Workgroup: avtcore
Internet-Draft:
draft-gudumasu-avtcore-decoder-energy-
reduction-00
Published: 12 March 2023
Intended Status: Standards Track
Expires: 13 September 2023
Authors: S. Gudumasu F. Aumont E. Francois
InterDigital InterDigital InterDigital
C. Herglotz
Friedrich-Alexander-Universität Erlangen-Nürnberg

RTP Control Protocol (RTCP) Messages for Decoder Energy Reduction

Abstract

This document describes an RTCP feedback message format for the second type of green metadata defined by the ISO/IEC International Standard 23001-11, known as Energy Efficient Media Consumption (Green metadata), developed by the ISO/IEC JTC 1/SC 29/WG 3 MPEG System. The RTCP feedback messages specified in this specification is compatible and complimentary with the other draft on green metadata and enables receivers to provide feedback to the senders for decoder power reduction and thus allows feedback-based energy efficient mechanisms to be implemented. The feedback message has broad applicability in real-time video communication services.

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 https://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 13 September 2023.

Copyright Notice

Table of Contents

1. Introduction

ISO/IEC 23001-11 specification, Energy Efficient Media Consumption (Green metadata) [GreenMetadata], specifies metadata that facilitates reduction of energy usage during media consumption. Two main types of metadata are defined in the specification. The first type consists of metadata generated by a video encoder which provides information about the decoding complexity of the delivered bitstream and about the quality of the decoded content. This first type of metadata is conveyed via the supplemental enhancement information (SEI) message mechanism specified in the video coding standard ITU-T Recommendation H.264 and ISO/IEC 14496-10 [AVC], H. 265 and ISO/IEC 23008-5 [HEVC], H.266 and ISO/IEC 23090-3 [VVC]. The document [I-D.draft-ietf-avtcore-rtcp-green-metadata] focuses on

this first type of metadata. It describes the spatial and temporal resolution request and notification feedback messages .

The second type consists of metadata generated by a decoder as feedback conveyed to the encoder to adapt the decoder energy consumption. This document focuses on this second type of metadata which is conveyed as extension of RTCP feedback messages [RFC4585]. The feedback includes decoder operations reduction Request, and coding tools configuration request.

This document describes a new RTCP feedback message (a decoder operation reduction request) that enables receivers to provide feedback to the senders and thus allows the sender for short-term adaptation and feedback-based energy efficient mechanisms to be implemented.

Both types of metadata can be used concurrently to further reduce energy consumption. Therefore, the messages described in this document can be used concurrently with messages described in [I-D.draft-ietf-avtcore-rtcp-green-metadata].

2. Conventions

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. Abbreviations

AVPF: The extended RTP profile for RTCP-based feedback

FCI: Feedback Control Information [RFC4585]

FMT: Feedback Message Type [RFC4585]

PSFB: Payload-specific FB message [RFC4585]

DORR: Decoder Operation Reduction Request

DORN: Decoder Operation Reduction Notification

4. Format of RTCP feedback messages

This document extends the RTCP feedback messages defined in the RTP/AVPF [RFC4585] and [RFC5104] and the [I-D.draft-ietf-avtcore-rtcp-green-metadata] by defining new Decoder Operation Reduction feedback messages. The RTCP feedback messages can be used by the receiver to inform the sender of the desirable decoding operation reduction of the bitstream delivered or the coding tools that shall be disabled within the bitstream delivered,

and by the sender to indicate the decoding operation reduction and the disabled coding tools it will use henceforth.

RTCP Green Metadata feedback message follows a similar message format as RTCP Temporal-Spatial Trade-off Request and Notification [RFC5104]. The message may be sent in a regular full compound RTCP packet or in an early RTCP packet, as per the RTP/AVPF rules.

This document specifies two additional payload-specific feedback messages: Decoder Operation Reduction Request (DORR) and Decoder Operation Reduction Notification (DORN).

