when there is packet loss, it is necessary to have loss concealment mechanisms at the decoder. Video loss concealment is a range of techniques to mask the effects of packet loss in video communications.

In some applications, reporting the information of receivers applying video loss concealment could give monitors or senders useful information on the Quality of Experience (QoE) of the application. One example is no-reference video quality evaluation. Video probes located upstream from the video endpoint or terminal may not see loss occurring between the probe and the endpoint, and also may not be fully aware of the specific loss concealment methods being dynamically applied by the video endpoint. Evaluating error concealment is important in this circumstance to estimate the subjective impact of impairments.

This document defines one new block type for video loss concealment to augment those defined in [RFC3611] and [RFC7294] for use in a range of RTP video applications. The metrics defined in this document belong to the class of transport-related terminal metrics defined in [RFC6792].

1.1. RTCP and RTCP XR Reports

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

1.2. Performance Metrics Framework

The Performance Metrics Framework [RFC6390] provides guidance on the definition and specification of performance metrics. The RTP monitoring framework [RFC6792] provides guidelines for the reporting block format using RTCP XR. The XR block type described in this document is in accordance with the guidelines in [RFC6390] and [RFC6792].

1.3. Applicability

These metrics are applicable to video applications the video component of audio/video applications using RTP and applying packet loss concealment mechanisms that are incorporated into the receiving endpoint to mitigate the impact of network impairments on QoE. For example, in an IPTV system, set-top boxes could use this RTCP XR block to report loss and loss concealment metrics to an IPTV management system to enable the service provider to monitor the quality of the IPTV service being delivered to end users.

2. Terminology

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].

3. Video Loss Concealment Methods

Video loss concealment mechanisms can be classified into 4 types as follows:

a) Frame freeze

The impaired video frame is not displayed; instead, the previously displayed frame is frozen for the duration of the loss event.

b) Interframe extrapolation

If an area of the video frame is damaged by loss, the same area from the previous frame(s) can be used to estimate what the missing pixels would have been. This can work well in a scene

with no motion but can be very noticeable if there is significant movement from one frame to another. Simple decoders can simply reuse the pixels that were in the missing area, while more complex decoders can try to use several frames to do a more complex extrapolation. Another example of a sophisticated form of interframe repair is to estimate the motion of the damaged region based on the motion of surrounding regions, and use that to select what part of the previous frame to use for repair. Some important frames, such as Instantaneous Decoding Refresh (IDR) frames, may not depend on any other frames and may be involved in a scene change. Using the interframe extrapolation method to conceal the loss of these frames may not obtain a satisfactory result.

c) Interpolation

A decoder uses the undamaged pixels in the video frame to estimate what the missing block of pixels should have.

d) Error-resilient encoding

The sender encodes the message in a redundant way so that the receiver can correct errors using the redundant information. There are usually two kinds of error-resilient encoding: One is that the redundant data useful for error resiliency performed at the decoder can be embedded into the compressed image/video bitstream. The other is encoding at the bitstream level, e.g., Forward Error Correction (FEC).

Usually, methods b, c, and d are deployed together to provide comprehensive loss concealment in complex decoders, while method a is relatively independent and may be applied in some simple decoders. Moreover, the frame-freeze method repairs video based on frames, while the other methods repair video based on fine-grained elements, such as macroblocks or bitstreams; this will cause the measurement metrics of frame-freeze and the other methods to be slightly different. Thus, In this document, we differentiate between frame-freeze and the other 3 loss concealment mechanisms.

4. Video Loss Concealment Report Block

This block reports the video loss concealment metrics to complement the audio metrics defined in [RFC7294]. The report block MUST be sent in conjunction with the information from the Measurement Information Block [RFC6776]. Instances of this metric block refer by synchronization source (SSRC) to the separate auxiliary Measurement Information Block [RFC6776]. The Video Loss Concealment Report Block relies on the measurement period in the Measurement Information Block indicating the span of the report. If the measurement period is not

received in the same compound RTCP packet as this metric block, this metric block MUST be discarded at the receiving side. The metrics in this report block are based on measurements that are typically made at the time that a video frame is decoded and rendered for playout.

