Network Working Group I. Johansson
Internet-Draft M. Westerlund
Intended status: Standards Track Ericsson AB
Expires: August 11, 2008 Feb 8, 2008

Support for non-compound RTCP, opportunities and consequences
draft-ietf-avt-rtcp-non-compound-02

Status of this Memo

Copyright Notice

Abstract

This memo discusses benefits and issues that arise when allowing RTCP packets to be transmitted as non-compound packets, i.e not follow the rules of RFC 3550. Based on that analysis this memo proposes changes to the rules to allow feedback messages to be sent as non-compound RTCP packets when using the RTP AVPF profile (RFC 4585) under certain conditions.

Requirements 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

Table of Contents

1. Introduction

In RTP [RFC3550] it is currently mandatory to always use RTCP compound packets containing at least Sender Reports or Receiver reports, and a SDES packet containing at least the CNAME item. There are good reasons for this as discussed below (see Section 2). However this do result in that the minimal RTCP packets are quite large. The RTP profile AVPF [RFC4585] specifies new RTCP packet types for feedback messages. Some of these feedback messages would benefit from being transmitted with minimal delay and AVPF do provide some mechanism to enable this. However for environments with low-bitrate links this still consumes quite large amount of resources and introduce extra delay in the time it takes to completely send the compound packet in the network. There are also other benefits as discussed in Section 3.

The use of non-compound packets is not without issues. This is discussed in Section 4. These issues needs to be considered and are part of the motivation for this document.

In addition this document proposes how AVPF could be updated to allow the transmission of non-compound packets in a way that would not substantially affect the mechanisms that compound packets provide. The connection to AVPF is motivated by the fact that non-compound RTCP is mainly intended for event driven feedback purposes and that the AVPF early and immediate modes make this possible.

2. RTCP Compound Packets

Section 6.1 in RFC3550 [RFC3550] specifies that an RTCP packet must be sent in a compound packet consisting of at least two individual packets, first an Sender Report (SR) or Receiver Report (RR), followed by additional packets including a mandatory SDES packet containing a CNAME Item for the transmitting source identifier (SSRC). Lets examine what these RTCP packet types are used for.

1. The sender and receiver reports (see Section 6.4 of RFC 3550 [RFC3550]) provides the RTP session participant with the Sender Source Identifier (SSRC) of all RTCP senders. Having all participants send these packets periodically allows everyone to determine the current number of participants. This information is used in the transmission scheduling algorithm. Thus this is particularly important for new participants so that they quickly can establish a good estimate of the group size. Failure to do this would result in RTCP senders consuming to much bandwidth.

2. The sender and receiver reports contain some basic statistics usable for monitoring of the transport and thus enable adaptation. These reports become more useful if sent regularly as the receiver of a report can perform analysis to find trends between the individual reports. When used for media transmission adaptation the information become more useful the more frequently it is received, at least until one report per round-trip time (RTT) is achieved. Therefore there are most cases no reason to not include the sender or receiver report in all RTCP packets.

3. The CNAME SDES item (See Section 6.5.1 of RFC 3550 [RFC3550]) exists to allow receivers to determine which media flows that should be synchronized with each other between different RTP sessions carrying different media types. Thus it is important to quickly receive this for each media sender in the session when joining an RTP session.

4. Sender Reports (SR) is used in combination with the above SDES CNAME mechanism to synchronize multiple RTP streams, such as audio and video. After having determined which media streams should be synchronized using the CNAME field, the receiver uses the Sender Report's NTP and RTP timestamp fields to establish synchronization.

Reviewing the above it is obvious that both SR/RR and the CNAME are very important for new session participants to be able to utilize any received media and to avoid flooding the network with RTCP reports. In addition, if not sent regularly the dynamic nature of the information provided would make it less and less useful.

3. Benefits with non-compound packets

As mentioned in the introduction, most advantages of using non-compound packets exists in cases when the available RTCP bit-rate is limited. This because non-compound packets will be substantially smaller than compound packets. A compound packet is forced to contain both an RR or an SR and the CNAME SDES item. The RR containing a report block for a single source is 32 bytes, an SR is 52 bytes. Both may be larger if they contain report blocks for multiple sources. The SDES packet containing a CNAME item will be 10 bytes plus the CNAME string length. Here it is reasonable that the CNAME string is at least 10 bytes to get a decent collision resistance. And if the recommended form of user@host is used, then most strings will be longer than 20 characters. Thus a non-compound packets can become at least 70-80 bytes smaller than the compound packet.

