

Version 0.3
May 6, 2011

(1) Introduction

This document provides recommendations and guidelines for RTP and RTCP in the context of SIPREC. In order to communicate most effectively, the Session Recording Client (SRC) and the Session Recording (SRS) should utilize the mechanisms provided by RTP in a well defined and predicable manner. It is the goal of this document to make the reader aware of these mechanisms and provide recommendations and guidelines. This document is completely informational. It includes no requirements and no normative language.

(2) Definitions

This document uses the definitions provided in draft-ietf-siprec-architecture for all SIPREC modules. A Session Recording Client (SRC), as defined within SIPREC, is expected to implement one or more of a variety of RTP roles within the context of various SIPREC use cases. These roles include, but are not limited to, acting as an end system, a mixer, or a translator. This document uses the definitions provided in "RTP: A Transport Protocol for Real-Time Application" [RFC 3550] for these roles.

----- R3.x RTP Roles

An SRC has the task of gathering RTP media from the various participants in a CS and forwarding the information to the SRS in the context of the RS. For a given media stream the SRC has a number of ways of doing this and may choose different ways for different media streams.

R3.x.1 SRC acting as an RTP Translator.

The SRC may act as a translator, as defined in [RFC3550]. In this case, the SRC forwards RTP packets to the SRS with their synchronization source identifier unchanged. Optionally an SRC acting as a translator can perform transcoding (from one codec to another), and this can result in a different rate of packets.

With this approach, RTP from different sources (i.e., from different participants) cannot be mixed by the SRC and must be sent separately to the SRS.

In this case all RTCP reports are passed on by the SRC to the participants which will be able to detect any SSRC collisions.

RTCP packets Sender Reports will generated by a participant sending a stream and will need to be forwarded to the SRS (in addition to being sent to receiving participants). RTCP Receiver Reports will be generated by the SRS and need to be forwarded to the relevant participant, the SRC in the case may need to manipulate the RTCP Receiver Reports to take account of any transcoding that has taken place.

OPEN ISSUE. With this approach, can SRTP be used? Can SRTP be used on the CS and not on the RS and vice versa? If SRTP is used on both the CS and the RS, can decryption and re-encryption occur and/or can decryption and re-encryption be avoided (clearly not when transcoding is involved or if different keys are to be used?

R3.x.2 SRC acting as an RTP Mixer.

In the case of the SRC acting as a RTP mixer, as defined in [RFC3550], the SRC combines RTP streams from different participants and sends them towards the SRS using its own synchronization source identifier. It will make timing adjustments among the received streams and generate its own timing on the stream sent to the SRS. Optionally an SRC acting as a mixer can perform transcoding, and can even cope with different codings received from different participants. RTCP Sender Reports and Receiver Reports are not forwarded by an SRC acting as mixer, but there are requirements for forwarding RTCP Source Description (SDES) packets. The use of SRTP between the SRC and the SRS is independent of the use of SRTP on the CS.

R3.x.3 SRC acting as an RTP Endpoint.

The case of the SRC acting as an RTP endpoint, as defined in [RFC3550], is similar to the mixer case, except that the RTP session between the SRC and the SRS is considered completely independent from the RTP session that is part of the CS. The SRC can, but need not, mix RTP streams from different participants prior to sending to the SRS. RTCP between the SRC and the SRS is completely independent of RTCP on the CS. The use of SRTP between the SRC and the SRS is independent of the use of SRTP on the CS.

(3) RTCP

The RTP data transport is augmented by a control protocol (RTCP) to allow monitoring of the data delivery. RTCP, as defined in [RFC3550], is based on the periodic transmission of control packets to all participants in the RTP session, using the same distribution mechanism as the data packets. Support for RTCP is required by RFC 3550, and it provides, among other things, the following important functionality in relation to SIPREC:

- 1) Feedback on the quality of the data distribution. This feedback is important for flow and congestion control functions, and to get feedback from the receivers to diagnose faults in the distribution. As such, RTCP is a well defined and efficient mechanism for the SRS to inform the SRC or any issues that arise in the reception of media that is to be recorded.
- 2) Carries a persistent transport-level identifier for an RTP source called the canonical name or CNAME. The SSRC identifier may change if a conflict is discovered or a program is restarted; in which case receivers can use the CNAME to keep track of each participant. Receivers may also use the CNAME to associate multiple data streams from a given participant in a set of related RTP sessions, for example to synchronize audio and video. Synchronization of media streams is also facilitated by the NTP and RTP timestamps included in RTCP packets by data senders.

(4) RTP Profile

The recommended RTP profiles for both the SRC and SRS are "Extended Secure RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/SAVPF)", [RFC5124] when using encrypted RTP streams, and "Extended RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback

(RTP/AVPF)", [RFC4585] when using non encrypted media streams. However, as this is not a requirement, some implementations may use "The Secure Real-time Transport Protocol (SRTP)", [RFC 3711] and "RTP Profile for Audio and Video Conferences with Minimal Control", AVP [RFC 3551]. Therefore, it is recommended that the SRC and SRS not rely entirely on SAVPF or AVPF for core functionality that may be at least partially achievable using SAVP and AVP.

AVPF and SAVPF provide an improved RTCP timer model that allows more flexible transmission of RTCP packets as response to events, rather than strictly according to bandwidth. AVPF based codec control messages provide efficient mechanisms for an SRC and SRS to handle events such as scene changes, error recovery, and dynamic bandwidth adjustments. These messages are discussed in more detail later in this document.

