

Audio/Video Transport Working
Group
Internet-Draft
Intended status: Informational
Expires: November 19, 2010

G. Hunt
P. Arden
BT
May 18, 2010

Monitoring Architectures for RTP
draft-hunt-avt-monarch-01.txt

Abstract

This memo proposes an architecture for extending RTCP with a new RTCP XR (RFC3611) block type to report new metrics regarding media transmission or reception quality, as proposed in draft-ietf-avt-rtcp-guidelines (work in progress [replace with RFC number]). This memo suggests that a new block should contain a single metric or a small number of metrics relevant to a single parameter of interest or concern, rather than containing a number of metrics which attempt to provide full coverage of all those parameters of concern to a specific application. Applications may then "mix and match" to create a set of blocks which covers their set of concerns. Where possible, a specific block should be designed to be re-usable across more than one application, for example, for all of voice, streaming audio and video.

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 November 19, 2010.

Copyright Notice

Copyright (c) 2010 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. Requirements notation	4
3. Using small blocks	5
4. The identity block	6
5. An example of a metric block	10
6. Application to translators	11
7. Application to RFC 5117 topologies	12
8. Expanding the RTCP XR block namespace	13
9. IANA Considerations	14
10. Security Considerations	15
11. Informative References	16
Authors' Addresses	17

1. Introduction

Any proliferation of metrics for transport and application quality monitoring has been identified as a potential problem for RTP/RTCP interoperability. Different applications layered on RTP may have some monitoring requirements in common, which should be satisfied by a common design. The objective here is to define an extensible framework and a small number of re-usable metrics to reduce implementation costs and to maximise inter-operability. Work-in-progress on [GUIDELINES] has stated that, where RTCP is to be extended with a new metric, the preferred mechanism is by the addition of a new RTCP XR [RFC3611] block. This memo assumes that any requirement for a new metric to be transported in RTCP will use a new RTCP XR block.

[GUIDELINES] provides advice on when and how new metrics should be introduced, including recommending that metrics are based on existing standards whenever possible.

Section 3 describes the key proposal of this memo, the use of small metrics blocks each of which addresses a single parameter of interest which may be "mixed and matched", rather than providing a large block to address all the parameters which might be of interest to a broad class of applications (for example, all VoIP applications).

Section 4 describes an optimisation to avoid repetition of identification information, which becomes desirable when small blocks are used.

Section 5 provides an example of the application of these principles to a specific case, that of a metric block to report packet delay variation.

Section 6 draws attention to the guidance in [RFC3550] concerning RTCP and translators.

Section 7 discusses the potential application of RTCP XR metrics blocks to the conferencing topologies discussed in [RFC5117].

Section 8 consists (in this draft) only of an "Editor's note" asking whether the limited namespace available for RTCP XR blocks is a concern, and if so whether it would be desirable to work on a standardised means to expand it.

2. Requirements notation

This memo is informative and as such contains no normative requirements.

3. Using small blocks

Different applications using RTP for media transport certainly have differing requirements for metrics transported in RTCP to support their operation. For many applications, the basic metrics for transport impairments provided in RTCP SR and RR packets [RFC3550] (together with source identification provided in RTCP SDES packets) are sufficient. For other applications additional metrics may be required or at least sufficiently useful to justify the overheads, both of processing in endpoints and of increased session bandwidth. For example an IPTV application using Forward Error Correction (FEC) might use either a metric of post-repair loss or a metric giving detailed information about pre-repair loss bursts to optimise payload bandwidth and the strength of FEC required for changing network conditions. However there are many metrics available. It is likely that different applications or classes of applications will wish to use different metrics. Any one application is likely to require metrics for more than one parameter but if this is the case, different applications will almost certainly require different combinations of metrics. If larger blocks are defined containing multiple metrics to address the needs of each application, it becomes likely that many different such larger blocks are defined, which becomes a danger to interoperability.