The following benefits exist for the smaller non-compound packets:

1. Shorter serialization time, i.e the time it takes the link to transmit the packet. For slower links this time can be substantial. For example transmitting 120 bytes over an link interface capable of 30 kbps takes 32 milliseconds (ms) assuming uniform transmission rate.

2. For links where the packet loss rate grows with the packet size, smaller packets will be less likely to be dropped. An example of such links are radio links. In the cellular world there exist links that are optimized to handle RTP packets sized for carrying compressed speech, which increases the capacity and coverage for voice services in a given wireless network. Minimum sized compound RTCP packets are commonly 2-3 times the size of a RTP packet carrying compressed speech. If the speech packet over such a bearer has a packet loss probability of p, then the RTCP packet will experience a loss probability of 1- (1-p)^x where x is the number of fragments the compound packet will be split on the link layer, i.e. commonly into 2 or 3 fragments.

3. Independently of the link type there are additional benefits with sending feedback in small non-compound RTCP. One such example is applications that use RTCP AVPF in early or immediate mode to send frequent event driven feedback. Under these circumstances non-compound RTCP reduces the risk that the RTCP bandwidth becomes too high during periods of heavy adaptation feedback signaling.

4. In cases when regular feedback is needed, such as the profile under development for TCP friendly rate control (TFRC) for RTP [I-D.ietf-avt-tfrc-profile], the size of compound RTCP can result in very high bandwidth requirements if the round trip time is short. For this particular application non-compound RTCP gives a very substantial improvement.

In cases when non-compound packets carry important and time sensitive feedback both shorter serialization time and the lower loss probability are important to enable the best possible functionality. Having a packet loss rate that is much higher for the feedback packets compared to media packets hurts when trying to perform media adaptation, to for example handle the changed performance present at the cell border in cellular system.

For high bit-rate applications there is usually no problem of supplying RTCP with sufficient bit-rates. When using AVPF one can use the "trr-int" parameter to restrict the regular reporting interval to approximately once per RTT or less often. As in most

cases there are no reasons to provide regular reports with higher density than this. Any additional bandwidth can then be used for feedback messages. The benefit of non-compound packets in this case is limited, but exists. One typical example is video using generic NACK in cases where the RTT is low. Using non-compound packets would reduce the total amount of bits used for RTCP. This is primarily applicable if the number of non-compound packets is large. This would also result in lower processing delay and less complexity for the feedback packets as they do not need to query the RTCP database to construct the right messages.

4. Issues with non-compound RTCP packets

This section describes some of the known issues with non-compound RTCP packets

4.1. Middle boxes

Middle boxes in the network may discard RTCP packets that do not follow the rules outlined in section 6.1 of RFC3550. The effect of this might for instance be that compound RTCP packets would get through while the non-compound feedback packets would be lost.

4.2. Packet Validation

A non-compound packet will be discarded by the packet validation code in Appendix A of RFC 3550 [RFC3550]. This has several impacts as described in the following sub sections.

4.2.1. Old RTCP Receivers

Any RTCP receiver without updated packet validation code will discard the non-compound packets. Thus these receivers will not see the feedback contained in the these non-compound packets. The effect of this depends on the type of feedback message and the role of the receiver. For example this may cause complete function loss in the case of attempting to use a non-compound NACK message (see Section 6.2.1 of RFC 4585 [RFC4585]) to non updated media sender in a session using the retransmission scheme defined by RFC 4588 [RFC4588].

This type of discarding would also effect the feedback suppression defined in AVPF. The result would be a partitioning of the receivers within the session between old ones only seeing the compound RTCP feedback messages and the newer ones seeing both. Where the old ones may send feedback messages for events already reported on in non-compound packets.

4.2.2. Weakened Packet Validation

The packet validation code needs to be rewritten to accept non-compound packets. One potential effect of this change is much weaker validation that received packets actually are RTCP packets, and not packets of some other type being wrongly delivered. Thus some consideration should be done to ensure the best possible validation is available. For example restricting non-compound packets to contain only some specific RTCP packet types, that is preferably signalled on a session basis. A solution to this is presented in Section 6.2

4.2.3. Bandwidth consideration

The discarding of non-compound RTCP packets would effect the RTCP transmission calculation in the following way; the avg_rtcp_size value would become larger than for RTP receivers that exclude the non-compound in this calculation (assuming that non-compound packets are smaller than compound ones). Therefore these senders would under-utilize the available bit-rate and send with a longer interval than updated receivers. For most sessions this should not be an issue. However for sessions with a large portion of non-compound packets may result in that the updated receivers time out non-updated senders prematurely. A solution to this is presented in Section 6.2.