4.1. Decoder Operation Reduction Request (DORR)

The DORR feedback message is identified by RTCP packet type value PT=PSFB and FMT=11.

The FCI field MUST contain one or more DORR FCI entries.

4.1.1. Message format

The content of the FCI entry for the Decoder Operation Reduction Request is depicted in Figure 1.

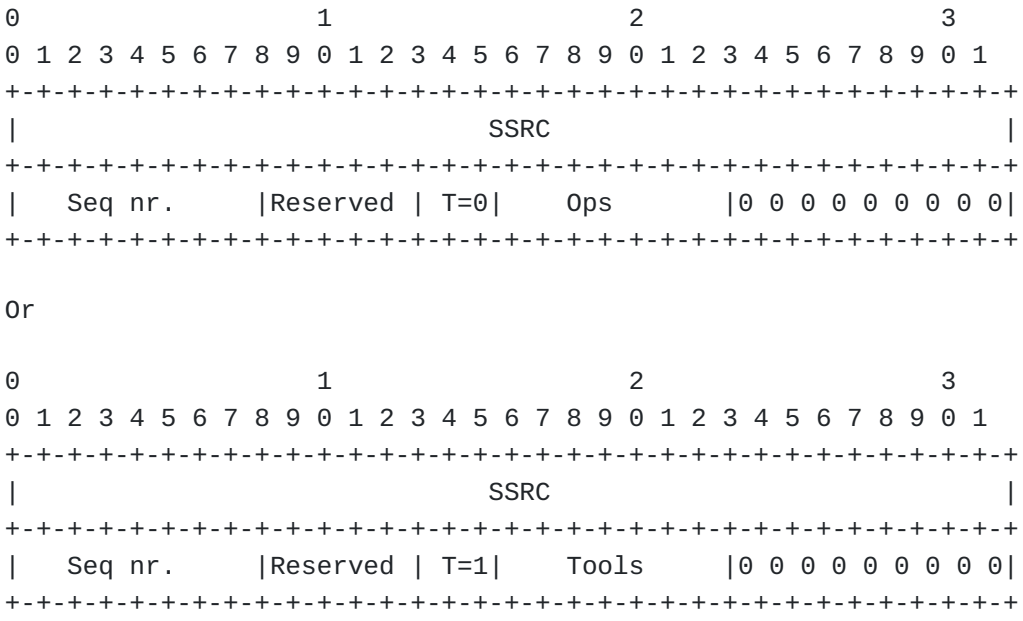

Figure 1: Syntax of an FCI Entries in the DORR Message

SSRC (32 bits): The Synchronization Source (SSRC) of the media sender that is requested to apply the Decoder Operation.

Seq nr. (8 bits): Request sequence number. The sequence number space is unique for pairing of the SSRC of request source and the

SSRC of the request target. The sequence number SHALL be increased by 1 modulo 256 for each new command. A repetition SHALL NOT increase the sequence number. The initial value is arbitrary. Reserved (4 bits): All bits SHALL be set to 0 by the sender and SHALL be ignored on reception.

T(2 bits): Decoding power reduction type. This field specifies the type of the decoder power reduction method as defined in clause 6.3 of [GreenMetadata]. Two modes are defined for expressing the required decoding reduction at the decoding side. Type T=0 indicates the relative value of decoding operations reduced compared to the previous decoding operations. Type T=1 indicates the required decoding operations are reduced based on the enabling or disabling selected coding tools. Others type values SHALL be ignored on reception.

Ops (6 bits): Type field value is T=0, this field specifies the variation of decoding operations relative to the decoding operations since the last DORN feedback message received by the transmitter, or since the start of the video session as defined in clause 6.3 of [GreenMetadata].