The Video Loss Concealment Report Block has the following format:

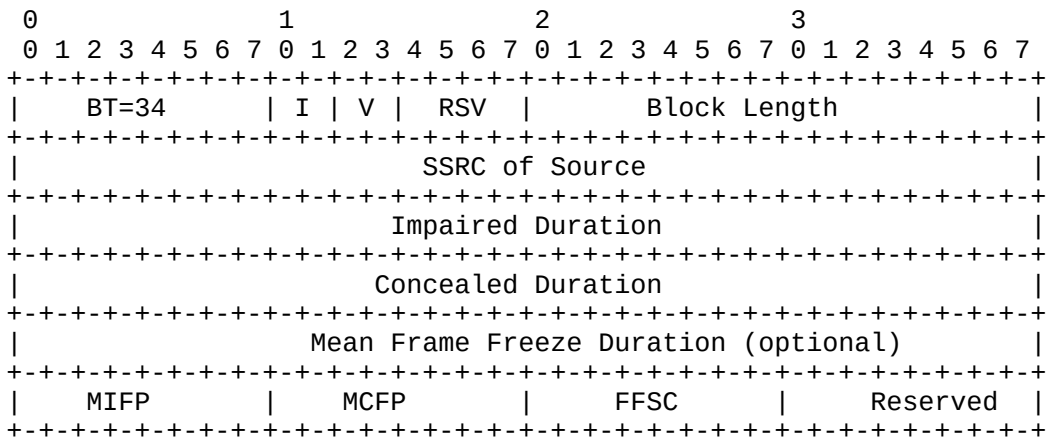

Figure 1: Format for the Video Loss Concealment Report Block

Block Type (BT): 8 bits

A Video Loss Concealment Report Block is identified by the constant 34.

Interval Metric Flag (I): 2 bits

This field indicates whether the reported metrics are interval, cumulative, or sampled metrics [RFC6792]:

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 - this value MUST NOT be used for this block type.

I=00: Reserved.

Video Loss Concealment Method Type (V): 2 bits

This field is used to identify the video loss concealment method type used at the receiver. The value is defined as follows:

V=10: Frame-freeze
V=11: Other Loss Concealment Method
V=01 and V=00: Reserved

If frame-freeze and another loss concealment method are used together for the media stream, two report blocks (one with V=10 for frame freeze and one with V=11 for the other loss concealment method) SHOULD be compounded together to report complete concealment information.

RSV: 4 bits

These bits are reserved for future use. They MUST be set to zero by senders and ignored by receivers (see Section 4.2 of [RFC6709]).

Block Length: 16 bits

This field is in accordance with the definition in [RFC3611]. In this report block, it MUST be set to 5 when V=10 and set to 4 when V=11. The block MUST be discarded if the block length is set to a different value.

SSRC of Source: 32 bits

As defined in Section 4.1 of [RFC3611].

Impaired Duration: 32 bits

The total duration, expressed in units of RTP timestamp from the sending side of the reporting block, of video impaired by transmission loss before applying any loss concealment methods.

Two values are reserved: A value of 0xFFFFFFFE indicates out of range (that is, a measured value exceeding 0xFFFFFFFD), and a value of 0xFFFFFFFF indicates that the measurement is unavailable.

Concealed Duration: 32 bits

The total duration, expressed in units of RTP timestamp from the sending side of the reporting block, of concealed damaged video pictures on which the loss concealment method corresponding to the Video Loss Concealment Method Type is applied.

Two values are reserved: A value of 0xFFFFFFFE indicates out of range (that is, a measured value exceeding 0xFFFFFFFD), and a value of 0xFFFFFFFF indicates that the measurement is unavailable.

Mean Frame-Freeze Duration: 32 bits

Mean Frame-Freeze Duration is the mean duration, expressed in units of RTP timestamp from the sending side of the reporting block, of the frame-freeze events. The value of Mean Frame-Freeze Duration is calculated by summing the total duration of all frame freeze events and dividing by the number of events. This metric is optional. It only exists when Video Loss Concealment Method Type=10.

Mean Impaired Frame Proportion (MIFP): 8 bits

Mean Impaired Frame Proportion is the mean proportion of each video frame impaired by loss before applying any loss concealment method during the interval, expressed as a fixed-point number with the binary point at the left edge of the field. It is calculated by summing the impaired proportion of each video frame and dividing by the number of frames during this period. The impaired proportion of each video frame is obtained by dividing the number of missing macroblocks from this video frame by the total macroblock number of the video frame, which is equivalent to multiplying the result of the division by 256, limiting the maximum value to 255 (to avoid overflow), and taking the integer part.