4.2.4. Computation of avg_rtcp_size

Long intervals between compound RTCP packets and many non-compound RTCP packets in between may lead to a computation of a value for avg_rtcp_size that varies greatly over time. This is discussed more in Section 6.2.

4.3. Header compression

The classifiers for header compression algorithms such as RoHC [RFC3095] and its profiles must be aware of the fact that, with the proposed non-compound RTCP packets, the first RTCP packet type might differ from 200 or 201. Otherwise they would likely wrongly classify the packets as something else than RTCP. However, as no header compression technology defined in IETF compresses RTCP, this should have no real impact.

4.4. RTP and RTCP multiplex on the same port

In applications which multiplex RTP and RTCP on the same port, as defined in [I-D.ietf-avt-rtp-and-rtcp-mux], care must be taken to ensure that the de-multiplexing is done properly even though RTCP packets are non-compound.

4.5. Encryption/authentication

SRTP presents a problem for non-compound RTCP. Section 3.4 in [RFC3711] states "SRTCP MUST be given packets according to that requirement in the sense that the first part MUST be a sender report or a receiver report".

However the same text also states that the encryption prefix that is present in the receiver and sender reports should not be used by SRTP. The conclusion is therefore that it is possible to use non-compound RTCP with SRTP.

5. Use cases for non-compound RTCP

Below are listed a few use cases for non-compound RTCP. It is worth noting here that the current uses of non-compound RTCP are thoroughly specified in other standardization bodies and are limited to specific services such as PoC or 3GPP-MTSI. A general definition of the use of non-compound RTCP for e.g control plane or codec control signaling would probably need to be specified within the IETF.

5.1. Control plane signaling

Open Mobile Alliance (OMA) Push-to-talk over Cellular (PoC) [OMA-PoC] makes use of non-compound packets when transmitting certain events. The OMA POC service is primarily used over cellular links capable of IP transport, such as the GSM GPRS.

5.2. Codec control signaling

Examples of codec control usage for non-compound RTCP are found in [3GPP-MTSI].

Another example that can be used with non-compound RTCP is e.g TMMBR messages as specified in [I-D.ietf-avt-avpf-ccm] which signal a request for a change in codec bitrate. The benefit of non-compound RTCP for these messages is that in bad channel conditions, a non-compound RTCP can be considerably more likely to be received than larger compound RTCP messages. This is critical as these messages predominantly occur when channel conditions are poor.

5.3. Feedback

The feedback scenario is best presented as a Video stream with generic NACK. In cases where the RTT is shorter than the receiver buffer depth, generic NACK can be used to request retransmission of missing packets, thus improving play out quality considerably. If

the generic NACK packets are transmitted as non-compound packets, the bandwidth requirement for RTCP will be minimal, enabling more frequent feedback. Like in the Codec control case it is important that these packets can be transmitted with as little delay as possible. The RTCP bandwidth reduction and transmission speed are equally useful when retransmission is not used for loss recovery.

Another interesting use for non-compound RTCP is in cases when regular feedback is needed, such as the profile under development for TCP friendly rate control (TFRC) for RTP [I-D.ietf-avt-tfrc-profile]. The size of compound RTCP can result in very high bandwidth requirements for the feedback when the round trip time is short. For this particular application non-compound RTCP may give a very substantial improvement.

5.4. Status reports

One idea proposed is to transmit small measurement or status reports in non-compound RTCP, and to be able to split the sub-packets of a minimum compound RTCP and transmit them separately. The status reports can be used either by the endpoints or by other network monitoring boxes in the network.

The benefit is that with some radio access technologies small packets are more robust to poor radio conditions than large packets. Additionally, with small (report) packets there is a smaller risk that the report packets will affect the channel that they report upon.

Even though this may be an interesting use case a few issues needs to be considered.

- o A risk exists that it opens up for a whole set of incompatible metrics and reports devised in various standardization fora leading to a potential interoperability problems.

- o Middle boxes or third party network monitoring equipment may fail to understand the new reports or even discard these new report types.

- o There may arise a need to verify that these "special" reports reach the intended recipient in case middle boxes in the network discards unknown reports. In many cases it is difficult to verify that for instance sender reports are received correctly, missing SR may as well be an indication that the other endpoint has terminated.