Tools (6 bits): Coding tools enabled or disabled. When the Type filed value is T=1, this 6-bit field represents the enabling or disabling of selected coding tools. When the receiver requests an encoder to disable loop filtering coding tools to reduce the decoding operations, 1st LSB bit is set to 1, otherwise it is set to 0. When the receiver requests an encoder to disable bi-directional prediction coding tools to reduce the decoding operations, 2nd bit is set to 1, otherwise it is set to 0. When the receiver requests an encoder to disable usage of intra prediction in a B frame coding tool to reduce the decoding operations, 3rd bit is set to 1, otherwise it is set to 0. When receiver requests an encoder to disable usage of fractional-pel interpolation filter coding tools to reduce the decoding operations, 4th bit is set to 1, otherwise it is set to 0. The 5th and 6th bits represent optional coding tools an encoder can disable to reduce decoder-side operations.

4.1.2. Semantics

A decoder can suggest a decoder operation reduction by sending a DORR message to an encoder. The decoder indicates the requested variation of local decoding operations since the start of the session or since the last DORR message. If the encoder is capable of adjusting, it SHOULD take into account the received DORR message for future coding of pictures.

The reaction to the reception of more than one DORR message by a media sender from different media receivers is left open to the implementation. The selected Ops SHALL be communicated to the media receivers by means of the DORN message (see Section 4.4).

Within the common packet header for feedback messages (as defined in section 6.1 of [RFC4585]), the "SSRC of packet sender" field indicates the source of the request, and the "SSRC of media source" is not used and SHALL be set to 0. The SSRCs of the media senders to which the DORR applies are in the corresponding FCI entries.

A DORR message MAY contain requests to multiple media senders, using one FCI entry per target media sender.

4.1.3. Timing Rules

The timing follows the rules outlined in section 3 of [RFC4585]. This request message is not time critical and SHOULD be sent using regular RTCP timing. Only if it is known that the user interface requires quick feedback, the message MAY be sent with early or immediate feedback timing.

4.1.4. Handling of Message in Mixers and Translators

A mixer or media translator that encodes content sent to the session participant issuing the DORR SHALL consider the request to determine if it can fulfill it by changing its own encoding parameters. A media translator unable to fulfill the request MAY forward the request unaltered towards the media sender. A mixer encoding for multiple session participants will need to consider the joint needs of these participants before generating a DORR on its own behalf towards the media sender.

4.2. Decoder Operation Reduction Notification (DORN)

The DORN message is identified by RTCP packet type value PT=PSFB and FMT=12.

The FCI field SHALL contain one or more DORN FCI entries.

4.2.1. Message format

The content of the FCI entry for the Decoder Operation Reduction notification is depicted in Figure 2.

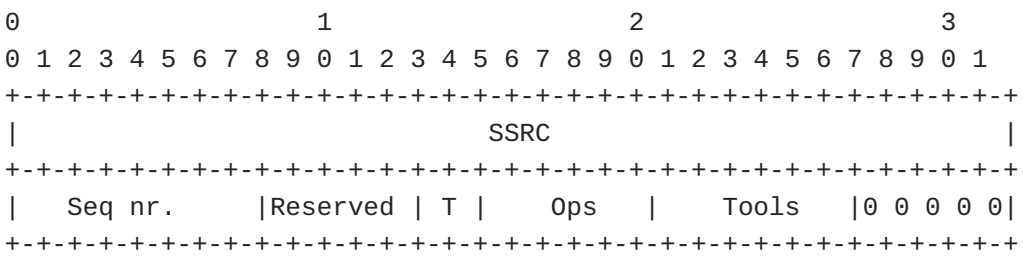

Figure 2: Syntax of an FCI Entry in the DORN Message

SSRC (32 bits): The Synchronization Source (SSRC) of the source of the DORR that resulted in this notification.

Seq nr. (8 bits): The sequence number value from the DORR that is being acknowledged.

T(2 bits): This field indicates the presence of the fields Ops and Tools in the DORN message. Possible values are 01 (Ops only), 10 (Tools only) or 11 (both fields).