To avoid this pitfall, this memo proposes the use of small RTCP XR metrics blocks each containing a very small number of individual metrics characterising only one parameter of interest to an application running over RTP. For example, at the RTP transport layer, the parameter of interest might be packet delay variation, and specifically the metric "IPDV" defined by [Y1540]. See Section 5 for architectural considerations for a metrics block, using as an example a metrics block to report packet delay variation.

4. The identity block

Any measurement must be identified. However if metrics are delivered in small blocks there is a danger of inefficiency arising from repeating this information in a number of metrics blocks within the same RTCP packet, in cases where the same identification information applies to multiple metrics blocks.

An instance of a metric must be identified using information which is likely to include most of the following:

- o the node at which it was measured,
- o the source of the measured stream (for example, its CNAME),
- o the SSRC of the measured stream,
- o the sequence number of the first packet of the RTP session,
- o the extended sequence numbers of the first packet of the current measurement interval, and the last packet included in the measurement,
- o the duration of the most recent measurement interval and
- o the duration of the interval applicable to cumulative measurements (which may be the duration of the RTP session to date).

[Editor's note: this set of information overlaps with, but is more extensive than, that in the union of similar information in RTCP RR packets. Should we assume that RR information is always present if XR is sent, and that measurement intervals are exactly coincident? If so, state assumption and remove overlaps. What were the design considerations which led to the additional information *not* being present in RRs? The reason for the additional information here is the perceived difficulty of "locating" the *start* of the RTP session (sequence number of 1st packet, duration of interval applicable to cumulative measurements) using only RR. Is this a misconception? It leads to redundant information in this design because equivalent information is provided multiple times, once in *every* identification packet. Though this ensures immunity to packet loss, the design is ugly and the overhead is not completely trivial.]

This section proposes an approach to minimise the inefficiency of providing this identification information, assuming that an architecture based on small blocks means that a typical RTCP packet will contain more than one metrics block needing the same identification. The choice of identification information to be

provided is discussed in [IDENTITY] (work in progress).

The approach is to define a stand-alone block containing only identification information, and to tag this identification block with a number which is unique within the scope of the containing RTCP XR packet. The "containing RTCP XR packet" is defined here as the RTCP XR header with PT=XR=207 defined in Section 2 of [RFC3611] and the associated payload defined by the length field of this RTCP XR header. The RTCP XR header itself includes the SSRC of the node at which all of the contained metrics were measured, hence this SSRC need not be repeated in the stand-alone identification block. A single containing RTCP XR packet may contain multiple identification blocks limited by the range of the tag field. Typically there will be one identification block per monitored source SSRC, but the use of more than one identification block for a single monitored source SSRC within a single containing RTCP XR packet is not ruled out.

There will be zero or more metrics blocks dependent on each identification block. The dependence of an instance of a metrics block on an identification block is established by the metrics block's having the same numeric value of the tag field as its identification block (in the same containing RTCP XR packet).

Figure 1 below illustrates this principle using as an example an RTCP XR packet containing four metrics blocks, reporting on streams from two sources. The measurement identity information is provided in two blocks with Block Type NMI, and tag values 0 and 1 respectively.

Note: in this example, RTCP XR block type values for four proposed new block types (work in progress) are given as NMI, NPDV, NBGL and NDEL. These represent numeric block type codepoints to be allocated by IANA at the conclusion of the work.

Each of these two identity blocks will specify the SSRC of one of the monitored streams, as well as information about the span of the measurement. There are two metrics blocks with tag=0 indicating their association with the measurement identity block which also has tag=0. These are the two blocks following the identity block with tag=0, though this positioning is not mandatory. There are also two metrics blocks with tag=1 indicating their association with the measurement identity block which also has tag=1, and these are the two blocks following the identity block with tag=1.

[Editor's note: if we mandated that metrics blocks associated with an identity block must always follow the identity block we could save the tag field and possibly simplify processing. Is this preferable to cross-referencing with a numeric tag?]