6. Rules and guidelines for non-compound packets in AVPF

Based on the above analysis it seems feasible to allow transmission of non-compound RTCP under some restrictions. First of all it is important that compound packets are regularly sent to ensure the feedback reporting works. The tracking of session size and number of participants is also important as this ensures that the RTCP bandwidth remain bounded independent of the number of session participants. As the compound packets also are used to establish the synchronization, any newly joining participant in a session would need to receive a compound packet from the media sender. In summary the regular usage of compound packets must be maintained throughout the complete session. Thus non-compound packets should be restricted to be used as extra feedback packets sent in cases when a regular compound packet would not have been sent.

The usage of non-compound RTCP packet SHALL only be done in RTP sessions operating in AVPF [RFC4585] Early RTCP or Immediate feedback mode. Non-compound packets SHALL NOT be sent until at least one compound packet has been sent. In Immediate feedback mode all feedback messages MAY be sent as non-compound packets. In early RTCP mode a feedback message scheduled for transmission as an Early RTCP packet, i.e not a Regular RTCP packet, MAY be sent as a non-compound packet. All packets that scheduled for transmission as Regular RTCP packets SHALL be sent as (full) compound RTCP packets as indicated by AVPF [RFC4585].

6.1. Verification of the delivery of non-compound packets

If an application is to use non-compound packets it is important to verify that they actually reaches the session participants. As outlined above in Section 4.1 and Section 4.2 packets may be discarded along the path or in the end-point. The end-points can be resolved by introducing signaling that informs if all session participants are capable of non-compound packets or not. The middle box issue is more difficult and here one will be required to use heuristics to determine if the non-compound packets are delivered or not. However in many cases the feedback messages sent using non-compound packets will result in either explicit or implicit indications that they have been received. Example of such are the RTP retransmission [RFC4588] that result from a NACK message [RFC4585], the Temporary Maximum Media Bit-rate Notification message resulting from a Temporary Maximum Media Bit-rate Request [I-D.ietf-avt-avpf-ccm], or the presence of a Decoder Refresh Point [I-D.ietf-avt-avpf-ccm] in the video media stream resulting from the Full Intra Request sent.

A proposed algorithm to detect consistent failure of delivery of non-

compound packets needs to be written. The details of this algorithm is application dependent and therefore outside the scope of this document.

If the verification fails it is strongly RECOMMENDED that only compound RTCP according to the rules outlined in RFC3550 is transmitted.

6.2. Algorithm modifications

6.2.1. Distinction between compound and non-compound RTCP

One question that arise is how to distinguish between small (non-compound) and large (compound) RTCP. A few alternatives:

- o Payload type: A non-compound RTCP may have a (first) PT number that differs from the PT numbers for SR or RR. This may be a weak alternative as some interest to be able to split minimum compound RTCP is expressed, see Status reports (Section 5.4). A possible problem here is also that this distinction does not actually tell the size of the RTCP.

- o Fixed size, set in specification. For instance one may base the distinction on the likely minimum size of a minimal compound RTCP. Assuming that such a packet will contain at least an SR (32 bytes) and a SDES CNAME (likely 16 bytes or more) one can conclude that 48 bytes (+IP/UDP overhead) is probably the smallest realistic size of a compound RTCP.

- o Fixed size, set in session setup : Some sessions may e.g use RTCP-XR or some other RTCP reporting on occasions that may give very large packet sizes, it may be desirable to adjust the threshold

- o Variable size: As non-compound RTCP is by definition RTCP that does not follow the rules for compound RTCP as they are specified in RFC3550, the size can be determined "on the fly".

- o Comparison against total UDP packet size: If the size of the first RTCP packet is smaller then the UDP packet size it is a compound RTCP. If the sizes match then it is a non-compound RTCP. This method is an indication of the number of RTCP packets in a given UDP payload, if 2 RTCP packets or more then it is a compound RTCP.

6.2.2. Modified bandwidth algorithms

The use of non-compound RTCP does not imply any specific need for algorithm modifications. A few possible algorithm modifications have been tested and even though the modifications may improve the performance when feedback is transmitted the benefits are not judged large enough to justify the relatively large changes in the algorithms.

A more extensive report covering various test-cases and different algorithm modifications can be downloaded from: http://www.ijdata.com/rtcp-non-compound-rtcp-evaluation.doc

The report also covers tests with other averaging factors than the specified 1/16, and it shows that it is beneficial to use slower averaging (1/32 or 1/64) as it makes the estimate avg_rtcp_size more stable and does not degrade the feedback transmission performance. This modification is however not critical.

6.2.3. Immediate mode

Section 3.3 in RFC4585 gives the option to use AVPF Immediate mode as long as the groupsize is below a certain limit. As feedback using non-compound RTCP becomes smaller it opens up for a more liberal use of immediate mode.

