Internet Engineering Task Force (IETF)
Request for Comments: 7866
Category: Standards Track
ISSN: 2070-1721

L. Portman
NICE Systems
H. Lum, Ed.
Genesys
C. Eckel
Cisco
A. Johnston
Illinois Institute of Technology
A. Hutton
Unify

May 2016

Session Recording Protocol

Abstract

This document specifies the use of the Session Initiation Protocol (SIP), the Session Description Protocol (SDP), and the Real-time Transport Protocol (RTP) for delivering real-time media and metadata from a Communication Session (CS) to a recording device. The Session Recording Protocol specifies the use of SIP, SDP, and RTP to establish a Recording Session (RS) between the Session Recording Client (SRC), which is on the path of the CS, and a Session Recording Server (SRS) at the recording device. This document considers only active recording, where the SRC purposefully streams media to an SRS and all participating user agents (UAs) are notified of the recording. Passive recording, where a recording device detects media directly from the network (e.g., using port-mirroring techniques), is outside the scope of this document. In addition, lawful intercept is outside the scope of this document.

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/rfc7866.

Table of Contents

1. Introduction

This document specifies the mechanism to record a Communication Session (CS) by delivering real-time media and metadata from the CS to a recording device. In accordance with the architecture [RFC7245], the Session Recording Protocol specifies the use of SIP, the Session Description Protocol (SDP), and RTP to establish a Recording Session (RS) between the Session Recording Client (SRC), which is on the path of the CS, and a Session Recording Server (SRS) at the recording device. SIP is also used to deliver metadata to the recording device, as specified in [RFC7865]. Metadata is information that describes recorded media and the CS to which they relate. The Session Recording Protocol intends to satisfy the SIP-based Media Recording (SIPREC) requirements listed in [RFC6341]. In addition to the Session Recording Protocol, this document specifies extensions for user agents (UAs) that are participants in a CS to receive recording indications and to provide preferences for recording.

This document considers only active recording, where the SRC purposefully streams media to an SRS and all participating UAs are notified of the recording. Passive recording, where a recording device detects media directly from the network (e.g., using port-mirroring techniques), is outside the scope of this document. In addition, lawful intercept is outside the scope of this document, in accordance with [RFC2804].

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. Definitions

This document refers to the core definitions provided in the architecture document [RFC7245].

Section 8 uses the definitions provided in "RTP: A Transport Protocol for Real-Time Applications" [RFC3550].

4. Scope

The scope of the Session Recording Protocol includes the establishment of the RSs and the reporting of the metadata. The scope also includes extensions supported by UAs participating in the CS, such as an indication of recording. The UAs need not be recording aware in order to participate in a CS being recorded.

The items in the following list, which is not exhaustive, do not represent the protocol itself and are considered out of scope for the Session Recording Protocol:

- Delivering recorded media in real time as the CS media
- Specifications of criteria to select a specific CS to be recorded or triggers to record a certain CS in the future
- Recording policies that determine whether the CS should be recorded and whether parts of the CS are to be recorded
- Retention policies that determine how long a recording is stored
- Searching and accessing the recorded media and metadata
- Policies governing how CS users are made aware of recording
- Delivering additional RS metadata through a non-SIP mechanism

5. Overview of Operations

This section is informative and provides a description of recording operations.

Section 6 describes the SIP communication in an RS between an SRC and an SRS, as well as the procedures for recording-aware UAs participating in a CS. Section 7 describes SDP handling in an RS, and the procedures for recording indications and recording preferences. Section 8 describes RTP handling in an RS. Section 9 describes the mechanism to deliver recording metadata from the SRC to the SRS.

As mentioned in the architecture document [RFC7245], there are a number of types of call flows based on the location of the SRC. The sample call flows discussed in Section 5.1 provide a quick overview of the operations between the SRC and the SRS.

5.1. Delivering Recorded Media

When a SIP Back-to-Back User Agent (B2BUA) with SRC functionality routes a call from UA A to UA B, the SRC has access to the media path between the UAs. When the SRC is aware that it should be recording the conversation, the SRC can cause the B2BUA to relay the media between UA A and UA B. The SRC then establishes the RS with the SRS and sends replicated media towards the SRS.

An endpoint may also have SRC functionality, where the endpoint itself establishes the RS to the SRS. Since the endpoint has access to the media in the CS, the endpoint can send replicated media towards the SRS.

The example call flows in Figures 1 and 2 show an SRC establishing an RS towards an SRS. Figure 1 illustrates UA A acting as the SRC. Figure 2 illustrates a B2BUA acting as the SRC. Note that the SRC can choose when to establish the RS independent of the CS, even though the example call flows suggest that the SRC is establishing the RS (message (5) in Figure 2) after the CS is established.

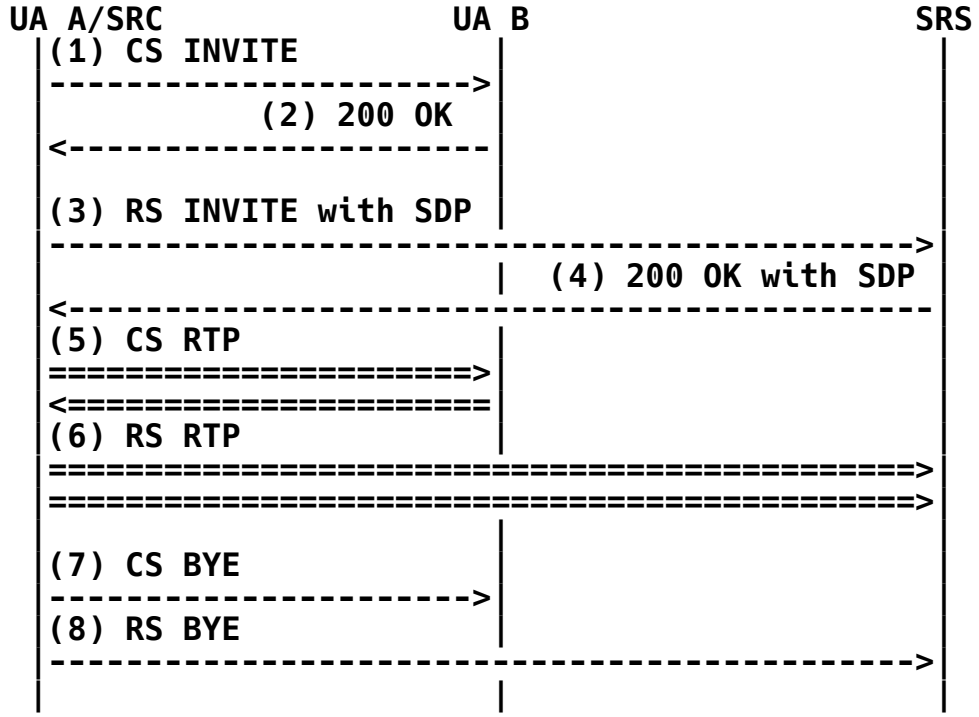

Figure 1: Basic Recording Call Flow with UA as SRC

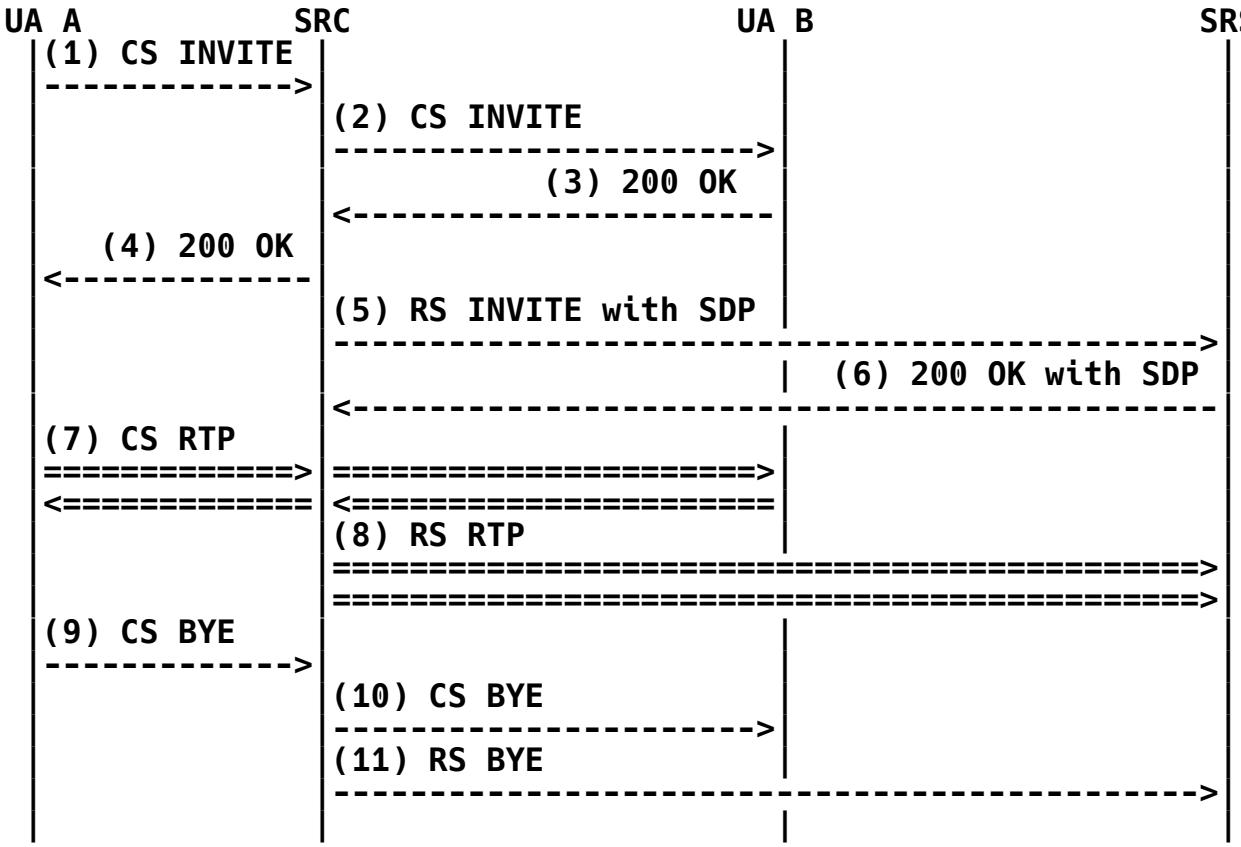

Figure 2: Basic Recording Call Flow with B2BUA as SRC

The call flow shown in Figure 2 can also apply to the case of a centralized conference with a mixer. For clarity, ACKs to INVITEs and 200 OKs to BYEs are not shown. The conference focus can provide the SRC functionality, since the conference focus has access to all the media from each conference participant. When a recording is requested, the SRC delivers the metadata and the media streams to the SRS. Since the conference focus has access to a mixer, the SRC may choose to mix the media streams from all participants as a single mixed media stream towards the SRS.