In the example, the block types of the metrics blocks associated with tag=0 are BT=NPDV (a PDV metrics block) and BT=NBGL (a burst and gap loss metrics block). The block types of the metrics blocks associated with tag=1 are BT=NPDV (a second PDV metrics block) and BT=NDEL (a delay metrics block). This illustrates that:

- o multiple instances of the same metrics block may occur within a containing RTCP XR packet, associated with different identification information, and
- o differing measurements may be made, and reported, for the different streams arriving at an RTP system.

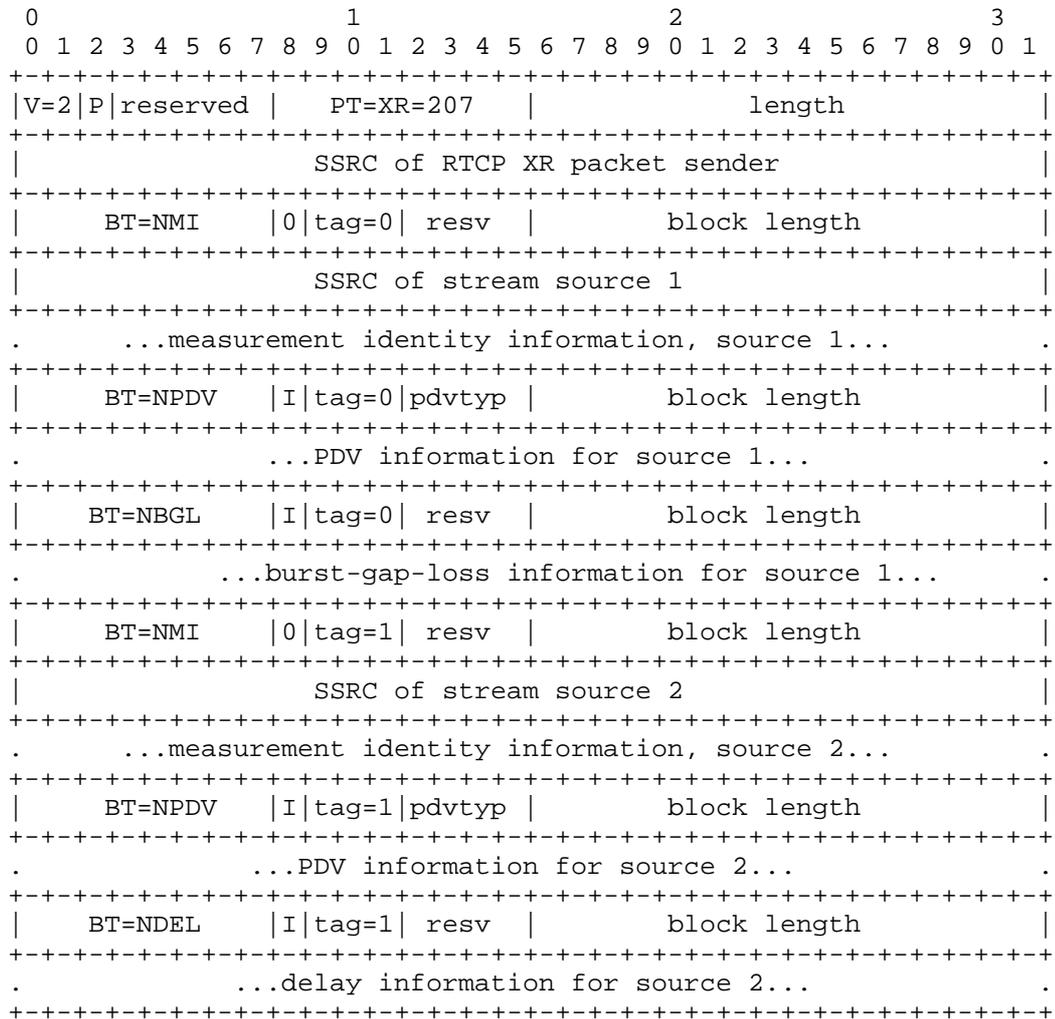


Figure 1: RTCP XR block with identity blocks

This approach of separating the identification information is more costly than providing identification in each metrics block if only a single metrics block is sent in an RTCP packet, but becomes beneficial as soon as more than one metrics block shares common identification.