If a video frame is totally lost, a value of 0xFF SHOULD be used for the frame when calculating the MIFP.

Mean Concealed Frame Proportion (MCFP): 8 bits

Mean Concealed Frame Proportion is the mean proportion of each video frame to which loss concealment (depicted as "V" in the definition of "Video Loss Concealment Method Type") was applied during the interval, expressed as a fixed-point number with the binary point at the left edge of the field. It is calculated by summing the concealed proportion of each video frame and dividing by the number of frames during this period. The concealed proportion of each video frame is obtained by dividing the number of concealed macroblocks from this video frame by the total macroblock number of the video frame, which is equivalent to multiplying the result of the division by 256, limiting the maximum value to 255 (to avoid overflow), and taking the integer part.

When calculating the MCFP, a value of 0xFF SHOULD be used for a lost frame that is totally concealed, and a value of 0 SHOULD be used for the frame if there are no concealed macroblocks in it. For Video Loss Concealment Method Type=10, each frame covered in the period of frame freeze is considered to be totally concealed; this means a value of 0xFF MUST be assigned.

Fraction of Frames Subject to Concealment (FFSC): 8 bits

Fraction of Frames Subject to Concealment is calculated by dividing the number of frames to which loss concealment (using Video Loss Concealment Method Type) was applied by the total number of frames and expressing this value as a fixed-point number with the binary point at the left edge of the field. It is equivalent to multiplying the result of the division by 256, limiting the maximum value to 255 (to avoid overflow), and taking the integer part.

A value of 0 indicates that there were no concealed frames, and a value of 0xFF indicates that the frames in the entire measurement interval are all concealed.

Reserved: 8 bits

These bits are reserved for future use. They MUST be set to zero by senders and ignored by receivers (see Section 4.2 of [RFC6709]).

5. SDP Signaling

[RFC3611] defines the use of the Session Description Protocol (SDP) for signaling the use of RTCP XR blocks.

5.1. SDP rtcp-xr-attrib Attribute Extension

This session augments the SDP attribute "rtcp-xr" defined in Section 5.1 of [RFC3611] by providing an additional value of "xr-format" to signal the use of the report block defined in this document. The ABNF [RFC5234] syntax is as follows.

```
xr-format =/ xr-vlc-block

xr-vlc-block = "vlc"
```

5.2. Offer/Answer Usage

When SDP is used in an offer/answer context, the SDP Offer/Answer usage defined in Section 5.2 of [RFC3611] for the unilateral "rtcp-xr" attribute parameters applies. For detailed usage of Offer/Answer for unilateral parameters, refer to Section 5.2 of [RFC3611].

6. Security Considerations

It is believed that this RTCP XR 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 paragraph 3 of Section 7 of [RFC3611] does not apply.

An attacker is likely to put incorrect information in the Video Loss Concealment reports; this will affect the estimation of the performance of video loss concealment mechanisms and the QoE of users. Implementers SHOULD consider the guidance in [RFC7202] for using appropriate security mechanisms, i.e., where security is a concern, the implementation SHOULD apply encryption and authentication to the report block. For example, this can be achieved by using the AVPF profile together with the Secure RTP profile as defined in [RFC3711]; an appropriate combination of the two profiles (an "SAVPF") is specified in [RFC5124]. However, other mechanisms also exist (documented in [RFC7201]) and might be more suitable.

7. IANA Considerations

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

7.1. New RTCP XR Block Type Value

This document assigns the block type value 34 to Video Loss Concealment Metric Report Block in the IANA "RTP Control Protocol Extended Reports (RTCP XR) Block Type Registry".

7.2. New RTCP XR SDP Parameter

This document also registers a new parameter "video-loss-concealment" in the "RTP Control Protocol Extended Reports (RTCP XR) Session Description Protocol (SDP) Parameters Registry".

7.3. Contact Information for Registrations

The contact information for the registration is:

RAI Area Directors <rai-ads@ietf.org>

8. References

8.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550, DOI 10.17487/RFC3550, July 2003, <http://www.rfc-editor.org/info/rfc3550>.

[RFC3611] Friedman, T., Ed., Caceres, R., Ed., and A. Clark, Ed., "RTP Control Protocol Extended Reports (RTCP XR)", RFC 3611, DOI 10.17487/RFC3611, November 2003, <http://www.rfc-editor.org/info/rfc3611>.