An SRC can use a single RS to record multiple CSs. Every time the SRC wants to record a new call, the SRC updates the RS with a new SDP offer to add new recorded streams to the RS and to correspondingly also update the metadata for the new call.

An SRS can also establish an RS to an SRC, although it is beyond the scope of this document to define how an SRS would specify which calls to record.

5.2. Delivering Recording Metadata

The SRC is responsible for the delivery of metadata to the SRS. The SRC may provide an initial metadata snapshot about recorded media streams in the initial INVITE content in the RS. Subsequent metadata updates can be represented as a stream of events in UPDATE [RFC3311] or re-INVITE requests sent by the SRC. These metadata updates are normally incremental updates to the initial metadata snapshot to optimize on the size of updates. However, the SRC may also decide to send a new metadata snapshot at any time.

Metadata is transported in the body of INVITE or UPDATE messages. Certain metadata, such as the attributes of the recorded media stream, is located in the SDP of the RS.

The SRS has the ability to send a request to the SRC to ask for a new metadata snapshot update from the SRC. This can happen when the SRS fails to understand the current stream of incremental updates for whatever reason -- for example, when the SRS loses the current state due to internal failure. The SRS may optionally attach a reason along with the snapshot request. This request allows both the SRC and the SRS to synchronize the states with a new metadata snapshot so that further incremental metadata updates will be based on the latest metadata snapshot. Similar to the metadata content, the metadata snapshot request is transported as content in UPDATE or INVITE messages sent by the SRS in the RS.

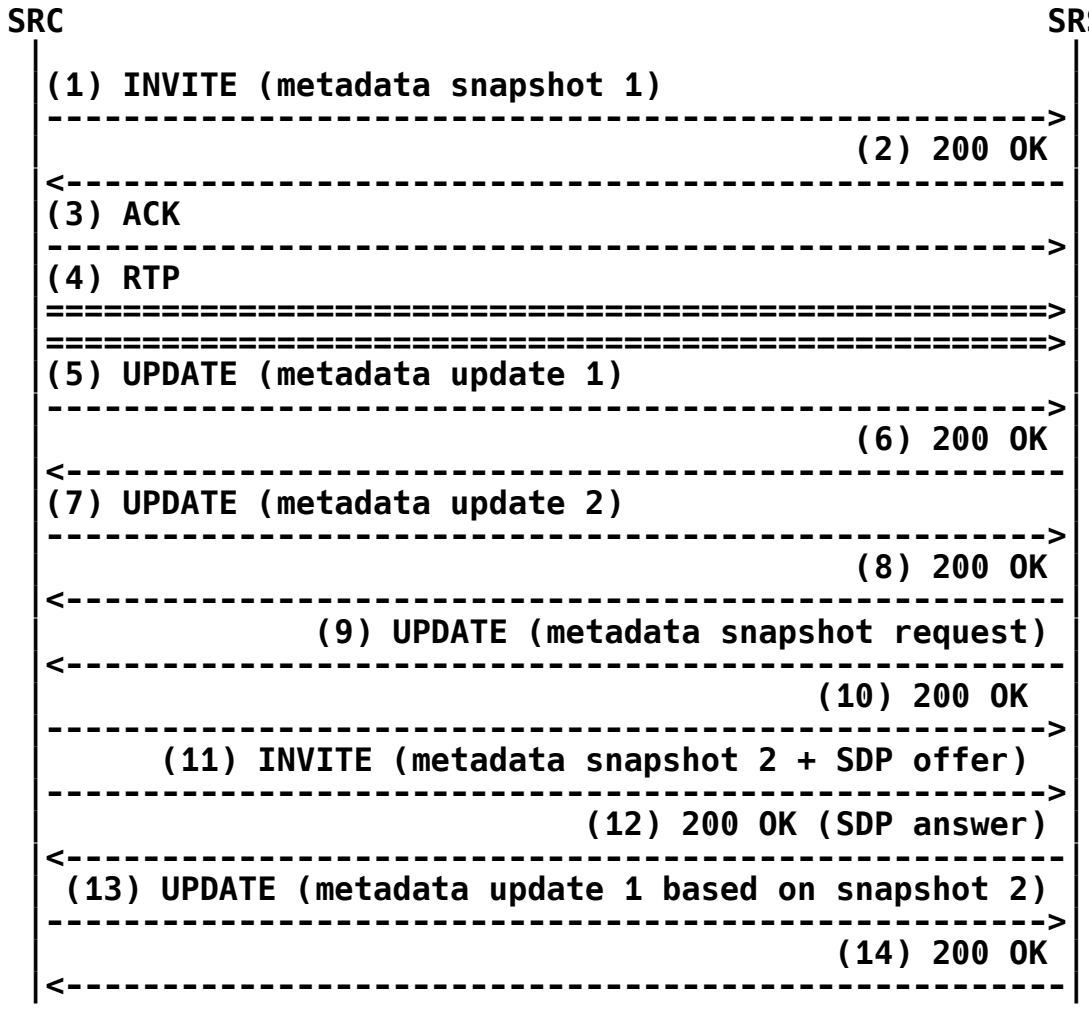

Figure 3: Delivering Metadata via SIP UPDATE

5.3. Receiving Recording Indications and Providing Recording Preferences

The SRC is responsible for providing recording indications to the participants in the CS. A recording-aware UA supports receiving recording indications via the SDP "a=record" attribute, and it can specify a recording preference in the CS by including the SDP "a=recordpref" attribute. The recording attribute is a declaration by the SRC in the CS to indicate whether recording is taking place. The recording preference attribute is a declaration by the recording-aware UA in the CS to indicate its recording preference. A UA that does not want to be recorded may still be notified that recording is occurring, for a number of reasons (e.g., it was not capable of

indicating its preference, its preference was ignored). If this occurs, the UA's only mechanism to avoid being recorded is to terminate its participation in the session.

To illustrate how the attributes are used, if UA A is initiating a call to UA B and UA A is also an SRC that is performing the recording, then UA A provides the recording indication in the SDP offer with a=record:on. Since UA A is the SRC, UA A receives the recording indication from the SRC directly. When UA B receives the SDP offer, UA B will see that recording is happening on the other endpoint of this session. Since UA B is not an SRC and does not provide any recording preference, the SDP answer does not contain a=record or a=recordpref.

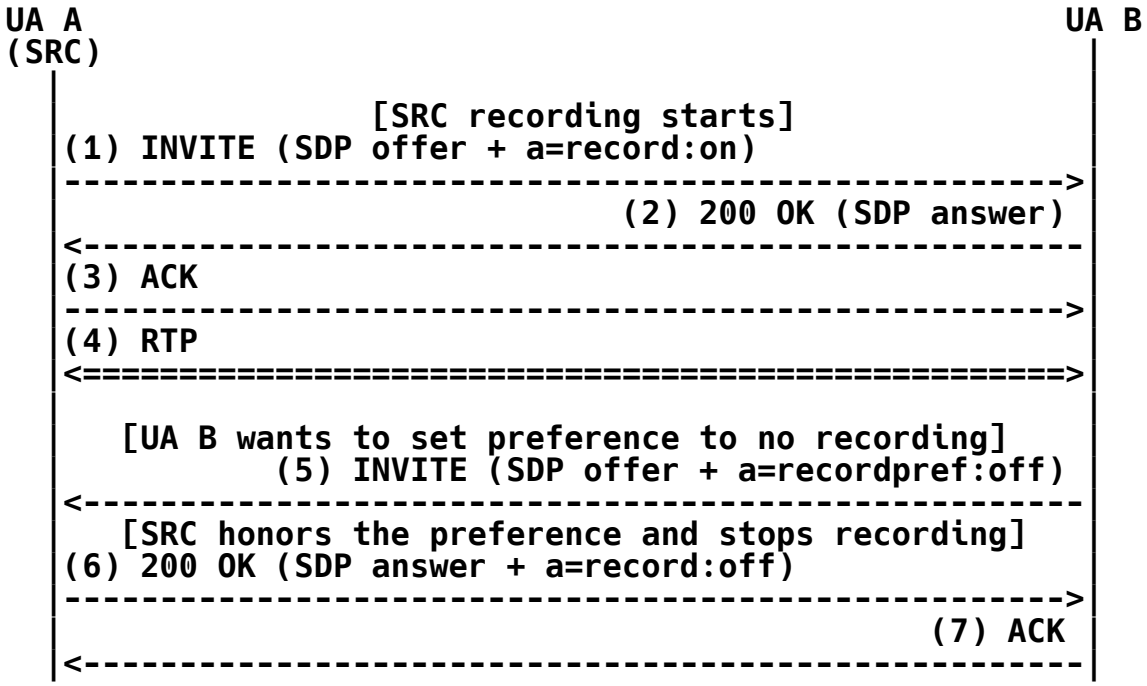

Figure 4: Recording Indication and Recording Preference

After the call is established and recording is in progress, UA B later decides to change the recording preference to no recording and sends a re-INVITE with the "a=recordpref" attribute. It is up to the SRC to honor the preference, and in this case the SRC decides to stop the recording and updates the recording indication in the SDP answer.

Note that UA B could have explicitly indicated a recording preference in (2), the 200 OK for the original INVITE. Indicating a preference of no recording in an initial INVITE or an initial response to an INVITE may reduce the chance of a user being recorded in the first place.