5. An example of a metric block

This section uses the example of an existing proposed metrics block to illustrate the application of the principles set out in Section 3.

The example [PDV] (work in progress) is a block to convey information about packet delay variation (PDV) only, consistent with the principle that a metrics block should address only one parameter of interest. One simple metric of PDV is available in the RTCP RR packet as the "jit" field. There are other PDV metrics which may be more useful to certain applications. Two such metrics are the IPDV metric ([Y1540], [RFC3393]) and the MAPDV2 metric [G1020]. Use of these metrics is consistent with the principle in Section 5 of [GUIDELINES] that metrics should usually be defined elsewhere, so that RTCP standards define only the transport of the metric rather than its nature. The purpose of this section is to illustrate the architecture using the example of [PDV] (work in progress) rather than to document the design of the PDV metrics block or to provide a tutorial on PDV in general.

Given the availability of at least three metrics for PDV, there are design options for the allocation of metrics to RTCP XR blocks:

- o provide an RTCP XR block per metric
- o provide a single RTCP XR block which contains all three metrics
- o provide a single RTCP block to convey any one of the three metrics, together with a identifier to inform the receiving RTP system of the specific metric being conveyed

In choosing between these options, extensibility is important, because additional metrics of PDV may well be standardised and require inclusion in this framework. The first option is extensible but only by use of additional RTCP XR blocks, which may consume the limited namespace for RTCP XR blocks at an unacceptable rate. The second option is not extensible, so could be rejected on that basis, but in any case a single application is quite unlikely to require transport of more than one metric for PDV. Hence the third option was chosen. This implies the creation of a subsidiary namespace to enumerate the PDV metrics which may be transported by this block, as discussed further in [PDV] (work in progress).

6. Application to translators

Section 7.2 of [RFC3550] describes processing of RTCP by translators. RTCP XR is within the scope of the recommendations of [RFC3550]. Some RTCP XR metrics blocks may usefully be measured at, and reported by, translators. As described in [RFC3550] this creates a requirement for the translator to allocate an SSRC for itself so that it may populate the SSRC in the RTCP XR packet header (although the translator is not a Synchronisation Source in the sense of originating RTP media packets). It must also supply this SSRC and the corresponding CNAME in RTCP SDES packets.

In RTP sessions where one or more translators generate any RTCP traffic towards their next-neighbour RTP system, other translators in the session have a choice as to whether they forward a translator's RTCP packets. Forwarding may provide additional information to other RTP systems in the connection but increases RTCP bandwidth and may in some cases present a security risk. RTP translators may have forwarding behaviour based on local policy, which might differ between different interfaces of the same translator.

[Editor's note: for bidirectional unicast, an RTP system may usually detect RTCP from a translator by noting that the sending SSRC is not present in any RTP media packet. However even for bidirectional unicast there is a possibility of a source sending RTCP before it has sent any RTP media (leading to transient mis-categorisation of an RTP end system or RTP mixer as a translator), and for multicast sessions - or unidirectional/streaming unicast - there is a possibility of a receive-only end system being permanently mis-categorised as a translator. Is there a need for a translator to declare itself explicitly? Needs further thought.]

7. Application to RFC 5117 topologies

An RTP system (end system, mixer or translator) which originates, terminates or forwards RTCP XR blocks is expected to handle RTCP, including RTCP XR, as specified in [RFC3550] for that class of RTP systems. Provided this expectation is met, an RTP system using RTCP XR is architecturally no different from an RTP system of the same class (end system, mixer, or translator) which does not use RTCP XR. This statement applies to the topologies investigated in [RFC5117], where they use RTP end systems, RTP mixers and RTP translators as these classes are defined in [RFC3550].