SAVP and SAVPF provide media encryption, integrity protection, replay protection, and a limited form of source authentication. They do not contain or require a specific keying mechanism.

(5) SRSC

The synchronization source (SSRC), as defined in [RFC3550], is carried in the RTP header and in various fields of RTCP packets. It is a random 32-bit number that is required to be globally unique within an RTP session. It is crucial that the number be chosen with care in order that participants on the same network or starting at the same time are not likely to choose the same number. Guidelines regarding SSRC value selection and conflict resolution are provided in [RFC3550].

The SSRC may also be used to separate different sources of media within a single RTP session. For this reason as well as for conflict resolution, it is important that the SRC and SRS handle changes in SSRC values and properly identify the reason of the change. The CNAME values carried in RTCP facilitate this identification.

(6) CSRC

The contributing source (CSRC), as defined in [RFC3550], identifies the source of a stream of RTP packets that has contributed to the combined stream produced by an RTP mixer. The mixer inserts a list of the SSRC identifiers of the sources that contributed to the generation of a particular packet into the RTP header of that packet. This list is called the CSRC list. It is recommended that a SRC, when acting a mixer, sets the CSRC list accordingly, and that the SRS interprets the CSRC list appropriately when received.

(7) SDES

The Source Description (SDES), as defined in [RFC3550], contains an SSRC/CSRC identifier followed by a list of zero or more items, which carry information about the SSRC/CSRC. End systems send one SDES packet containing their own source identifier (the same as the SSRC in the fixed RTP header). A mixer sends one SDES packet containing a chunk for each contributing source from which it is receiving SDES information, or multiple complete SDES packets.

(8) CNAME

The Canonical End-Point Identifier (CNAME), as defined in [RFC3550], provide the binding from the SSRC identifier to an identifier for the source (sender or receiver) that remains constant. It is important the an SRC and SRS generate CNAMEs appropriately and use them for this purpose. Guidelines for generating CNAME values are provided in "Guidelines for Choosing RTP Control Protocol (RTCP) Canonical Names (CNAMEs)" [RFC6222].

(9) Keepalive

It is anticipated that media streams in SIPREC may exist in inactive states for extended periods of times for an of a number of valid reasons. In order to the bindings and any pinholes in NATs/firewalls to remain active during such intervals, it is recommended to follow the keep-alive procedures in "Application Mechanism for keeping alive the Network Address Translator (NAT) mappings associated to RTP/RTCP flows", [draft-ietf-avt-app-rtp-keepalive] for all RTP media streams.

(10) RTCP Feedback Messages

"Codec Control Messages in the RTP Audio-Visual Profile with Feedback (AVPF)" [RFC 5104] specifies extensions to the messages defined in AVPF [RFC4585]. Support for and proper usage of these messages is important to SRC and SRS implementations. Note that these messages are applicable only when using the AVFP or SAVPF RTP profiles.

(10.1) Full Intra Request

A Full Intra Request (FIR) Command, when received by the designate media sender, requires that the media sender sends a Decoder Refresh Point at the earliest opportunity. Using a decoder refresh point implies refraining from using any picture sent prior to that point as a reference for the encoding process of any subsequent picture sent in the stream.

Decoder refresh points, especially Intra or IDR pictures for H.264 video codecs, are in general several times larger in size than predicted pictures. Thus, in scenarios in which the available bit rate is small, the use of a decoder refresh point implies a delay that is significantly longer than the typical picture duration.

(10.1.1) SIP INFO for FIR

"XML Schema for Media Control" [RFC5168] defines an Extensible Markup Language (XML) Schema for video fast update. Implementations are discouraged from using the method described except for backward compatibility purposes. Implementations should use FIR messages, as described in (10.1) instead.

(10.2) Picture Loss Indicator

Using the FIR command to recover from errors is explicitly disallowed, and instead the PLI message defined in AVPF [RFC4585] should be used. The PLI message reports lost pictures and has been included in AVPF for precisely that purpose.

(10.3) Temporary Maximum Media Stream Bit Rate Request

A receiver, translator, or mixer uses the Temporary Maximum Media Stream Bit

Rate Request (TMMBR) to request a sender to limit the maximum bit rate for a media stream to the provided value. Appropriate use of TMMBR facilitates rapid adaptation to changes in available bandwidth.

(10.3.1) Renegotiation of SDP bandwidth attribute

If it is likely that the new value indicated by TMMBR will be valid for the remainder of the session, the TMMBR sender is expected to perform a renegotiation of the session upper limit using the session signaling protocol. Therefore for SIPREC, implementations are recommended to use TMMBR for temporary changes, and renegotiation of bandwidth via SDP offer/answer of more permanent changes.

(11) Symmetric RTP/RTCP for Sending and Receiving

Within the SDP offer/answer, RTP entities choose the RTP and RTCP transport addresses (i.e., IP addresses and port numbers) on which to receive packets. When sending RTP packets; however, they may use a different IP address or port number for RTP, RTCP, or both. Symmetric RTP/RTCP requires that the IP address and port number for sending and receiving RTP/RTCP packets are identical. Using Symmetric RTP and RTCP, as defined in "Symmetric RTP / RTP Control Protocol (RTCP)" [RFC4961] is recommended. It should not be confused with, or used in place of, "Multiplexing RTP Data and Control Packets on a Single Port" [RFC5761].