6. SIP Handling

6.1. Procedures at the SRC

6.1.1. Initiating a Recording Session

An RS is a SIP session with specific extensions applied, and these extensions are listed in the procedures below for the SRC and the SRS. When an SRC or an SRS receives a SIP session that is not an RS, it is up to the SRC or the SRS to determine what to do with the SIP session.

The SRC can initiate an RS by sending a SIP INVITE request to the SRS. The SRC and the SRS are identified in the From and To headers, respectively.

The SRC MUST include the "+sip.src" feature tag in the Contact URI, defined in this specification as an extension to [RFC3840], for all RSs. An SRS uses the presence of the "+sip.src" feature tag in dialog creating and modifying requests and responses to confirm that the dialog being created is for the purpose of an RS. In addition, when an SRC sends a REGISTER request to a registrar, the SRC MAY include the "+sip.src" feature tag to indicate that it is an SRC.

Since SIP Caller Preferences extensions are optional to implement for routing proxies, there is no guarantee that an RS will be routed to an SRC or SRS. A new option tag, "siprec", is introduced. As per [RFC3261], only an SRC or an SRS can accept this option tag in an RS. An SRC MUST include the "siprec" option tag in the Require header when initiating an RS so that UAs that do not support the Session Recording Protocol extensions will simply reject the INVITE request with a 420 (Bad Extension) response.

When an SRC receives a new INVITE, the SRC MUST only consider the SIP session as an RS when both the "+sip.srs" feature tag and the "siprec" option tag are included in the INVITE request.

6.1.2. SIP Extensions for Recording Indications and Preferences

For the CS, the SRC MUST provide recording indications to all participants in the CS. A participant UA in a CS can indicate that it is recording aware by providing the "record-aware" option tag, and the SRC MUST provide recording indications in the new SDP "a=record" attribute described in Section 7 below. In the absence of the "record-aware" option tag -- meaning that the participant UA is not recording aware -- an SRC MUST provide recording indications through other means, such as playing a tone in-band or having a signed participant contract in place.

An SRC in the CS may also indicate itself as a session recording client by including the "+sip.src" feature tag. A recording-aware participant can learn that an SRC is in the CS and can set the recording preference for the CS with the new SDP "a=recordpref" attribute described in Section 7.

6.2. Procedures at the SRS

When an SRS receives a new INVITE, the SRS MUST only consider the SIP session as an RS when both the "+sip.src" feature tag and the "siprec" option tag are included in the INVITE request.

The SRS can initiate an RS by sending a SIP INVITE request to the SRC. The SRS and the SRC are identified in the From and To headers, respectively.

The SRS MUST include the "+sip.srs" feature tag in the Contact URI, as per [RFC3840], for all RSs. An SRC uses the presence of this feature tag in dialog creation and modification requests and responses to confirm that the dialog being created is for the purpose of an RS (REQ-030 in [RFC6341]). In addition, when an SRS sends a REGISTER request to a registrar, the SRS SHOULD include the "+sip.srs" feature tag to indicate that it is an SRS.

An SRS MUST include the "siprec" option tag in the Require header as per [RFC3261] when initiating an RS so that UAs that do not support the Session Recording Protocol extensions will simply reject the INVITE request with a 420 (Bad Extension) response.

6.3. Procedures for Recording-Aware User Agents

A recording-aware UA is a participant in the CS that supports the SIP and SDP extensions for receiving recording indications and for requesting recording preferences for the call. A recording-aware UA MUST indicate that it can accept the reporting of recording indications provided by the SRC with a new "record-aware" option tag

when initiating or establishing a CS; this means including the "record-aware" option tag in the Supported header in the initial INVITE request or response.

A recording-aware UA MUST provide a recording indication to the end user through an appropriate user interface, indicating whether recording is on, off, or paused for each medium. Appropriate user interfaces may include real-time notification or previously established agreements that use of the device is subject to recording. Some UAs that are automatons (e.g., Interactive Voice Response (IVR), media server, Public Switched Telephone Network (PSTN) gateway) may not have a user interface to render a recording indication. When such a UA indicates recording awareness, the UA SHOULD render the recording indication through other means, such as passing an in-band tone on the PSTN gateway, putting the recording indication in a log file, or raising an application event in a VoiceXML dialog. These UAs MAY also choose not to indicate recording awareness, thereby relying on whatever mechanism an SRC chooses to indicate recording, such as playing a tone in-band.

7. SDP Handling

7.1. Procedures at the SRC

The SRC and SRS follow the SDP offer/answer model described in [RFC3264]. The procedures for the SRC and SRS describe the conventions used in an RS.

7.1.1. SDP Handling in the RS

Since the SRC does not expect to receive media from the SRS, the SRC typically sets each media stream of the SDP offer to only send media, by qualifying them with the "a=sendonly" attribute, according to the procedures in [RFC3264].

The SRC sends recorded streams of participants to the SRS, and the SRC MUST provide a "label" attribute ("a=label"), as per [RFC4574], on each media stream in order to identify the recorded stream with the rest of the metadata. The "a=label" attribute identifies each recorded media stream, and the label name is mapped to the Media Stream Reference in the metadata as per [RFC7865]. The scope of the "a=label" attribute only applies to the SDP and metadata conveyed in the bodies of the SIP request or response that the label appeared in. Note that a recorded stream is distinct from a CS stream; the metadata provides a list of participants that contribute to each recorded stream.

Figure 5 shows an example SDP offer from an SRC with both audio and video recorded streams. Note that this example contains unfolded lines longer than 72 characters; these lines are captured between <allOneLine> tags.

```
v=0
o=SRC 2890844526 2890844526 IN IP4 198.51.100.1
s=-
c=IN IP4 198.51.100.1
t=0 0
m=audio 12240 RTP/AVP 0 4 8
a=sendonly
a=label:1
m=video 22456 RTP/AVP 98
a=rtpmap:98 H264/90000
<allOneLine>
a=fmtp:98 profile-level-id=42A01E;
          sprop-parameter-sets=Z0IACpZTBYmI,aMljiA==
</allOneLine>
a=sendonly
a=label:2
m=audio 12242 RTP/AVP 0 4 8
a=sendonly
a=label:3
m=video 22458 RTP/AVP 98
a=rtpmap:98 H264/90000
<allOneLine>
a=fmtp:98 profile-level-id=42A01E;
          sprop-parameter-sets=Z0IACpZTBYmI,aMljiA==
</allOneLine>
a=sendonly
a=label:4
```

Figure 5: Sample SDP Offer from SRC with Audio and Video Streams

7.1.1.1. Handling Media Stream Updates

Over the lifetime of an RS, the SRC can add and remove recorded streams to and from the RS for various reasons -- for example, when a CS stream is added to or removed from the CS, or when a CS is created or terminated if an RS handles multiple CSs. To remove a recorded stream from the RS, the SRC sends a new SDP offer where the port of the media stream to be removed is set to zero, according to the procedures in [RFC3264]. To add a recorded stream to the RS, the SRC sends a new SDP offer by adding a new media stream description or by reusing an old media stream that had been previously disabled, according to the procedures in [RFC3264].

The SRC can temporarily discontinue streaming and collection of recorded media from the SRC to the SRS for reasons such as masking the recording. In this case, the SRC sends a new SDP offer and sets the media stream to inactive (a=inactive) for each recorded stream to be paused, as per the procedures in [RFC3264]. To resume streaming and collection of recorded media, the SRC sends a new SDP offer and sets the media stream to sendonly (a=sendonly). Note that a CS may itself change the media stream direction by updating the SDP -- for example, by setting a=inactive for SDP hold. Media stream direction changes in the CS are conveyed in the metadata by the SRC. When a CS media stream is changed to or from inactive, the effect on the corresponding RS media stream is governed by SRC policy. The SRC MAY have a local policy to pause an RS media stream when the corresponding CS media stream is inactive, or it MAY leave the RS media stream as sendonly.

7.1.2. Recording Indication in the CS

While there are existing mechanisms for providing an indication that a CS is being recorded, these mechanisms are usually delivered on the CS media streams, such as playing an in-band tone or an announcement to the participants. A new "record" SDP attribute is introduced to allow the SRC to indicate recording state to a recording-aware UA in a CS.

The "record" SDP attribute appears at the media level or session level in either an SDP offer or answer. When the attribute is applied at the session level, the indication applies to all media streams in the SDP. When the attribute is applied at the media level, the indication applies to that one media stream only, and that overrides the indication if also set at the session level. Whenever the recording indication needs to change, such as termination of recording, the SRC MUST initiate a re-INVITE or UPDATE to update the SDP "a=record" attribute.

The following is the ABNF [RFC5234] of the "record" attribute:

```
attribute =/ record-attr
; attribute defined in RFC 4566

record-attr = "record:" indication
indication = "on" / "off" / "paused"
```

on: Recording is in progress.

off: No recording is in progress.

paused: Recording is in progress but media is paused.

7.1.3. Recording Preference in the CS

When the SRC receives the "a=recordpref" SDP in an SDP offer or answer, the SRC chooses to honor the preference to record based on local policy at the SRC. If the SRC makes a change in recording state, the SRC MUST report the new recording state in the "a=record" attribute in the SDP answer or in a subsequent SDP offer.

7.2. Procedures at the SRS

Typically, the SRS only receives RTP streams from the SRC; therefore, the SDP offer/answer from the SRS normally sets each media stream to receive media, by setting them with the "a=recvonly" attribute, according to the procedures of [RFC3264]. When the SRS is not ready to receive a recorded stream, the SRS sets the media stream as inactive in the SDP offer or answer by setting it with an "a=inactive" attribute, according to the procedures of [RFC3264]. When the SRS is ready to receive recorded streams, the SRS sends a new SDP offer and sets the media streams with an "a=recvonly" attribute.