6.2.4. Enforcing compound RTCP

As discussed earlier it is important that the sending of compound RTCP packets do happen at regular interval. However, this will occur as long as the RTCP senders follow the AVPF scheduling algorithm defined in Section 3.5 in [RFC4585]. This as all regular RTCP packets must be full compound RTCP packets. Note that also in immediate mode is there a requirement on sending regular RTCP packets.

6.3. Open issues

This section contains a list of the yet unresolved issues.

- o SRTP : Does non-compound RTCP break SRTP behavior ?

6.4. SDP Signalling Attribute

We request to define the a "a=rtcp-nc" [RFC4566] attribute to indicate if the session participant is capable of supporting non-compound packets. It is a required that a participant that proposes the use of non-compound RTCP itself supports the reception of non-

compound RTCP.

An offering client that wish to use non-compound RTCP MUST include the attribute "a=rtcp-nc" in the SDP offer. If the other client does not support non-compound RTCP the attribute MUST be removed from the answer SDP.

7. IANA Considerations

IANA will be required to register the SDP signalling attribute defined in Section 6.4.

8. Security Considerations

The security considerations of RTP [RFC3550] and AVPF [RFC4585] will apply also to non-compound packets. The reduction in validation strength for received packets on the RTCP port may result in a higher degree of acceptance of spurious data as real RTCP packets. This vulnerability can mostly be addressed by usage of an security mechanism that provide authentication, e.g. SRTP[RFC3711].

9. Acknowledgements

The authors would like to thank all the people who gave feedback on this document.

This document also contain some text copied from [RFC3550], [RFC4585]and [RFC3711]. We take the opportunity to thank the authors of said documents.

10. References

10.1. Normative References

[RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550, July 2003.

[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, July 2006.

10.2. Informative References

[3GPP-MTSI]
3GPP, "Specification : 3GPP TS 26.114 (v7.1.0 preliminary), http://www.3gpp.org/ftp/tsg_sa/WG4_CODEC/Specs_update_after_SA36/26114-710.zip", March 2007.

[I-D.ietf-avt-avpf-ccm]
Wenger, S., Chandra, U., Westerlund, M., and B. Burman, "Codec Control Messages in the RTP Audio-Visual Profile with Feedback (AVPF)", draft-ietf-avt-avpf-ccm-10 (work in progress), October 2007.

[I-D.ietf-avt-rtp-and-rtcp-mux]
Perkins, C. and M. Westerlund, "Multiplexing RTP Data and Control Packets on a Single Port", draft-ietf-avt-rtp-and-rtcp-mux-07 (work in progress), August 2007.

[I-D.ietf-avt-tfrc-profile]
Gharai, L., "RTP with TCP Friendly Rate Control", draft-ietf-avt-tfrc-profile-10 (work in progress), July 2007.

[OMA-PoC] Open Mobile Alliance, "Specification : Push to talk Over Cellular User Plane, http://www.openmobilealliance.org/release_program/docs/PoC/V1_0_1-20061128-A/OMA-TS-PoC-UserPlane-V1_0_1-20061128-A.pdf", November 2006.

[RFC3095] Bormann, C., Burmeister, C., Degermark, M., Fukushima, H., Hannu, H., Jonsson, L-E., Hakenberg, R., Koren, T., Le, K., Liu, Z., Martensson, A., Miyazaki, A., Svanbro, K., Wiebke, T., Yoshimura, T., and H. Zheng, "RObust Header Compression (ROHC): Framework and four profiles: RTP, UDP, ESP, and uncompressed", RFC 3095, July 2001.

[RFC3711] Baugher, M., McGrew, D., Naslund, M., Carrara, E., and K. Norrman, "The Secure Real-time Transport Protocol (SRTP)", RFC 3711, March 2004.

[RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, July 2006.

[RFC4588] Rey, J., Leon, D., Miyazaki, A., Varsa, V., and R. Hakenberg, "RTP Retransmission Payload Format", RFC 4588, July 2006.

Authors' Addresses

Ingemar Johansson
Ericsson AB
Laboratoriegrand 11
SE-971 28 Lulea
SWEDEN

Phone: +46 73 0783289
Email: ingemar.s.johansson (AT) ericsson.com

Magnus Westerlund
Ericsson AB
Torshamnsgatan 21-23
SE-164 83 Stockholm
SWEDEN

Phone: +46 8 7190000
Email: magnus.westerlund (AT) ericsson.com

Full Copyright Statement

Intellectual Property

Acknowledgment

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).