[RFC3711] Baugher, M., McGrew, D., Naslund, M., Carrara, E., and K. Norrman, "The Secure Real-time Transport Protocol (SRTP)", RFC 3711, DOI 10.17487/RFC3711, March 2004, <http://www.rfc-editor.org/info/rfc3711>.

[RFC5124] Ott, J. and E. Carrara, "Extended Secure RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/SAVPF)", RFC 5124, DOI 10.17487/RFC5124, February 2008, <http://www.rfc-editor.org/info/rfc5124>.

[RFC5234] Crocker, D., Ed., and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, RFC 5234, DOI 10.17487/RFC5234, January 2008, <http://www.rfc-editor.org/info/rfc5234>.

[RFC6776] Clark, A. and Q. Wu, "Measurement Identity and Information Reporting Using a Source Description (SDES) Item and an RTCP Extended Report (XR) Block", RFC 6776, DOI 10.17487/RFC6776, October 2012, <http://www.rfc-editor.org/info/rfc6776>.

[RFC7294] Clark, A., Zorn, G., Bi, C., and Q. Wu, "RTP Control Protocol (RTCP) Extended Report (XR) Blocks for Concealment Metrics Reporting on Audio Applications", RFC 7294, DOI 10.17487/RFC7294, July 2014, <http://www.rfc-editor.org/info/rfc7294>.

8.2. Informative References

[RFC6390] Clark, A. and B. Claise, "Guidelines for Considering New Performance Metric Development", BCP 170, RFC 6390, DOI 10.17487/RFC6390, October 2011, <http://www.rfc-editor.org/info/rfc6390>.

[RFC6709] Carpenter, B., Aboba, B., Ed., and S. Cheshire, "Design Considerations for Protocol Extensions", RFC 6709, DOI 10.17487/RFC6709, September 2012, <http://www.rfc-editor.org/info/rfc6709>.

[RFC6792] Wu, Q., Ed., Hunt, G., and P. Arden, "Guidelines for Use of the RTP Monitoring Framework", RFC 6792, DOI 10.17487/RFC6792, November 2012, <http://www.rfc-editor.org/info/rfc6792>.

[RFC7201] Westerlund, M. and C. Perkins, "Options for Securing RTP Sessions", RFC 7201, DOI 10.17487/RFC7201, April 2014, <http://www.rfc-editor.org/info/rfc7201>.

[RFC7202] Perkins, C. and M. Westerlund, "Securing the RTP Framework: Why RTP Does Not Mandate a Single Media Security Solution", RFC 7202, DOI 10.17487/RFC7202, April 2014, <http://www.rfc-editor.org/info/rfc7202>.

Appendix A. Metrics Represented Using the Template from RFC 6390

a. Video Impaired Duration Metric

* Metric Name: Video Impaired Duration Metric

* Metric Description: The total duration of the video impaired by transmission loss before applying any loss concealment methods.

* Method of Measurement or Calculation: The metric is based on measurements that are typically made at the time that a video frame is decoded and rendered for playout.

* Units of Measurement: This metric is expressed in units of RTP timestamp.

* Measurement Point(s) with Potential Measurement Domain: It is measured at the receiving end of the RTP stream.

* Measurement Timing: See paragraph 1 of Section 4.

* Use and Applications: The metric is applicable to video applications of RTP and the video component of audio/video applications in which packet loss concealment mechanisms are applied to the receiving endpoint to mitigate the impact of network impairments on QoE.

b. Video Concealed Duration Metric

* Metric Name: Video Concealed Duration Metric

* Metric Description: The total duration of concealed damaged video pictures on which loss concealment method corresponding to Video Loss Concealment Method Type is applied.

* Method of Measurement or Calculation: The metric is based on measurements that are typically made at the time that a video frame is decoded and rendered for playout.

* Units of Measurement: This metric is expressed in units of RTP timestamp.

* Measurement Point(s) with Potential Measurement Domain: It is measured at the receiving end of the RTP stream.

* Measurement Timing: See paragraph 1 of Section 4.

* Use and Applications: These metrics are applicable to video applications of RTP and the video component of audio/video applications in which packet loss concealment mechanisms are incorporated into the receiving endpoint to mitigate the impact of network impairments on QoE.

c. Mean Video Frame-Freeze Duration Metric

* Metric Name: Mean Video Frame-Freeze Duration Metric

* Metric Description: The mean duration of the frame-freeze events.