Figure 6 shows an example of an SDP answer from the SRS for the SDP offer from Figure 5. Note that this example contains unfolded lines longer than 72 characters; these lines are captured between <allOneLine> tags.

```
v=0
o=SRS 0 0 IN IP4 198.51.100.20
s=-
c=IN IP4 198.51.100.20
t=0 0
m=audio 10000 RTP/AVP 0
a=recvonly
a=label:1
m=video 10002 RTP/AVP 98
a=rtpmap:98 H264/90000
<allOneLine>
a=fmtp:98 profile-level-id=42A01E;
          sprop-parameter-sets=Z0IACpZTBYmI,aMljiA==
</allOneLine>
a=recvonly
a=label:2
m=audio 10004 RTP/AVP 0
a=recvonly
a=label:3
m=video 10006 RTP/AVP 98
a=rtpmap:98 H264/90000
<allOneLine>
a=fmtp:98 profile-level-id=42A01E;
          sprop-parameter-sets=Z0IACpZTBYmI,aMljiA==
</allOneLine>
a=recvonly
a=label:4
```

Figure 6: Sample SDP Answer from SRS with Audio and Video Streams

Over the lifetime of an RS, the SRS can remove recorded streams from the RS for various reasons. To remove a recorded stream from the RS, the SRS sends a new SDP offer where the port of the media stream to be removed is set to zero, according to the procedures in [RFC3264].

The SRS MUST NOT add recorded streams in the RS when the SRS sends a new SDP offer. Similarly, when the SRS starts an RS, the SRS MUST initiate the INVITE without an SDP offer to let the SRC generate the SDP offer with the streams to be recorded.

The sequence diagram in Figure 7 shows an example where the SRS is initially not ready to receive recorded streams and later updates the RS when the SRS is ready to record.

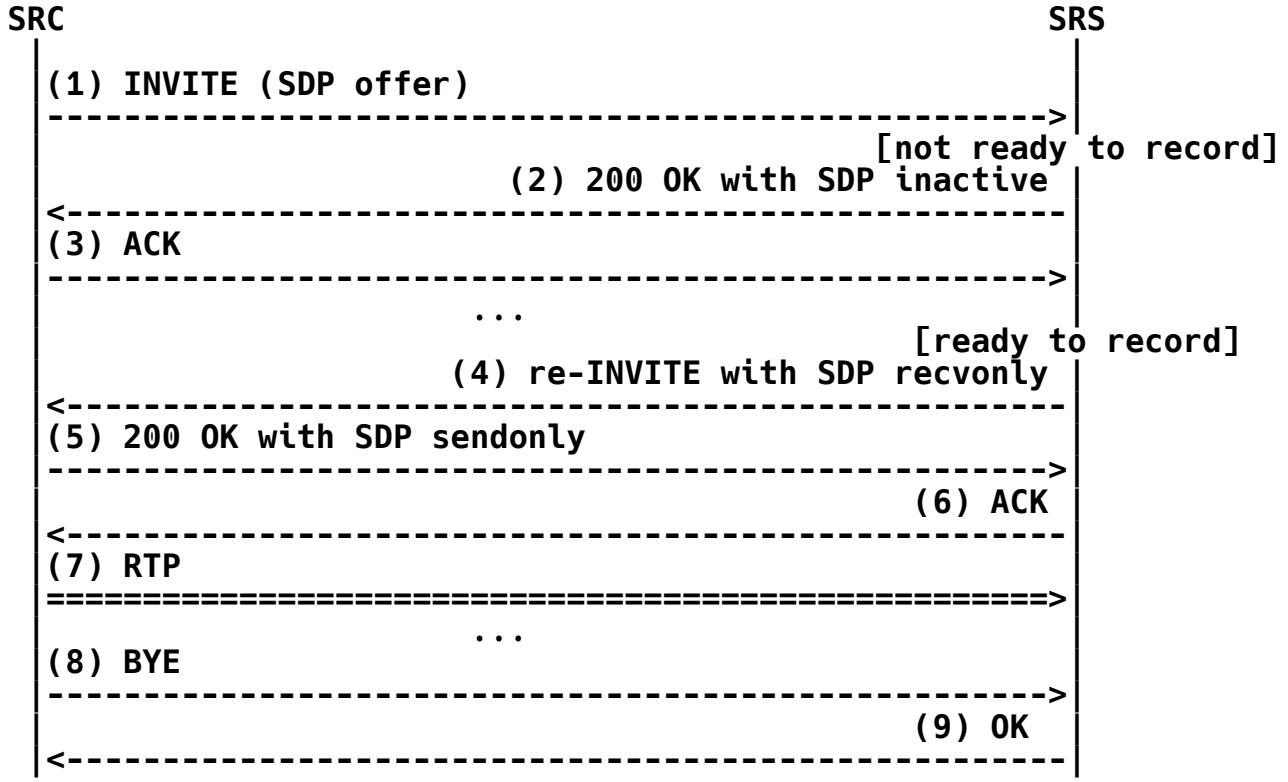

Figure 7: SRS Responding to Offer with a=inactive

7.3. Procedures for Recording-Aware User Agents

7.3.1. Recording Indication

When a recording-aware UA receives an SDP offer or answer that includes the "a=record" attribute, the UA provides to the end user an indication as to whether the recording is on, off, or paused for each medium, based on the most recently received "a=record" SDP attribute for that medium.

When a CS is traversed through multiple UAs such as a B2BUA or a conference focus, each UA involved in the CS that is aware that the CS is being recorded MUST provide the recording indication through the "a=record" attribute to all other parties in the CS.

It is possible that more than one SRC is in the call path of the same CS, but the recording indication attribute does not provide any hint as to which SRC or how many SRCs are recording. An endpoint knows only that the call is being recorded. Furthermore, this attribute is not used as a request for a specific SRC to start or stop recording.

7.3.2. Recording Preference

A participant in a CS MAY set the recording preference in the CS to be recorded or not recorded at session establishment or during the session. A new "recordpref" SDP attribute is introduced, and the participant in the CS may set this recording preference attribute in any SDP offer/answer at session establishment time or during the session. The SRC is not required to honor the recording preference from a participant, based on local policies at the SRC, and the participant can learn the recording indication through the "a=record" SDP attribute as described in Section 7.3.1.

The SDP "a=recordpref" attribute can appear at the media level or session level and can appear in an SDP offer or answer. When the attribute is applied at the session level, the recording preference applies to all media streams in the SDP. When the attribute is applied at the media level, the recording preference applies to that one media stream only, and that overrides the recording preference if also set at the session level. The UA can change the recording preference by changing the "a=recordpref" attribute in a subsequent SDP offer or answer. The absence of the "a=recordpref" attribute in the SDP indicates that the UA has no recording preference.

The following is the ABNF of the "recordpref" attribute:

```
attribute =/ recordpref-attr
; attribute defined in RFC 4566

recordpref-attr = "a=recordpref:" pref
pref = "on" / "off" / "pause" / "nopreference"
```

on: Sets the preference to record if it has not already been started. If the recording is currently paused, the preference is to resume recording.

off: Sets the preference for no recording. If recording has already been started, then the preference is to stop the recording.

pause: If the recording is currently in progress, sets the preference to pause the recording.

nopreference:
Indicates that the UA has no preference regarding recording.

8. RTP Handling

This section provides recommendations and guidelines for RTP and the Real-time Transport Control Protocol (RTCP) in the context of SIPREC [RFC6341]. In order to communicate most effectively, the SRC, the SRS, and any recording-aware UAs should utilize the mechanisms provided by RTP in a well-defined and predictable manner. It is the goal of this document to make the reader aware of these mechanisms and to provide recommendations and guidelines.

8.1. RTP Mechanisms

This section briefly describes important RTP/RTCP constructs and mechanisms that are particularly useful within the context of SIPREC.

8.1.1. 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, per [RFC3550], 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 from the receivers may be used to diagnose faults in the distribution. As such, RTCP is a well-defined and efficient mechanism for the SRS to inform the SRC, and for the SRC to inform recording-aware UAs, of issues that arise with respect to the reception of media that is to be recorded.

2) Including a persistent transport-level identifier -- the CNAME, or canonical name -- for an RTP source

   The synchronization source (SSRC) [RFC3550] 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.

8.1.2. RTP Profile

The RECOMMENDED RTP profiles for the SRC, SRS, and recording-aware UAs 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 these are not requirements, some implementations may use "The Secure Real-time Transport Protocol (SRTP)" [RFC3711] and "RTP Profile for Audio and Video Conferences with Minimal Control" [RFC3551]. Therefore, it is RECOMMENDED that the SRC, SRS, and recording-aware UAs not rely entirely on RTP/SAVPF or RTP/AVPF for core functionality that may be at least partially achievable using RTP/SAVP and RTP/AVP.

AVPF and SAVPF provide an improved RTCP timer model that allows more flexible transmission of RTCP packets in response to events, rather than strictly according to bandwidth. AVPF-based codec control messages provide efficient mechanisms for an SRC, an SRS, and recording-aware UAs 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.

8.1.3. SSRC

The 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, SRS, and recording-aware UAs handle changes in SSRC values and properly identify the reason for the change. The CNAME values carried in RTCP facilitate this identification.

8.1.4. 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 an SRC or recording-aware UA, when acting as a mixer, set the CSRC list accordingly, and that the SRC and SRS interpret the CSRC list per [RFC3550] when received.

8.1.5. SDES

The Source Description (SDES), as defined in [RFC3550], contains an SSRC/CSRC identifier followed by a list of zero or more items that 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 CSRC from which it is receiving SDES information, or multiple complete SDES packets if there are more than 31 such sources.

The ability to identify individual CSRCs is important in the context of SIPREC. Metadata [RFC7865] provides a mechanism to achieve this at the signaling level. SDES provides a mechanism at the RTP level.

8.1.5.1. CNAME

The Canonical End-Point Identifier (CNAME), as defined in [RFC3550], provides the binding from the SSRC identifier to an identifier for the source (sender or receiver) that remains constant. It is important that the SRC and recording-aware UAs generate CNAMEs appropriately and that the SRC and SRS interpret and use them for this purpose. Guidelines for generating CNAME values are provided in "Guidelines for Choosing RTP Control Protocol (RTCP) Canonical Names (CNAMEs)" [RFC7022].

8.1.6. Keepalive

It is anticipated that media streams in SIPREC may exist in an inactive state for extended periods of time for any of a number of valid reasons. In order for the bindings and any pinholes in NATs/firewalls to remain active during such intervals, it is RECOMMENDED that the SRC, SRS, and recording-aware UAs follow the keepalive procedure recommended in "Application Mechanism for Keeping Alive the NAT Mappings Associated with RTP / RTP Control Protocol (RTCP) Flows" [RFC6263] for all RTP media streams.