These topologies are specifically Topo-Point-to-Point, Topo-Multicast, Topo-Translator (both variants, Topo-Trn-Translator and Topo-Media-Translator, and combinations of the two), and Topo-Mixer.

The topologies based on systems which do not behave according to [RFC3550], that is Topo-Video-Switch-MCU and Topo-RTCP-terminating-MCU, suffer from the difficulties described in [RFC5117]. These difficulties apply to systems sending, and expecting to receive, RTCP XR blocks as much as to systems using other RTCP packet types. For example, a participant RTP end system may send media to a video switch MCU. If the media stream is not selected for forwarding by the switch, neither RTCP RR packets nor RTCP XR blocks referring to the end system's generated stream will be received at the RTP end system. Strictly the RTP end system can only conclude that its RTP has been lost in the network, though an RTP end system complying with the robustness principle of [RFC1122] should survive with essential functions unimpaired.

8. Expanding the RTCP XR block namespace

[Editor's note: the RTCP XR block namespace is limited by the 8-bit block type field in the RTCP XR header (Section 3 of [RFC3611]). IESG have noted that this is potentially restrictive. It would be possible to standardise an expansion mechanism, probably based on use of a new field near the start of the variable-length "type-specific block contents" field. Clearly this could apply only to new block types, so might be standardised to apply to some subrange of the current 8-bit range, for example the range 128 through 191 might be used. At time of writing, block types 12 to 254 are unassigned and 255 is reserved for future expansion. Is there a consensus for, or against, work to allow expansion? One potential use is through hierarchical control, where one or a few codepoints at the top level are given to other SDOs who may then define a number of metrics distinguished by values in the (so far hypothetical) new field.]

9. IANA Considerations

None.

10. Security Considerations

This document itself contains no normative text and hence should not give rise to any new security considerations, to be confirmed.

11. Informative References

- [G1020] ITU-T, "ITU-T Rec. G.1020, Performance parameter definitions for quality of speech and other voiceband applications utilizing IP networks", July 2006.
- [GUIDELINES] Ott, J., "Guidelines for Extending the RTP Control Protocol (RTCP)", ID draft-ott-avt-rtcp-guidelines-03, February 2010.
- [IDENTITY] Hunt, G., "RTCP XR Report Block for Measurement Identity", ID draft-ietf-avt-rtcp-xr-meas-identity-02, May 2009.
- [PDV] Hunt, G., "RTCP XR Report Block for Packet Delay Variation Metric Reporting", ID draft-ietf-avt-rtcp-xr-pdv-03, May 2009.
- [RFC1122] Braden, R., "Requirements for Internet Hosts -- Communication Layers", RFC 1122, October 1989.
- [RFC3393] Demichelis, C., "IP Packet Delay Variation Metric for IP Performance Metrics (IPPM)", RFC 3393, November 2002.
- [RFC3550] Schulzrinne, H., "RTP: A Transport Protocol for Real-Time Applications", RFC 3550, July 2003.
- [RFC3611] Friedman, T., "RTP Control Protocol Extended Reports (RTCP XR)", RFC 3611, November 2003.
- [RFC5117] Westerlund, M., "RTP Topologies", RFC 5117, January 2008.
- [Y1540] ITU-T, "ITU-T Rec. Y.1540, IP packet transfer and availability performance parameters", November 2007.

Authors' Addresses

Geoff Hunt
BT
Orion 1 PP2
Adastral Park
Martlesham Heath
Ipswich, Suffolk IP5 3RE
United Kingdom

Phone: +44 1473 651704
Email: geoff.hunt@bt.com

Philip Arden
BT
Orion 3/7 PP4
Adastral Park
Martlesham Heath
Ipswich, Suffolk IP5 3RE
United Kingdom

Phone: +44 1473 644192
Email: philip.arden@bt.com