* Method of Measurement or Calculation: The metric is based on measurements that are typically made at the time that a video frame is decoded and rendered for playout. The metric is calculated by summing the total duration of all frame-freeze events and dividing by the number of events.

* Units of Measurement: This metric is expressed in units of RTP timestamp.

* Measurement Point(s) with Potential Measurement Domain: It is measured at the receiving end of the RTP stream.

* Measurement Timing: See paragraph 1 of Section 4.

* Use and Applications: These metrics are applicable to video applications of RTP and the video component of audio/video applications in which packet loss concealment mechanisms are incorporated into the receiving endpoint to mitigate the impact of network impairments on QoE.

d. Mean Impaired Video Frame Proportion Metric

* Metric Name: Mean Impaired Video Frame Proportion Metric

* Metric Description: Mean proportion of each video frame impaired by loss before applying any loss concealment method during the interval.

* Method of Measurement or Calculation: The metric is based on measurements that are typically made at the time that a video frame is decoded and rendered for playout. It is calculated by summing the impaired proportion of each video frame and dividing by the number of frames during this period. The impaired proportion of each video frame is obtained by dividing the number of missing macroblocks from this video frame by the

total macroblock number of the video frame, which is equivalent to multiplying the result of the division by 256, limiting the maximum value to 255 (to avoid overflow), and taking the integer part.

* Units of Measurement: This metric is expressed as a fixed-point number with the binary point at the left edge of the field.

* Measurement Point(s) with Potential Measurement Domain: It is measured at the receiving end of the RTP stream.

* Measurement Timing: See paragraph 1 of Section 4.

* Use and Applications: These metrics are applicable to video applications of RTP and the video component of audio/video applications in which packet loss concealment mechanisms are incorporated into the receiving endpoint to mitigate the impact of network impairments on QoE.

e. Mean Concealed Video Frame Proportion Metric

* Metric Name: Mean Concealed Video Frame Proportion Metric

* Metric Description: Mean proportion of each video frame to which loss concealment (using Video Loss Concealment Method Type) was applied during the interval.

* Method of Measurement or Calculation: The metric is based on measurements that are typically made at the time that a video frame is decoded and rendered for playout. It is calculated by summing the concealed proportion of each video frame and dividing by the number of frames during this period. The concealed proportion of each video frame is obtained by dividing the number of concealed macroblocks from this video frame by the total macroblock number of the video frame, which is equivalent to multiplying the result of the division by 256, limiting the maximum value to 255 (to avoid overflow), and taking the integer part.

* Units of Measurement: This metric is expressed as a fixed-point number with the binary point at the left edge of the field.

* Measurement Point(s) with Potential Measurement Domain: It is measured at the receiving end of the RTP stream.

* Measurement Timing: See paragraph 1 of Section 4.

* Use and Applications: These metrics are applicable to video applications of RTP and the video component of audio/video applications in which packet loss concealment mechanisms are incorporated into the receiving endpoint to mitigate the impact of network impairments on QoE.

f. Fraction of Video Frames Subject to Concealment Metric

* Metric Name: Fraction of Video Frames Subject to Concealment Metric

* Metric Description: Proportion of concealed video frames to which loss concealment (using the Video Loss Concealment Method Type) was applied compared to the total number of frames during the interval.

* Method of Measurement or Calculation: The metric is based on measurements that are typically made at the time that a video frame is decoded and rendered for playout. This metric is calculated by dividing the number of frames to which loss concealment (using Video Loss Concealment Method Type) was applied by the total number of frames. It is equivalent to multiplying the result of the division by 256, limiting the maximum value to 255 (to avoid overflow), and taking the integer part.

* Units of Measurement: This metric is expressed as a fixed-point number with the binary point at the left edge of the field.

* Measurement Point(s) with Potential Measurement Domain: It is measured at the receiving end of the RTP stream.

* Measurement Timing: See paragraph 1 of Section 4.

* Use and Applications: These metrics are applicable to video applications of RTP and the video component of audio/video applications in which packet loss concealment mechanisms are incorporated into the receiving endpoint to mitigate the impact of network impairments on QoE.

Acknowledgements

The author would like to thank Colin Perkins and Roni Even for their valuable comments.

Authors' Addresses

Rachel Huang
Huawei
101 Software Avenue, Yuhua District
Nanjing 210012
China

Email: rachel.huang@huawei.com