8.1.7. RTCP Feedback Messages

"Codec Control Messages in the RTP Audio-Visual Profile with Feedback (AVPF)" [RFC5104] specifies extensions to the messages defined in AVPF [RFC4585]. Support for and proper usage of these messages are important to SRC, SRS, and recording-aware UA implementations. Note that these messages are applicable only when using the AVPF or SAVPF RTP profiles.

8.1.7.1. Full Intra Request

A Full Intra Request (FIR) command, when received by the designated media sender, requires that the media sender send 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 Instantaneous Decoding Refresh (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.

8.1.7.1.1. Deprecated Usage of SIP INFO Instead of 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 in [RFC5168], except for purposes of backward compatibility. Implementations SHOULD use FIR messages instead.

To make sure that a common mechanism exists between the SRC and SRS, the SRS MUST support both mechanisms (FIR and SIP INFO), using FIR messages when negotiated successfully with the SRC and using SIP INFO otherwise.

8.1.7.2. Picture Loss Indication

Picture Loss Indication (PLI), as defined in [RFC4585], informs the encoder of the loss of an undefined amount of coded video data belonging to one or more pictures. [RFC4585] recommends using PLI instead of FIR messages to recover from errors. FIR is appropriate only in situations where not sending a decoder refresh point would render the video unusable for the users. Examples where sending FIR messages is appropriate include a multipoint conference when a new

user joins the conference and no regular decoder refresh point interval is established, and a video-switching Multipoint Control Unit (MCU) that changes streams.

Appropriate use of PLI and FIR is important to ensure, with minimum overhead, that the recorded video is usable (e.g., the necessary reference frames exist for a player to render the recorded video).

8.1.7.3. Temporary Maximum Media Stream Bit Rate Request

A receiver, translator, or mixer uses the Temporary Maximum Media Stream Bit Rate Request (TMMBR) [RFC5104] 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.

8.1.7.3.1. 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 for more permanent changes.

8.1.8. Symmetric RTP/RTCP for Sending and Receiving

Within an SDP offer/answer exchange, RTP entities choose the RTP and RTCP transport addresses (i.e., IP addresses and port numbers) on which to receive packets. When sending packets, the RTP entities may use the same source port or a different source port than those signaled for receiving packets. When the transport address used to send and receive RTP is the same, it is termed "symmetric RTP" [RFC4961]. Likewise, when the transport address used to send and receive RTCP is the same, it is termed "symmetric RTCP" [RFC4961].

When sending RTP, the use of symmetric RTP is REQUIRED. When sending RTCP, the use of symmetric RTCP is REQUIRED. Although an SRS will not normally send RTP, it will send RTCP as well as receive RTP and RTCP. Likewise, although an SRC will not normally receive RTP from the SRS, it will receive RTCP as well as send RTP and RTCP.

> Note: Symmetric RTP and symmetric RTCP are different from RTP/RTCP multiplexing [RFC5761].

8.2. Roles

An SRC has the task of gathering media from the various UAs in one or more CSs and forwarding the information to the SRS within the context of a corresponding RS. There are numerous ways in which an SRC may do this, including, but not limited to, appearing as a UA within a CS, or as a B2BUA between UAs within a CS.

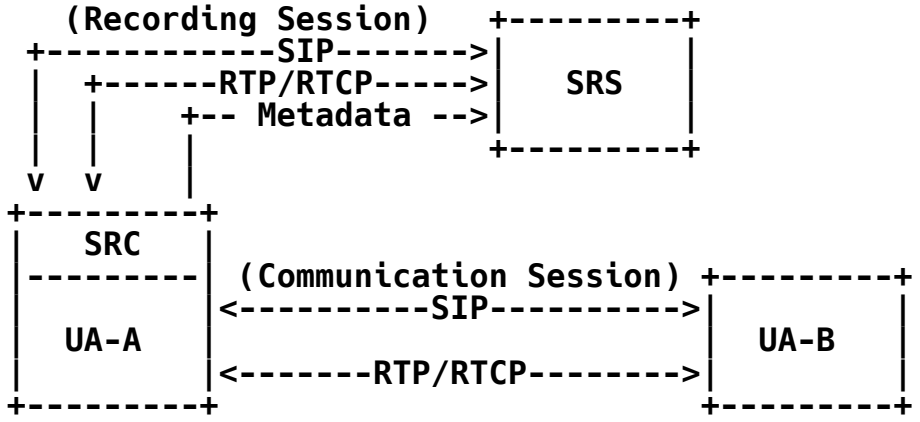

Figure 8: UA as SRC

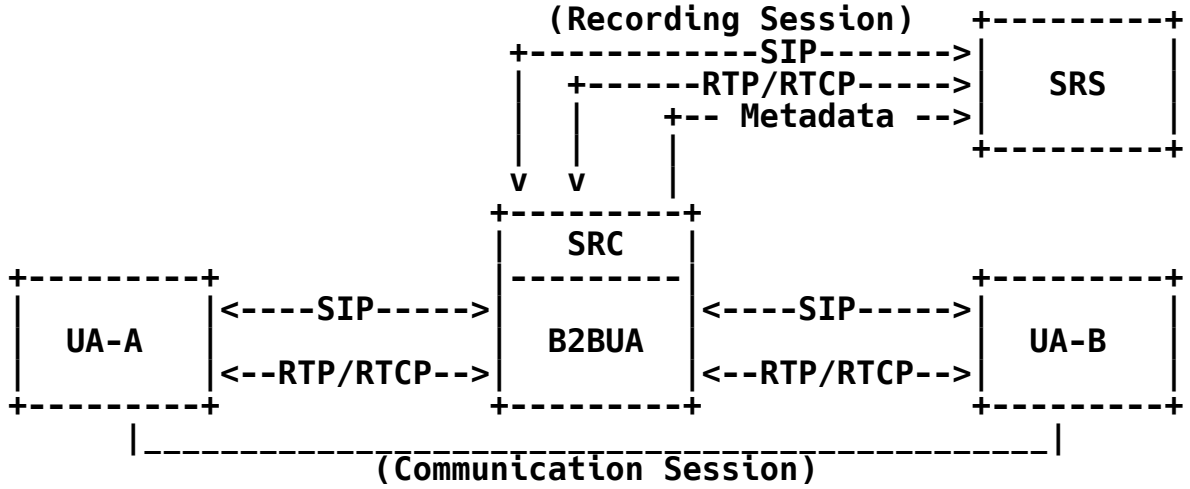

Figure 9: B2BUA as SRC

The following subsections define a set of roles an SRC may choose to play, based on its position with respect to a UA within a CS, and an SRS within an RS. A CS and a corresponding RS are independent sessions; therefore, an SRC may play a different role within a CS than it does within the corresponding RS.

8.2.1. SRC Acting as an RTP Translator

The SRC may act as a translator, as defined in [RFC3550]. A defining characteristic of a translator is that it forwards RTP packets with their SSRC identifier intact. There are two types of translators: one that simply forwards, and another that performs transcoding (e.g., from one codec to another) in addition to forwarding.

8.2.1.1. Forwarding Translator

When acting as a forwarding translator, RTP received as separate streams from different sources (e.g., from different UAs with different SSRCs) cannot be mixed by the SRC and MUST be sent separately to the SRS. All RTCP reports MUST be passed by the SRC between the UAs and the SRS, such that the UAs and SRS are able to detect any SSRC collisions.

RTCP Sender Reports generated by a UA sending a stream MUST be forwarded to the SRS. RTCP Receiver Reports generated by the SRS MUST be forwarded to the relevant UA.

UAs may receive multiple sets of RTCP Receiver Reports -- one or more from other UAs participating in the CS, and one from the SRS participating in the RS. A UA SHOULD process the RTCP Receiver Reports from the SRS if it is recording aware.

If SRTP is used on both the CS and the RS, decryption and/or re-encryption may occur. For example, if different keys are used, it will occur. If the same keys are used, it need not occur. Section 12 provides additional information on SRTP and keying mechanisms.

If packet loss occurs, either from the UA to the SRC or from the SRC to the SRS, the SRS SHOULD detect and attempt to recover from the loss. The SRC does not play a role in this, other than forwarding the associated RTP and RTCP packets.

8.2.1.2. Transcoding Translator

When acting as a transcoding translator, an SRC MAY perform transcoding (e.g., from one codec to another), and this may result in a different rate of packets between what the SRC receives on the CS and what the SRC sends on the RS. As when acting as a forwarding translator, RTP received as separate streams from different sources (e.g., from different UAs with different SSRCs) cannot be mixed by the SRC and MUST be sent separately to the SRS. All RTCP reports MUST be passed by the SRC between the UAs and the SRS, such that the UAs and SRS are able to detect any SSRC collisions.

RTCP Sender Reports generated by a UA sending a stream MUST be forwarded to the SRS. RTCP Receiver Reports generated by the SRS MUST be forwarded to the relevant UA. The SRC may need to manipulate the RTCP Receiver Reports to take into account any transcoding that has taken place.

UAs may receive multiple sets of RTCP Receiver Reports -- one or more from other UAs participating in the CS, and one from the SRS participating in the RS. A recording-aware UA SHOULD be prepared to process the RTCP Receiver Reports from the SRS, whereas a recording-unaware UA may discard such RTCP packets as irrelevant.

If SRTP is used on both the CS and the RS, decryption and/or re-encryption may occur. For example, if different keys are used, it will occur. If the same keys are used, it need not occur. Section 12 provides additional information on SRTP and keying mechanisms.

If packet loss occurs, either from the UA to the SRC or from the SRC to the SRS, the SRS SHOULD detect and attempt to recover from the loss. The SRC does not play a role in this, other than forwarding the associated RTP and RTCP packets.

8.2.2. SRC Acting as an RTP Mixer

In the case of the SRC acting as an RTP mixer, as defined in [RFC3550], the SRC combines RTP streams from different UAs and sends them towards the SRS using its own SSRC. The SSRCs from the contributing UA SHOULD be conveyed as CSRC identifiers within this stream. The SRC may 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 UAs. RTCP Sender Reports and Receiver Reports are not forwarded by an SRC acting as a mixer, but there are requirements for forwarding RTCP Source Description (SDES) packets. The SRC generates its own RTCP Sender Reports and Receiver Reports toward the associated UAs and SRS.