Ops (6 bits): Expected operation reduction variation to decode the bitstream delivered by the media sender.

Tools (6 bits): Coding tools the media sender is using henceforth.

It is to note that the returned value (Ops, Tools) may differ from the requested one, for example, in cases where a media encoder cannot change its coding configuration , or when pre-recorded content is used.

4.2.2. Semantics

This feedback message is used to acknowledge the reception of a DORR. For each DORR received targeted at the session participant, a DORN FCI entry SHALL be sent in a DORN feedback message. A single DORN message MAY acknowledge multiple requests using multiple FCI entries. The Ops and Tools value included SHALL be the same in all FCI entries of the DORN message. Including an FCI for each requestor allows each requesting entity to determine that the media sender received the request. The notification SHALL also be sent in response to DORR repetitions received. If the request receiver has received DORR with several different sequence numbers from a single requestor, it SHALL only respond to the request with the highest (modulo 256) sequence number. Note that the highest sequence number may be a smaller integer value due to the wrapping of the field. Appendix A.1 of [RFC3550] has an algorithm for keeping track of the highest received sequence number for RTP packets; it could be adapted for this usage.

The DORN SHALL include the Ops and Tools that will be used as a result of the request. This is not necessarily the same Ops and Tools as requested, as the media sender may need to aggregate requests from several requesting session participants. It may also have some other policies or rules that limit the selection.

Within the common packet header for feedback messages (as defined in section 6.1 of [RFC4585]), the "SSRC of packet sender" field indicates the source of the Notification, and the "SSRC of media source" is not used and SHALL be set to 0. The SSRCs of the requesting entities to which the Notification applies are in the corresponding FCI entries.

4.2.3. Timing Rules

The timing follows the rules outlined in section 3 of [RFC4585]. This acknowledgement message is not extremely time critical and SHOULD be sent using regular RTCP timing.

4.2.4. Handling of DORN in Mixers and Translators

A mixer or translator that acts upon a DORN SHALL also send the corresponding DORN. In cases where it needs to forward a DORR itself, the notification message MAY need to be delayed until the DORR has been responded to.

5. Security Considerations

The defined messages have certain properties that have security implications. These must be addressed and taken into account by users of this protocol.

Spoofed or maliciously created feedback messages of the type defined in this specification can have the following implications:

* severely reduced Ops value due to false DORR messages that sets the Number of operation to a very low value;

* severely Tools value due to false DORR messages that sets the Enabled tools to a state in which the video can't be decoded;

* severely reduced picture resolution due to false DORR messages that sets the picture width and height to a very low value;

* severely reduced frame rate due to false DORR messages that sets the frame rate to a very low value.

To prevent these attacks, there is a need to apply authentication and integrity protection of the feedback messages. This can be accomplished against threats external to the current RTP session

using the RTP profile that combines Secure RTP [SRTP] and AVPF into SAVPF [SAVPF]. In the mixer cases, separate security contexts and filtering can be applied between the mixer and the participants, thus protecting other users on the mixer from a misbehaving participant.

6. SDP Definitions

The capability of handling messages defined in this specification MAY be exchanged at a higher layer such as SDP. This specification follows all the rules defined in AVPF [RFC4585] and CCM [RFC5104] for an "rtcp-fb" attribute relating to the payload type in a session description.

6.1. Extension of the rtcp-fb Attribute

This specification defines a new parameter "DORR" to the "ccm" feedback value defined in CCM [RFC5104] to indicate support of the Decoder Operation Reduction Request/Notification (DORR/DORN). All the rules described in [RFC4585] for rtcp-fb attribute relating to payload type and to multiple rtcp-fb attributes in a session description also apply to the new feedback messages defined in this specification.

rtcp-fb-ccm-param =/ SP "dorr" ; Decoder Operation Reduction Request

6.2. Example

The following SDP describes a point-to-point video call with VVC RTP currently under definition in document RTP Payload Format for Versatile Video Coding(VVC), draft-ietf-avtcore-rtp-vvc-02, with the originator of the call declaring its capability to support the FIR and DORR/DORN codec control messages. The SDP is carried in a high-level signaling protocol like SIP.