The use of SRTP between the SRC and the SRS for the RS is independent of the use of SRTP between the UAs and the SRC for the CS. Section 12 provides additional information on SRTP and keying mechanisms.

If packet loss occurs from the UA to the SRC, the SRC SHOULD detect and attempt to recover from the loss. If packet loss occurs from the SRC to the SRS, the SRS SHOULD detect and attempt to recover from the loss.

8.2.3. 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 for the RS is independent of the use of SRTP between the UAs and SRC for the CS. Section 12 provides additional information on SRTP and keying mechanisms.

If packet loss occurs from the UA to the SRC, the SRC SHOULD detect and attempt to recover from the loss. If packet loss occurs from the SRC to the SRS, the SRS SHOULD detect and attempt to recover from the loss.

8.3. RTP Session Usage by SRC

There are multiple ways that an SRC may choose to deliver recorded media to an SRS. In some cases, it may use a single RTP session for all media within the RS, whereas in others it may use multiple RTP sessions. The following subsections provide examples of basic RTP session usage by the SRC, including a discussion of how the RTP constructs and mechanisms covered previously are used. An SRC may choose to use one or more of the RTP session usages within a single RS. For the purpose of base interoperability between SRC and SRS, an SRC MUST support separate m-lines in SDP, one per CS media direction. The set of RTP session usages described is not meant to be exhaustive.

8.3.1. SRC Using Multiple m-lines

When using multiple m-lines, an SRC includes each m-line in an SDP offer to the SRS. The SDP answer from the SRS MUST include all m-lines, with any rejected m-lines indicated with a zero port, per [RFC3264]. Having received the answer, the SRC starts sending media to the SRS as indicated in the answer. Alternatively, if the SRC deems the level of support indicated in the answer to be unacceptable, it may initiate another SDP offer/answer exchange in which an alternative RTP session usage is negotiated.

In order to preserve the mapping of media to participant within the CSs in the RS, the SRC SHOULD map each unique CNAME within the CSs to a unique CNAME within the RS. Additionally, the SRC SHOULD map each unique combination of CNAME/SSRC within the CSs to a unique CNAME/SSRC within the RS. In doing so, the SRC may act as an RTP translator or as an RTP endpoint.

Figure 10 illustrates a case in which each UA represents a participant contributing two RTP sessions (e.g., one for audio and one for video), each with a single SSRC. The SRC acts as an RTP translator and delivers the media to the SRS using four RTP sessions, each with a single SSRC. The CNAME and SSRC values used by the UAs within their media streams are preserved in the media streams from the SRC to the SRS.

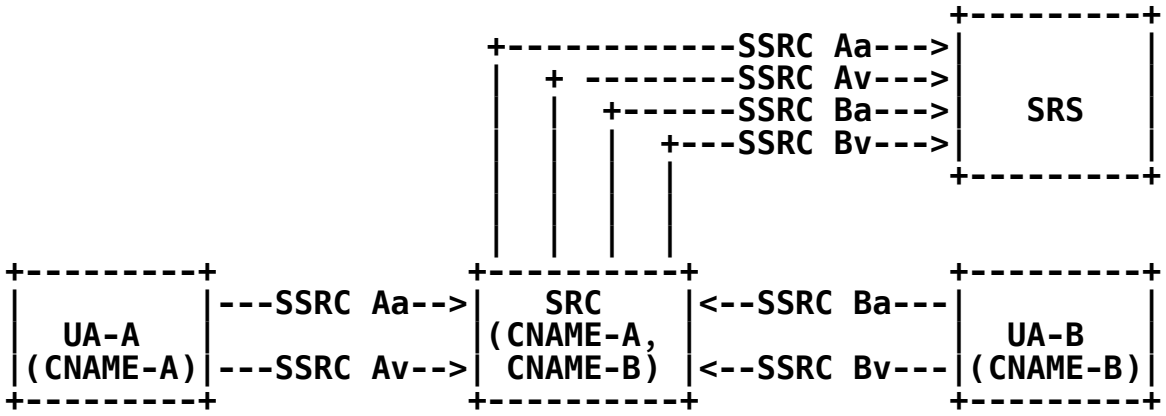

Figure 10: SRC Using Multiple m-lines

8.3.2. SRC Using Mixing

When using mixing, the SRC combines RTP streams from different participants and sends them towards the SRS using its own SSRC. The SSRCs from the contributing participants SHOULD be conveyed as CSRC identifiers. The SRC includes one m-line for each RTP session in an SDP offer to the SRS. The SDP answer from the SRS MUST include all m-lines, with any rejected m-lines indicated with a zero port, per [RFC3264]. Having received the answer, the SRC starts sending media to the SRS as indicated in the answer.

In order to preserve the mapping of media to participant within the CSs in the RS, the SRC SHOULD map each unique CNAME within the CSs to a unique CNAME within the RS. Additionally, the SRC SHOULD map each unique combination of CNAME/SSRC within the CSs to a unique

CNAME/SSRC within the RS. The SRC MUST avoid SSRC collisions, rewriting SSRCs if necessary when used as CSRCs in the RS. In doing so, the SRC acts as an RTP mixer.

In the event that the SRS does not support this usage of CSRC values, it relies entirely on the SIPREC metadata to determine the participants included within each mixed stream.

Figure 11 illustrates a case in which each UA represents a participant contributing two RTP sessions (e.g., one for audio and one for video), each with a single SSRC. The SRC acts as an RTP mixer and delivers the media to the SRS using two RTP sessions, mixing media from each participant into a single RTP session containing a single SSRC and two CSRCs.

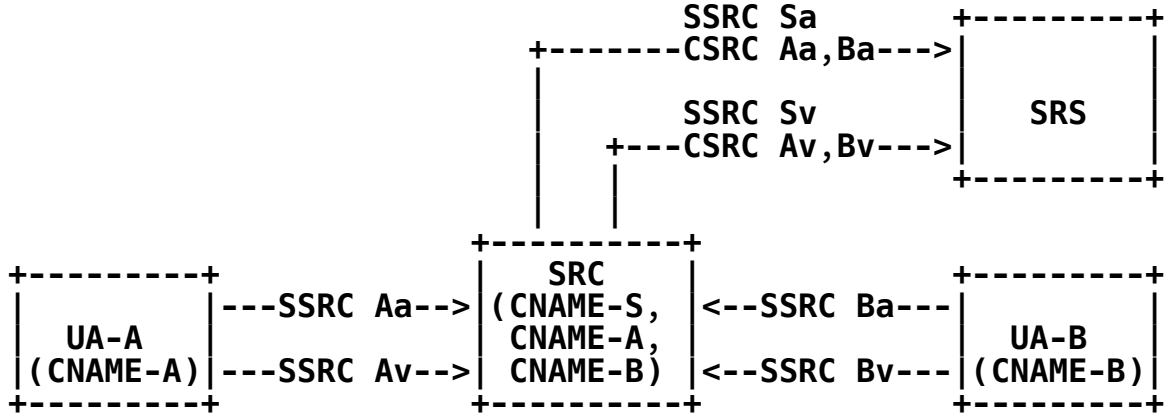

Figure 11: SRC Using Mixing

8.4. RTP Session Usage by SRS

An SRS that supports recording an audio CS MUST support SRC usage of separate audio m-lines in SDP, one per CS media direction. An SRS that supports recording a video CS MUST support SRC usage of separate video m-lines in SDP, one per CS media direction. Therefore, for an SRS supporting a typical audio call, the SRS has to support receiving at least two audio m-lines. For an SRS supporting a typical audio and video call, the SRS has to support receiving at least four total m-lines in the SDP -- two audio m-lines and two video m-lines.

These requirements allow an SRS to be implemented that supports video only, without requiring support for audio recording. They also allow an SRS to be implemented that supports recording only one direction of one stream in a CS -- for example, an SRS designed to record security monitoring cameras that only send (not receive) video without any audio. These requirements were not written to prevent

other modes from being implemented and used, such as using a single m-line and mixing the separate audio streams together. Rather, the requirements were written to provide a common base mode to implement for the sake of interoperability. It is important to note that an SRS implementation supporting the common base mode may not record all media streams in a CS if a participant supports more than one m-line in a video call, such as one for camera and one for presentation. SRS implementations may support other modes as well, but they have to at least support the modes discussed above, such that they interoperate in the common base mode for basic interoperability.

9. Metadata

Some metadata attributes are contained in SDP, and others are contained in a new content type called "application/rs-metadata". The format of the metadata is described as part of the mechanism in [RFC7865]. A new "disposition-type" of Content-Disposition is defined for the purpose of carrying metadata. The value is "recording-session", which indicates that the "application/rs-metadata" content contains metadata to be handled by the SRS.

9.1. Procedures at the SRC

The SRC MUST send metadata to the SRS in an RS. The SRC SHOULD send metadata as soon as it becomes available and whenever it changes. Cases in which an SRC may be justified in waiting temporarily before sending metadata include:

- waiting for a previous metadata exchange to complete (i.e., the SRC cannot send another SDP offer until the previous offer/answer completes and may also prefer not to send an UPDATE during this time).
- constraining the signaling rate on the RS.
- sending metadata when key events occur, rather than for every event that has any impact on metadata.

The SRC may also be configured to suppress certain metadata out of concern for privacy or perceived lack of need for it to be included in the recording.

Metadata sent by the SRC is categorized as either a full metadata snapshot or a partial update. A full metadata snapshot describes all metadata associated with the RS. The SRC MAY send a full metadata snapshot at any time. The SRC MAY send a partial update only if a full metadata snapshot has been sent previously.

The SRC MAY send metadata (either a full metadata snapshot or a partial update) in an INVITE request, an UPDATE request [RFC3311], or a 200 response to an offerless INVITE from the SRS. If the metadata contains a reference to any SDP labels, the request containing the metadata MUST also contain an SDP offer that defines those labels.

When a SIP message contains both an SDP offer and metadata, the request body MUST have content type "multipart/mixed", with one subordinate body part containing the SDP offer and another containing the metadata. When a SIP message contains only an SDP offer or metadata, the "multipart/mixed" container is optional.