Offer:

```
v=0;
o=alice xxxxx
s=Offer/Answer
m=video 49170 RTP/AVP 98
a=rtpmap:98 H266/90000
a=fmtp:98 profile-id=1;
sprop-vps=<"video parameter sets data">;
sprop-sps="<"sequence parameter set data">;
sprop-pps=<"picture parameter set data">;
a=rtcp-fb:98 ccm fir
a=rtcp-fb:98 ccm dorr
```

Answer:

```
v=0;
o=alice xxxxx
s=Offer/Answer
c=xxxx
m=video 49170 RTP/AVP 98
a=rtpmap:98 H266/90000
a=rtcp-fb:98 ccm dorr
```

In the above example, when the sender receives a DORR message from the remote party it is capable of adjusting the trade-off as indicated in the RTCP DORN feedback message.

7. IANA Considerations

Placeholder

8. Informative References

[AVC] ISO/IEC, "Advanced video coding, ITU-T Recommendation H. 264", ISO/IEC 14496-10, 2021, <https://www.itu.int/rec/T-REC-H.264>.

[GreenMetadata] ISO/IEC, "ISO/IEC DIS 23001-11, Information technology - MPEG Systems Technologies - Part 11: Energy-Efficient Media Consumption (Green Metadata)", ISO/IEC 23001-11, 2022, <https://www.iso.org/standard/83674.html>.

[HEVC] ISO/IEC, "High efficiency video coding, ITU-T Recommendation H.265", ISO/IEC 23008-5, 2021, <https://www.itu.int/rec/T-REC-H.265>.

[I-D.draft-ietf-avtcore-rtcp-green-metadata] He, Y., Herglotz, C., and E. Francois, "RTP Control Protocol (RTCP) Messages for Green Metadata", Work in Progress, Internet-Draft, draft-ietf-avtcore-rtcp-green-metadata-00, 19 January 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-avtcore-rtcp-green-metadata-00>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/rfc/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, <https://www.rfc-editor.org/rfc/rfc3550>.

[RFC4585] Ott, J., Wenger, S., Sato, N., Burmeister, C., and J. Rey, "Extended RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/AVPF)", RFC 4585, DOI 10.17487/RFC4585, July 2006, <https://www.rfc-editor.org/rfc/rfc4585>.

[RFC5104] Wenger, S., Chandra, U., Westerlund, M., and B. Burman, "Codec Control Messages in the RTP Audio-Visual Profile with Feedback (AVPF)", RFC 5104, DOI 10.17487/RFC5104, February 2008, <https://www.rfc-editor.org/rfc/rfc5104>.

[SAVPF] IETF, "Extended Secure RTP Profile for RTCP-based Feedback (RTP/SAVPF)", 2008, <https://datatracker.ietf.org/doc/pdf/rfc5124>.

[SRTP] IETF, "The Secure Real-time Transport Protocol(SRTP)", 2004, <https://datatracker.ietf.org/doc/pdf/rfc3711>.

[VVC] ISO/IEC, "Versatile Video Coding, ITU-T Recommendation H. 266", ISO/IEC 23090-3, 2022, <http://www.itu.int/rec/T-REC-H.266>.

Authors' Addresses

Srinivas Gudumasu
InterDigital
Canada

Email: srinivas.gudumasu@interdigital.com

Franck Aumont
InterDigital
France

Email: franck.aumont@interdigital.com

Edouard Francois
InterDigital
France

Email: edouard.francois@interdigital.com

Christian Herglotz
Friedrich-Alexander-Universität Erlangen-Nürnberg
Germany

Email: christian.herglotz@fau.de