The SRC SHOULD include a full metadata snapshot in the initial INVITE request establishing the RS. If metadata is not yet available (e.g., an RS established in the absence of a CS), the SRC SHOULD send a full metadata snapshot as soon as metadata becomes available.

If the SRC receives a snapshot request from the SRS, it MUST immediately send a full metadata snapshot.

Figure 12 illustrates an example of a full metadata snapshot sent by the SRC in the initial INVITE request:

```
INVITE sip:recorder@example.com SIP/2.0
Via: SIP/2.0/TCP src.example.com;branch=z9hG4bKdf6b622b648d9
From: <sip:2000@example.com>;tag=35e195d2-947d-4585-946f-09839247
To: <sip:recorder@example.com>
Call-ID: d253c800-b0d1ea39-4a7dd-3f0e20a
CSeq: 101 INVITE
Max-Forwards: 70
Require: siprec
Accept: application/sdp, application/rs-metadata
Contact: <sip:2000@src.example.com>;+sip.src
Content-Type: multipart/mixed;boundary=foobar
Content-Length: [length]

--foobar
Content-Type: application/sdp

v=0
o=SRS 2890844526 2890844526 IN IP4 198.51.100.1
s=-
c=IN IP4 198.51.100.1
t=0 0
m=audio 12240 RTP/AVP 0 4 8
a=sendonly
a=label:1

--foobar
Content-Type: application/rs-metadata
Content-Disposition: recording-session

[metadata content]
```

Figure 12: Sample INVITE Request for the Recording Session

9.2. Procedures at the SRS

The SRS receives metadata updates from the SRC in INVITE and UPDATE requests. Since the SRC can send partial updates based on the previous update, the SRS needs to keep track of the sequence of updates from the SRC.

In the case of an internal failure at the SRS, the SRS may fail to recognize a partial update from the SRC. The SRS may be able to recover from the internal failure by requesting a full metadata snapshot from the SRC. Certain errors, such as syntax errors or semantic errors in the metadata information, are likely caused by an

error on the SRC side, and it is likely that the same error will occur again even when a full metadata snapshot is requested. In order to avoid repeating the same error, the SRS can simply terminate the RS when a syntax error or semantic error is detected in the metadata.

The SRS MAY explicitly request a full metadata snapshot by sending an UPDATE request. This request MUST contain a body with Content-Disposition type "recording-session" and MUST NOT contain an SDP body. The SRS MUST NOT request a full metadata snapshot in an UPDATE response or in any other SIP transaction. The format of the content is "application/rs-metadata", and the body is an XML document, the format of which is defined in [RFC7865]. Figure 13 shows an example:

```
UPDATE sip:2000@src.example.com SIP/2.0
Via: SIP/2.0/UDP srs.example.com;branch=z9hG4bKdf6b622b648d9
To: <sip:2000@example.com>;tag=35e195d2-947d-4585-946f-098392474
From: <sip:recorder@example.com>;tag=1234567890
Call-ID: d253c800-b0d1ea39-4a7dd-3f0e20a
CSeq: 1 UPDATE
Max-Forwards: 70
Require: siprec
Contact: <sip:recorder@srs.example.com>;+sip.srs
Accept: application/sdp, application/rs-metadata
Content-Disposition: recording-session
Content-Type: application/rs-metadata
Content-Length: [length]

<?xml version="1.0" encoding="UTF-8"?>
  <requestsnapshot xmlns='urn:ietf:params:xml:ns:recording:1'>
    <requestreason xml:lang="it">SRS internal error</requestreason>
  </requestsnapshot>
```

Figure 13: Metadata Request

Note that UPDATE was chosen for the SRS to request a metadata snapshot, because it can be sent regardless of the state of the dialog. This was seen as better than requiring support for both UPDATE and re-INVITE messages for this operation.

When the SRC receives a request for a metadata snapshot, it MUST immediately provide a full metadata snapshot in a separate INVITE or UPDATE transaction. Any subsequent partial updates will not be dependent on any metadata sent prior to this full metadata snapshot.

The metadata received by the SRS can contain ID elements used to cross-reference one element to another. An element containing the definition of an ID and an element containing a reference to that ID will often be received from the same SRC. It is also valid for those elements to be received from different SRCs -- for example, when each endpoint in the same CS acts as an SRC to record the call and a common ID refers to the same CS. The SRS MUST NOT consider this an error.

10. Persistent Recording

Persistent recording is a specific use case addressing REQ-005 in [RFC6341], where an RS can be established in the absence of a CS. The SRC continuously records media in an RS to the SRS even in the absence of a CS for all UAs that are part of persistent recording. By allocating recorded streams and continuously sending recorded media to the SRS, the SRC does not have to prepare new recorded streams with a new SDP offer when a new CS is created and also does not impact the timing of the CS. The SRC only needs to update the metadata when new CSs are created.

When there is no CS running on the devices with persistent recording, there is no recorded media to stream from the SRC to the SRS. In certain environments where a Network Address Translator (NAT) is used, a minimum amount of flow activity is typically required to maintain the NAT binding for each port opened. Agents that support Interactive Connectivity Establishment (ICE) solve this problem. For non-ICE agents, in order not to lose the NAT bindings for the RTP/RTCP ports opened for the recorded streams, the SRC and SRS SHOULD follow the recommendations provided in [RFC6263] to maintain the NAT bindings.

11. IANA Considerations

11.1. Registration of Option Tags

This specification registers two option tags. The required information for this registration, as specified in [RFC3261], is as follows.

11.1.1. "siprec" Option Tag

Name: siprec

Description: This option tag is for identifying that the SIP session is for the purpose of an RS. This is typically not used in a Supported header. When present in a Require header in a request, it indicates that the UA is either an SRC or SRS capable of handling an RS.

11.1.2. "record-aware" Option Tag

Name: record-aware

Description: This option tag is to indicate the ability of the UA to receive recording indicators in media-level or session-level SDP. When present in a Supported header, it indicates that the UA can receive recording indicators in media-level or session-level SDP.

11.2. Registration of Media Feature Tags

This document registers two new media feature tags in the SIP tree per the process defined in [RFC2506] and [RFC3840].

11.2.1. Feature Tag for the SRC

Media feature tag name: sip.src

ASN.1 Identifier: 1.3.6.1.8.4.27

Summary of the media feature indicated by this tag: This feature tag indicates that the UA is a Session Recording Client for the purpose of an RS.

Values appropriate for use with this feature tag: boolean

The feature tag is intended primarily for use in the following applications, protocols, services, or negotiation mechanisms: This feature tag is only useful for an RS.

Examples of typical use: Routing the request to a Session Recording Server.

Security Considerations: Security considerations for this media feature tag are discussed in Section 11.1 of RFC 3840.

11.2.2. Feature Tag for the SRS

Media feature tag name: sip.srs

ASN.1 Identifier: 1.3.6.1.8.4.28

Summary of the media feature indicated by this tag: This feature tag indicates that the UA is a Session Recording Server for the purpose of an RS.

Values appropriate for use with this feature tag: boolean

The feature tag is intended primarily for use in the following applications, protocols, services, or negotiation mechanisms: This feature tag is only useful for an RS.

Examples of typical use: Routing the request to a Session Recording Client.

Security Considerations: Security considerations for this media feature tag are discussed in Section 11.1 of RFC 3840.

11.3. New Content-Disposition Parameter Registrations

This document registers a new "disposition-type" value in the Content-Disposition header: recording-session.

recording-session: The body describes either

* metadata about the RS

  or

* the reason for the metadata snapshot request

  as determined by the MIME value indicated in the Content-Type.

11.4. SDP Attributes

This document registers the following new SDP attributes.

11.4.1. "record" SDP Attribute

Contact names:
Leon Portman, leon.portman@nice.com;
Henry Lum, henry.lum@genesyslab.com

Attribute name: record

Long-form attribute name: Recording Indication

Type of attribute: session level or media level

Subject to charset: no

This attribute provides the recording indication for the session or media stream.

Allowed attribute values: on, off, paused

11.4.2. "recordpref" SDP Attribute

Contact names:
Leon Portman, leon.portman@nice.com;
Henry Lum, henry.lum@genesyslab.com

Attribute name: recordpref

Long-form attribute name: Recording Preference

Type of attribute: session level or media level

Subject to charset: no

This attribute provides the recording preference for the session or media stream.

Allowed attribute values: on, off, pause, nopreference

12. Security Considerations

The RS is fundamentally a standard SIP dialog [RFC3261]; therefore, the RS can reuse any of the existing SIP security mechanisms available for securing the session signaling, the recorded media, and the metadata. The use cases and requirements document [RFC6341] outlines the general security considerations, and this document describes specific security recommendations.

The SRC and SRS MUST support SIP with Transport Layer Security (TLS) version 1.2, SHOULD follow the best practices when using TLS as per [RFC7525], and MAY use Session Initiation Protocol Secure (SIPS) with TLS as per [RFC5630]. The RS MUST be at least as secure as the CS; this means using at least the same strength of cipher suite as the CS if the CS is secured. For example, if the CS uses SIPS for signaling and RTP/SAVP for media, then the RS may not use SIP or plain RTP unless other equivalent security measures are in effect, since doing so would mean an effective security downgrade. Examples of other potentially equivalent security mechanisms include mutually authenticated TLS for the RS signaling channel or an appropriately protected network path for the RS media component.

12.1. Authentication and Authorization

At the transport level, the RS uses TLS authentication to validate the authenticity of the SRC and SRS. The SRC and SRS MUST implement TLS mutual authentication for establishing the RS. Whether the SRC/SRS chooses to use TLS mutual authentication is a deployment decision. In deployments where a UA acts as its own SRC, this requires that the UA have its own certificate as needed for TLS mutual authentication. In deployments where the SRC and the SRS are in the same administrative domain and have some other means of assuring authenticity, the SRC and SRS may choose not to authenticate each other or to have the SRC authenticate the SRS only. In deployments where the SRS can be hosted on a different administrative domain, it is important to perform mutual authentication to ensure the authenticity of both the SRC and the SRS before transmitting any recorded media. The risk of not authenticating the SRS is that the recording may be sent to an entity other than the intended SRS, allowing a sensitive call recording to be received by an attacker. On the other hand, the risk of not authenticating the SRC is that an SRS will accept calls from an unknown SRC and allow potential forgery of call recordings.

There may be scenarios in which the signaling between the SRC and SRS is not direct, e.g., a SIP proxy exists between the SRC and the SRS. In such scenarios, each hop is subject to the TLS mutual authentication constraint, and transitive trust at each hop is

utilized. Additionally, an SRC or SRS may use other existing SIP mechanisms available, including, but not limited to, Digest authentication [RFC3261], asserted identity [RFC3325], and connected identity [RFC4916].

The SRS may have its own set of recording policies to authorize recording requests from the SRC. The use of recording policies is outside the scope of the Session Recording Protocol.

12.2. RTP Handling

In many scenarios, it will be critical for the media transported between the SRC and the SRS to be protected. Media encryption is an important element in the overall SIPREC solution; therefore, the SRC and the SRS MUST support RTP/SAVP [RFC3711] and RTP/SAVPF [RFC5124]. RTP/SAVP and RTP/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. At a minimum, the SRC and SRS MUST support the SDP security descriptions key negotiation mechanism [RFC4568]. For cases in which Datagram Transport Layer Security for Secure RTP (DTLS-SRTP) is used to encrypt a CS media stream, an SRC may use SRTP Encrypted Key Transport (EKT) [EKT-SRTP] in order to use SRTP-SDES in the RS without needing to re-encrypt the media.

> Note: When using EKT in this manner, it is possible for participants in the CS to send traffic that appears to be from other participants and have this forwarded by the SRC to the SRS within the RS. If this is a concern (e.g., the RS is intended for audit or compliance purposes), EKT is not an appropriate choice.

When RTP/SAVP or RTP/SAVPF is used, an SRC can choose to use the same keys or different keys in the RS than those used in the CS. Some SRCs are designed to simply replicate RTP packets from a CS media stream to the SRS, in which case the SRC will use the same key in the RS as the key used in the CS. In this case, the SRC MUST secure the SDP containing the keying material in the RS with at least the same level of security as in the CS. The risk of lowering the level of security in the RS is that it will effectively become a downgrade attack on the CS, since the same key is used for both the CS and the RS.

SRCs that decrypt an encrypted CS media stream and re-encrypt it when sending it to the SRS MUST use a different key than what is used for the CS media stream, to ensure that it is not possible for someone who has the key for the CS media stream to access recorded data they

are not authorized to access. In order to maintain a comparable level of security, the key used in the RS SHOULD be of equivalent strength to, or greater strength than, that used in the CS.

12.3. Metadata

Metadata contains sensitive information, such as the address of record of the participants and other extension data placed by the SRC. It is essential to protect the content of the metadata in the RS. Since metadata is a content type transmitted in SIP signaling, metadata SHOULD be protected at the transport level by SIPS/TLS.

12.4. Storage and Playback

While storage and playback of the call recording are beyond the scope of this document, it is worthwhile to mention here that it is also important for the recording storage and playback to provide a level of security that is comparable to the CS. It would defeat the purpose of securing both the CS and the RS mentioned in the previous sections if the recording can be easily played back with a simple, unsecured HTTP interface without any form of authentication or authorization.

13. References

13.1. 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>.

[RFC2506] Holtman, K., Mutz, A., and T. Hardie, "Media Feature Tag Registration Procedure", BCP 31, RFC 2506, DOI 10.17487/RFC2506, March 1999, <http://www.rfc-editor.org/info/rfc2506>.

[RFC3261] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., and E. Schooler, "SIP: Session Initiation Protocol", RFC 3261, DOI 10.17487/RFC3261, June 2002, <http://www.rfc-editor.org/info/rfc3261>.

[RFC3264] Rosenberg, J. and H. Schulzrinne, "An Offer/Answer Model with Session Description Protocol (SDP)", RFC 3264, DOI 10.17487/RFC3264, June 2002, <http://www.rfc-editor.org/info/rfc3264>.

[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>.

[RFC3840] Rosenberg, J., Schulzrinne, H., and P. Kyzivat, "Indicating User Agent Capabilities in the Session Initiation Protocol (SIP)", RFC 3840, DOI 10.17487/RFC3840, August 2004, <http://www.rfc-editor.org/info/rfc3840>.

[RFC4574] Levin, O. and G. Camarillo, "The Session Description Protocol (SDP) Label Attribute", RFC 4574, DOI 10.17487/RFC4574, August 2006, <http://www.rfc-editor.org/info/rfc4574>.

[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>.

[RFC7245] Hutton, A., Ed., Portman, L., Ed., Jain, R., and K. Rehor, "An Architecture for Media Recording Using the Session Initiation Protocol", RFC 7245, DOI 10.17487/RFC7245, May 2014, <http://www.rfc-editor.org/info/rfc7245>.

[RFC7865] Ravindranath, R., Ravindran, P., and P. Kyzivat, "Session Initiation Protocol (SIP) Recording Metadata", RFC 7865, DOI 10.17487/RFC7865, May 2016, <http://www.rfc-editor.org/info/rfc7865>.

13.2. Informative References

[EKT-SRTP] Mattsson, J., Ed., McGrew, D., Wing, D., and F. Andreasen, "Encrypted Key Transport for Secure RTP", Work in Progress, draft-ietf-avtcore-srtp-ekt-03, October 2014.

[RFC2804] IAB and IESG, "IETF Policy on Wiretapping", RFC 2804, DOI 10.17487/RFC2804, May 2000, <http://www.rfc-editor.org/info/rfc2804>.

[RFC3311] Rosenberg, J., "The Session Initiation Protocol (SIP) UPDATE Method", RFC 3311, DOI 10.17487/RFC3311, October 2002, <http://www.rfc-editor.org/info/rfc3311>.

[RFC3325] Jennings, C., Peterson, J., and M. Watson, "Private Extensions to the Session Initiation Protocol (SIP) for Asserted Identity within Trusted Networks", RFC 3325, DOI 10.17487/RFC3325, November 2002, <http://www.rfc-editor.org/info/rfc3325>.

[RFC3551] Schulzrinne, H. and S. Casner, "RTP Profile for Audio and Video Conferences with Minimal Control", STD 65, RFC 3551, DOI 10.17487/RFC3551, July 2003, <http://www.rfc-editor.org/info/rfc3551>.

[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>.

[RFC4568] Andreasen, F., Baugher, M., and D. Wing, "Session Description Protocol (SDP) Security Descriptions for Media Streams", RFC 4568, DOI 10.17487/RFC4568, July 2006, <http://www.rfc-editor.org/info/rfc4568>.

[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, <http://www.rfc-editor.org/info/rfc4585>.

[RFC4916] Elwell, J., "Connected Identity in the Session Initiation Protocol (SIP)", RFC 4916, DOI 10.17487/RFC4916, June 2007, <http://www.rfc-editor.org/info/rfc4916>.

[RFC4961] Wing, D., "Symmetric RTP / RTP Control Protocol (RTCP)", BCP 131, RFC 4961, DOI 10.17487/RFC4961, July 2007, <http://www.rfc-editor.org/info/rfc4961>.

[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, <http://www.rfc-editor.org/info/rfc5104>.

[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>.

[RFC5168] Levin, O., Even, R., and P. Hagendorf, "XML Schema for Media Control", RFC 5168, DOI 10.17487/RFC5168, March 2008, <http://www.rfc-editor.org/info/rfc5168>.

[RFC5630] Audet, F., "The Use of the SIPS URI Scheme in the Session Initiation Protocol (SIP)", RFC 5630, DOI 10.17487/RFC5630, October 2009, <http://www.rfc-editor.org/info/rfc5630>.

[RFC5761] Perkins, C. and M. Westerlund, "Multiplexing RTP Data and Control Packets on a Single Port", RFC 5761, DOI 10.17487/RFC5761, April 2010, <http://www.rfc-editor.org/info/rfc5761>.

[RFC6263] Marjou, X. and A. Sollaud, "Application Mechanism for Keeping Alive the NAT Mappings Associated with RTP / RTP Control Protocol (RTCP) Flows", RFC 6263, DOI 10.17487/RFC6263, June 2011, <http://www.rfc-editor.org/info/rfc6263>.

[RFC6341] Rehor, K., Ed., Portman, L., Ed., Hutton, A., and R. Jain, "Use Cases and Requirements for SIP-Based Media Recording (SIPREC)", RFC 6341, DOI 10.17487/RFC6341, August 2011, <http://www.rfc-editor.org/info/rfc6341>.

[RFC7022] Begen, A., Perkins, C., Wing, D., and E. Rescorla, "Guidelines for Choosing RTP Control Protocol (RTCP) Canonical Names (CNAMEs)", RFC 7022, DOI 10.17487/RFC7022, September 2013, <http://www.rfc-editor.org/info/rfc7022>.

[RFC7525] Sheffer, Y., Holz, R., and P. Saint-Andre, "Recommendations for Secure Use of Transport Layer Security (TLS) and Datagram Transport Layer Security (DTLS)", BCP 195, RFC 7525, DOI 10.17487/RFC7525, May 2015, <http://www.rfc-editor.org/info/rfc7525>.

Acknowledgements

We want to thank John Elwell, Paul Kyzivat, Partharsarathi R, Ram Mohan R, Hadriel Kaplan, Adam Roach, Miguel Garcia, Thomas Stach, Muthu Perumal, Dan Wing, and Magnus Westerlund for their valuable comments and inputs to this document.

Authors' Addresses

Leon Portman
NICE Systems
22 Zarhin Street
P.O. Box 690
Ra'anana 4310602
Israel

Email: leon.portman@gmail.com

Henry Lum (editor)
Genesys
1380 Rodick Road, Suite 201
Markham, Ontario L3R4G5
Canada

Email: henry.lum@genesyslab.com

Charles Eckel
Cisco
170 West Tasman Drive
San Jose, CA 95134
United States

Email: eckelcu@cisco.com

Alan Johnston
Illinois Institute of Technology
Bellevue, WA
United States

Email: alan.b.johnston@gmail.com

Andrew Hutton
Unify
Brickhill Street
Milton Keynes MK15 0DJ
United Kingdom

Email: andrew.hutton@unify.com